ARULMIGU PALANI ANDAVAR ARTS COLLEGE FOR WOMEN, PALANI

PG DEPARTMENT OF BOTANY

LEARNING RESOURCES

GENETICS AND MOLECULAR BIOLOGY

Gregor Johann Mendel

Gregor Johann Mendel 20 July 1822 -- 6 January 1884) was a German-Czech biologist, meteorologist, mathematician, Augustinian friar and abbot of St.Thomas' Abbey in Brno (*Brünn*), Margraviate of Moravia. Mendel was born in a German-speaking family in the Silesian part of the Austrian Empire (today's Czech Republic) and gained posthumous recognition as the founder of the modern science of genetics. Though farmers had known for millennia that crossbreeding of animals and plants could favor certain desirable traits, Mendel's pea plant experiments conducted between 1856 and 1863 established many of the rules of heredity, now referred to as the laws of Mendelian inheritance.

Mendel worked with seven characteristics of pea plants: plant height, pod shape and color, seed shape and color, and flower position and color. Taking seed color as an example, Mendel showed that when a true-breeding yellow pea and a true-breeding green pea were cross-bred their offspring always produced yellow seeds. However, in the next generation, the green peas reappeared at a ratio of 1 green to 3 yellow. To explain this phenomenon, Mendel coined the terms "recessive" and "dominant" in reference to certain traits. In the preceding example, the green trait, which seems to have vanished in the first filial generation, is recessive and the yellow is dominant. He published his work in 1866, demonstrating the actions of invisible "factors"—now called genes—in predictably determining the traits of an organism.

The profound significance of Mendel's work was not recognized until the turn of the 20th century (more than three decades later) with the rediscovery of his laws. Erich von Tschermak, Hugo de Vries and Carl Correns independently verified several of Mendel's experimental findings in 1900, ushering in the modern age of genetics.

Early Life and Education

Mendel was born into a German-speaking family in Heinzendorf bei Odrau (called *Hynčice* in Czech), at the Moravian-Silesian border, Austrian Empire (now Czechia). He was the son of Anton and Rosine (Schwirtlich) Mendel and had one older sister, Veronika, and one younger, Theresia. They lived and worked on a farm which had been owned by the Mendel family for at least 130 years (the house where Mendel was born is now a museum devoted to Mendel). During

his childhood, Mendel worked as a gardener and studied beekeeping. As a young man, he attended gymnasium in Troppau (Czech: *Opava*). He had to take four months off during his gymnasium studies due to illness. From 1840 to 1843, he studied practical and theoretical philosophy and physics at the Philosophical Institute of the University of Olomouc (German: *Olmütz*), taking another year off because of illness. He also struggled financially to pay for his studies, and Theresia gave him her dowry. Later he helped support her three sons, two of whom became doctors.

He became a monk in part because it enabled him to obtain an education without having to pay for it himself. As the son of a struggling farmer, the monastic life, in his words, spared him the "perpetual anxiety about a means of livelihood." Born Johann Mendel, he was given the name **Gregor** (*Řehoř* in Czech) when he joined the Order of Saint Augustine.

Gregor Johann Mendel

Academic career

When Mendel entered the Faculty of Philosophy, the Department of Natural History and Agriculture was headed by Johann Karl Nestler who conducted extensive research of hereditary traits of plants and animals, especially sheep. Upon recommendation of his physics teacher Friedrich Franz, Mendel entered the Augustinian St Thomas's Abbey in Brno and began his training as a priest. Mendel worked as a substitute high school teacher. In 1850, he failed the oral part, the last of three parts, of his exams to become a certified high school teacher. In 1851, he was sent to the University of Vienna to study under the sponsorship of Abbot Cyril František Napp so that he could get more formal education. At Vienna, his professor of physics was Christian Doppler. Mendel returned to his abbey in 1853 as a teacher, principally of physics. In 1854 he met Aleksander Zawadzki who encouraged his research in Brno. In 1856, he took the exam to become a certified teacher and again failed the oral part. In 1867, he replaced Napp as abbot of the monastery.

After he was elevated as abbot in 1868, his scientific work largely ended, as Mendel became overburdened with administrative responsibilities, especially a dispute with the civil government over its attempt to impose special taxes on religious institutions. Mendel died on 6 January 1884, at the age of 61, in Brno, from chronic nephritis. Czech composer Leoš Janáček played the organ at his funeral. After his death, the succeeding abbot burned all papers in Mendel's collection, to mark an end to the disputes over taxation. The exhumation of Mendel's corpse in 2021 delivered some physiognomic details like body height (168 cm (66 in). His genome was analysed, revealing that Mendel was predisposed to heart problems.

Experiments on plant hybridization

Dominant and recessive phenotypes. (1) Parental generation. (2) F1 generation. (3) F2 generation.

Mendel, known as the "father of modern genetics", chose to study variation in plants in his monastery's 2 hectares (4.9 acres) experimental garden. Mendel was assisted in his experimental design by Aleksander Zawadzki while his superior abbot Napp wrote to discourage him, saying that the Bishop giggled when informed of the detailed genealogies of peas.

After initial experiments with pea plants, Mendel settled on studying seven traits that seemed to be inherited independently of other traits: seed shape, flower color, seed coat tint, pod shape, unripe pod color, flower location, and plant height. He first focused on seed shape, which was either angular or round. Between 1856 and 1863 Mendel cultivated and tested some 28,000 plants, the majority of which were pea plants (*Pisum sativum*). This study showed that, when true-breeding different varieties were crossed to each other (e.g., tall plants fertilized by short plants), in the second generation, one in four pea plants had purebred recessive traits, two out of four were hybrids, and one out of four were purebred dominant. His experiments led him to make two generalizations, the Law of Segregation and the Law of Independent Assortment, which later came to be known as Mendel's Laws of Inheritance.

Initial reception of Mendel's work

Mendel presented his paper, *Versuche über Pflanzenhybriden* ("Experiments on Plant Hybridization"), at two meetings of the Natural History Society of Brno in Moravia on 8 February and 8 March 1865. It generated a few favorable reports in local newspapers, but was ignored by the scientific community. When Mendel's paper was published in 1866 in *Verhandlungen des naturforschenden Vereines in Brünn*, it was seen as essentially about hybridization rather than inheritance, had little impact, and was cited only about three times over the next thirty-five years. His paper was criticized at the time, but is now considered a seminal work. Notably, Charles Darwin was not aware of Mendel's paper, and it is envisaged that if he had been aware of it, genetics as it exists now might have taken hold much earlier. Mendel's scientific biography thus provides an example of the failure of obscure, highly original innovators to receive the attention they deserve.

Rediscovery of Mendel's work

About forty scientists listened to Mendel's two groundbreaking lectures, but it would appear that they failed to understand the implications of his work. Later, he also carried on a correspondence with Carl Nägeli, one of the leading biologists of the time, but Nägeli too failed to appreciate Mendel's discoveries. At times, Mendel must have entertained doubts about his work, but not always: "My time will come," he reportedly told a friend,Gustav von Niessl.

During Mendel's lifetime, most biologists held the idea that all characteristics were passed to the next generation through blending inheritance (indeed, many effectively are), in which the traits

from each parent are average .Instances of this phenomenon are now explained by the action of multiple genes with quantitative effects. Charles Darwin tried unsuccessfully to explain inheritance through a theory of pangenesis. It was not until the early 20th century that the importance of Mendel's ideas was realized.

By 1900, research aimed at finding a successful theory of discontinuous inheritance rather than blending inheritance led to independent duplication of his work by Hugo de Vries and Carl Correns, and the rediscovery of Mendel's writings and laws. Both acknowledged Mendel's priority, and it is thought probable that de Vries did not understand the results he had found until after reading Mendel. Though Erich von Tschermak was originally also credited with rediscovery, this is no longer accepted because he did not understand Mendel's laws. Though de Vries later lost interest in Mendelism, other biologists started to establish modern genetics as a science. All three of these researchers, each from a different country, published their rediscovery of Mendel's work within a two-month span in the spring of 1900.

Mendel's results were quickly replicated, and genetic linkage quickly worked out. Biologists flocked to the theory; even though it was not yet applicable to many phenomena, it sought to give a genotypic understanding of heredity which they felt was lacking in previous studies of heredity, which had focused on phenotypic approaches. Most prominent of these previous approaches was the biometric school of Karl Pearson and W. F. R. Weldon, which was based heavily on statistical studies of phenotype variation. The strongest opposition to this school came from William Bateson, who perhaps did the most in the early days of publicising the benefits of Mendel's theory (the word "genetics", and much of the discipline's other terminology, originated with Bateson). This debate between the biometricians and the Mendelians was extremely vigorous in the first two decades of the 20th century, with the biometricians claiming statistical and mathematical rigor, whereas the Mendelians claimed a better understanding of biology. Modern genetics shows that Mendelian heredity is in fact an inherently biological process, though not all genes of Mendel's experiments are yet understood.

In the end, the two approaches were combined, especially by work conducted by R. A. Fisher as early as 1918. The combination, in the 1930s and 1940s, of Mendelian genetics with Darwin's theory of natural selection resulted in the modern synthesis of evolutionary biology.

In the Soviet Union and China, Mendelian genetics was rejected in favor of Lamarckism, leading to imprisonment and even execution of Mendelian geneticists (see Lysenkoism).

Other experiments

Mendel began his studies on heredity using mice. He was at St. Thomas's Abbey but his bishop did not like one of his friars studying animal sex, so Mendel switched to plants. Mendel also bred bees in a bee house that was built for him, using bee hives that he designed. He also studied astronomy and meteorology, founding the 'Austrian Meteorological Society' in 1865. The majority of his published works were related to meteorology.

Mendel also experimented with hawkweed (*Hieracium*) and honeybees. He published a report on his work with hawkweed a group of plants of great interest to scientists at the time because of their diversity. However, the results of Mendel's inheritance study in hawkweeds was unlike his results for peas; the first generation was very variable and many of their offspring were identical to the maternal parent. In his correspondence with Carl Nägeli he discussed his results but was unable to explain them. It was not appreciated until the end of the nineteenth century that many hawkweed species were apomictic, producing most of their seeds through an asexual process.

None of his results on bees survived, except for a passing mention in the reports of Moravian Apiculture Society. All that is known definitely is that he used Cyprian and Carniolan bees, which were particularly aggressive to the annoyance of other monks and visitors of the monastery such that he was asked to get rid of them. Mendel, on the other hand, was fond of his bees, and referred to them as "my dearest little animals".

He also described novel plant species, and these are denoted with the botanical author abbreviation "Mendel".Mendel presented the results of his experiments with nearly **30,000 pea plants** to the local natural history society. He demonstrated that traits are transmitted faithfully from parents to offspring in specific patterns. In 1866, he published his work, *Experiments in Plant Hybridization,*[1] in the proceedings of the Natural History Society of Brünn.

Mendel's work went virtually unnoticed by the scientific community, which incorrectly believed that the process of inheritance involved a **blending** of parental traits that produced an intermediate physical appearance in offspring. This hypothetical process appeared to be correct because of what we know now as continuous variation. Continuous variation is the range of

small differences we see among individuals in a characteristic like human height. It does appear that offspring are a "blend" of their parents' traits when we look at characteristics that exhibit continuous variation. Mendel worked instead with traits that show **discontinuous variation**. Discontinuous variation is the variation seen among individuals when each individual shows one of two—or a very few—easily distinguishable traits, such as violet or white flowers. Mendel's choice of these kinds of traits allowed him to see experimentally that the traits were not blended in the offspring as would have been expected at the time, but that they were inherited as distinct traits. In 1868, Mendel became abbot of the monastery and exchanged his scientific pursuits for his pastoral duties. He was not recognized for his extraordinary scientific contributions during his lifetime; in fact, it was not until 1900 that his work was rediscovered, reproduced, and revitalized by scientists on the brink of discovering the chromosomal basis of heredity.

Mendel's Crosses

Mendel's seminal work was accomplished using the **garden pea**, *Pisum sativum*, to study inheritance. This species naturally **self-fertilizes**, meaning that pollen encounters ova within the same flower. The flower petals remain sealed tightly until pollination is completed to prevent the pollination of other plants. The result is highly inbred, or "true-breeding," pea plants. These are plants that always produce offspring that look like the parent. By experimenting with true-breeding pea plants, Mendel avoided the appearance of unexpected traits in offspring that might occur if the plants were not true breeding. The garden pea also grows to maturity within one season, meaning that several generations could be evaluated over a relatively short time. Finally, large quantities of garden peas could be cultivated simultaneously, allowing Mendel to conclude that his results did not come about simply by chance.

Mendel performed hybridizations, which involve **mating two true-breeding individuals** that have different traits. In the pea, which is naturally self-pollinating, this is done by manually transferring pollen from the anther of a mature pea plant of one variety to the stigma of a separate mature pea plant of the second variety.

Plants used in first-generation crosses were called **P, or parental generation**, plants (Figure 8.3). Mendel collected the seeds produced by the P plants that resulted from each cross and grew

them the following season. These offspring were called the F_1, or the first filial (filial = daughter or son), generation. Once Mendel examined the characteristics in the **F_1 generation of plants**, he allowed them to self-fertilize naturally. He then collected and grew the seeds from the F_1 plants to produce the F_2, or second filial, generation. Mendel's experiments extended beyond the **F_2 generation** to the F_3 generation, F_4 generation, and so on, but it was the ratio of characteristics in the P, F_1, and F_2 generations that were the most intriguing and became the basis of Mendel's postulates.

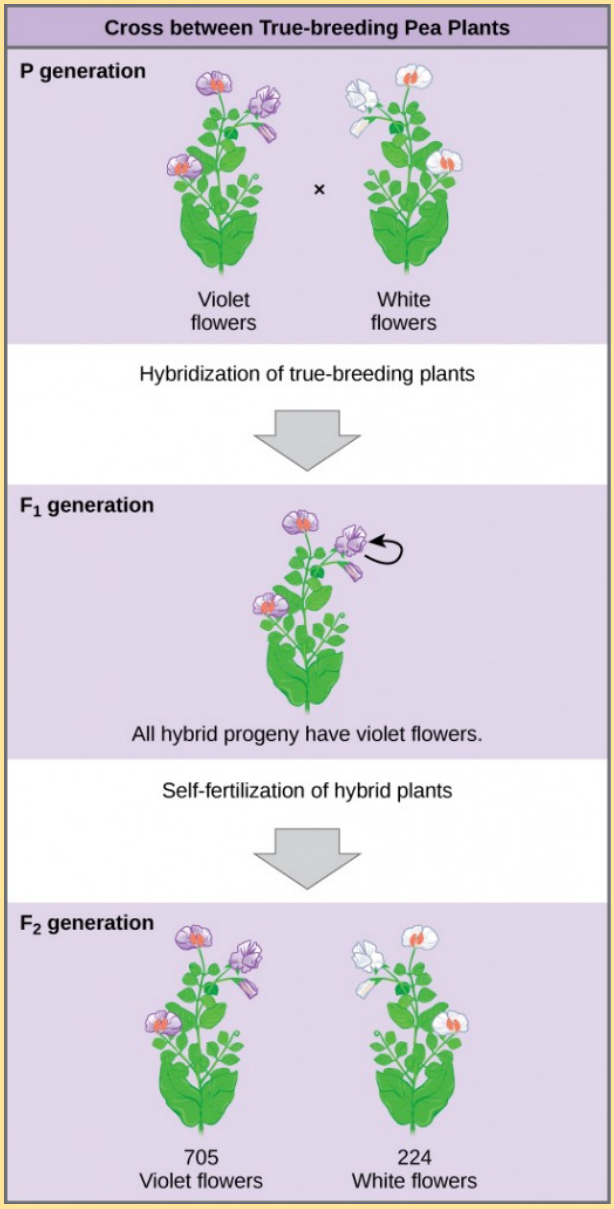


Figure 8.3 Mendel's process for performing crosses included examining flower color.

Garden Pea Characteristics Revealed the Basics of Heredity

In his 1865 publication, Mendel reported the results of his crosses involving seven different characteristics, each with two contrasting traits. A trait is defined as a variation in the physical appearance of a heritable characteristic. The characteristics included plant height, seed texture, seed color, flower color, pea-pod size, pea-pod color, and flower position. For the characteristic of flower color, for example, the two contrasting traits were white versus violet. To fully examine each characteristic, Mendel generated large numbers of F_1 and F_2 plants and reported results from thousands of F_2 plants.

What results did Mendel find in his crosses for flower color? First, Mendel confirmed that he was using plants that bred true for white or violet flower color. Irrespective of the number of generations that Mendel examined, all self-crossed offspring of parents with white flowers had white flowers, and all self-crossed offspring of parents with violet flowers had violet flowers. In addition, Mendel confirmed that, other than flower color, the pea plants were physically identical. This was an important check to make sure that the two varieties of pea plants only differed with respect to one trait, flower color.

Once these validations were complete, Mendel applied the pollen from a plant with violet flowers to the stigma of a plant with white flowers. After gathering and sowing the seeds that resulted from this cross, Mendel found that **100 percent of the F_1** hybrid generation had violet flowers. Conventional wisdom at that time would have predicted the hybrid flowers to be pale violet or for hybrid plants to have equal numbers of white and violet flowers. In other words, the contrasting parental traits were expected to blend in the offspring. Instead, Mendel's results demonstrated that the white flower trait had completely disappeared in the F_1 generation.

Importantly, Mendel did not stop his experimentation there. He allowed the F_1 plants to self-fertilize and found that 705 plants in the **F_2**generation had violet flowers and 224 had white flowers. This was a ratio of 3.15 violet flowers to one white flower, or **approximately 3:1**. When Mendel transferred pollen from a plant with violet flowers to the stigma of a plant with white flowers and vice versa, he obtained approximately the same ratio irrespective of which parent—male or female—contributed which trait. This is called a **reciprocal cross**—a paired cross in which the respective traits of the male and female in one cross become the respective

traits of the female and male in the other cross. For the other six characteristics that Mendel examined, the F_1 and F_2 generations behaved in the same way that they behaved for flower color. One of the two traits would disappear completely from the F_1 generation, only to reappear in the F_2 generation at a ratio of roughly 3:1 (Figure 8.4).

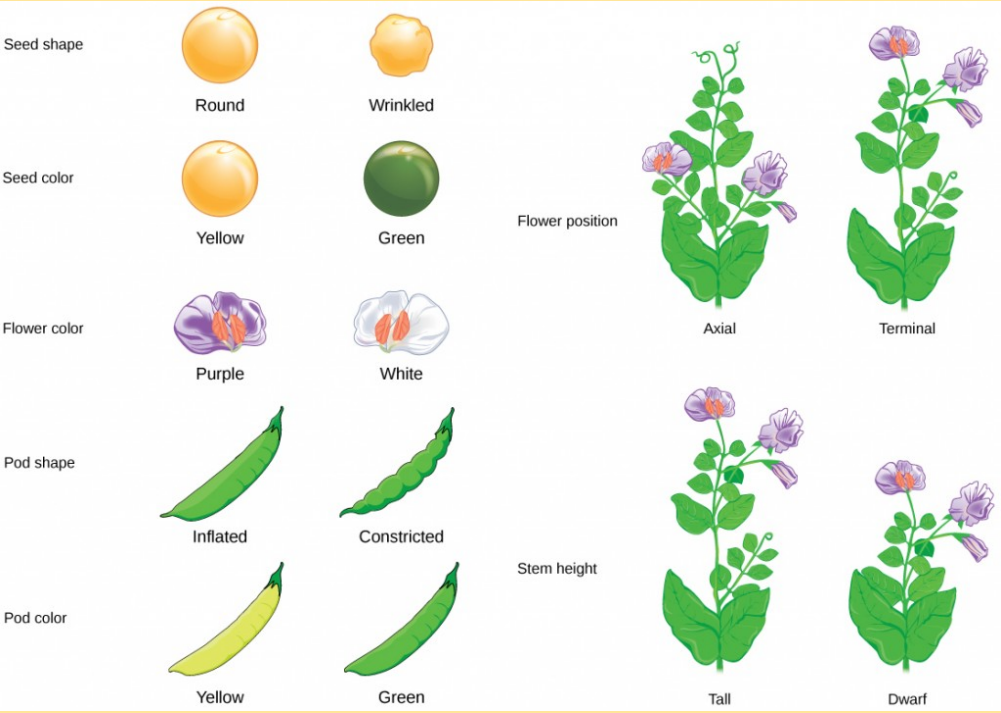


Figure 8.4 Mendel identified seven pea plant characteristics.

Upon compiling his results for many thousands of plants, Mendel concluded that the characteristics could be divided into expressed and latent traits. He called these **dominant and recessive traits**, respectively. Dominant traits are those that are inherited unchanged in a hybridization. Recessive traits become latent, or disappear in the offspring of a hybridization. The recessive trait does, however, reappear in the progeny of the hybrid offspring. An example of a dominant trait is the violet-colored flower trait. For this same characteristic (flower color), white-colored flowers are a recessive trait. The fact that the recessive trait reappeared in the F_2 generation meant that the traits remained separate (and were not blended) in the plants of the F_1 generation. Mendel proposed that this was because the plants possessed two copies of the trait for the flower-color characteristic, and that each parent transmitted one of their two copies to their offspring, where they came together. Moreover, the physical observation of a dominant trait could mean that the genetic composition of the organism included two dominant versions of the

characteristic, or that it included one dominant and one recessive version. Conversely, the observation of a recessive trait meant that the organism lacked any dominant versions of this characteristic.In this famous experiment, Mendel purposefully crosspollinated pea plants based on their different features to make important discoveries on how traits are inherited between generations. Seven traits were used by Mendel, including smooth or wrinkled ripe seeds, yellow or green seed albumen, purple or white flower, tall or dwarf stem length, and others.
Over the course of his experiments, Mendel made three important discoveries:

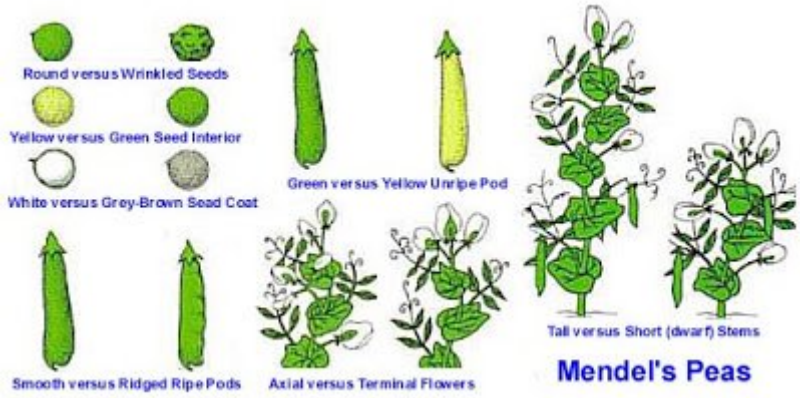
1. The Law of Segregation: offspring acquire one hereditary factor from each parent

2. The Law of Independent Assortment: different traits have an equal opportunity of occurring together (this was later shown to not entirely be true)

3. The Law of Dominance: offspring will inherit the dominant trait, and can only inherit the recessive trait if they inherit both recessive factors

His first step was to establish pea plant populations with two different features, such as tall vs. short height, breeding them until they always produced offspring identical to the parent. After this, he then bred them with each other to observe how the offspring inherited the traits. This first generation found that all the offspring shared one feature, which he called he dominant trait, and did not display the other type, the recessive trait. In this example, they were all tall. His next step in his experiment was to allow this generation that displayed only the dominant trait to self-fertilize, creating a new generation that displayed the hidden trait. The second generation that came from the first-generation self fertilizing had the recessive trait reappear in about 1 in 4 of the plants, with there being a 3:1 ratio for plants that showed the dominant trait for every plant that showed the recessive trait.

Contrasting characters that Mendel focussed on garden pea plant



Neomendelism

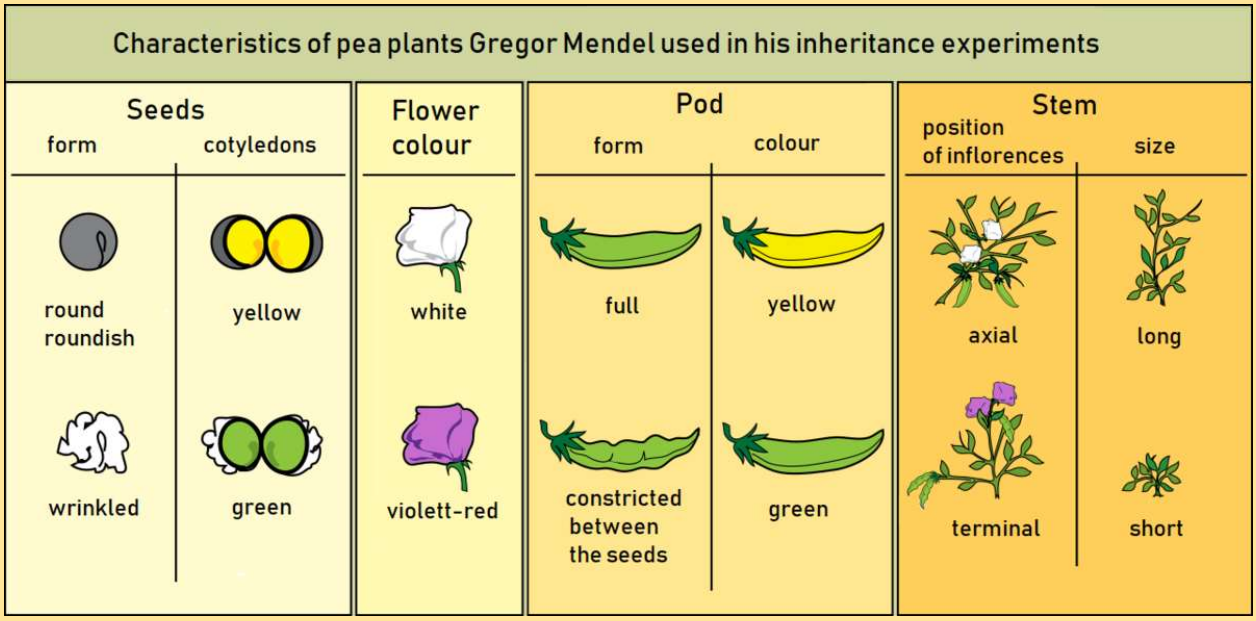
New discoveries made by modifying the Mendel's law is termed as neomendelism. They include incomplete dominance, co-dominance, epistasis etc.

Reasons for Mendel's success

1. The pea plant which Mendel chose for conducting experiments, is most ideal for controlled breeding, since it can easily be subjected to cross pollination.
2. He identified very clear contrasting characters in the pea plants.
3. He selected pure breeding plants for his experiments. He is said to have spent about 2 years to ascertain this characteristic feature.
4. Mendel concentrated at a time only on the inheritance on one particular trait, with the two contrasting conditions, instead of attempting the inheritance of entire set of characters in the plant.
5. He maintained an accurate record of all the observations he made on the breeding experiments that he had designed.
6. He pooled the data obtained from similar experiments for different characteristics and analysed the results by using statistical methods and applying the law of probability.
7. Mendel was also fortunate in the sense that the characters he had chosen in the pea plant did not show any interaction or linkage.
8. Mendel was fortunate enough in choosing the seven pairs of contrasting characters in pea plants. It was later discovered that the genes responsible for these characters are located on

separate chromosomes.

9. He was able to effectively check the flowers under investigation from contamination by unwanted pollen grains.



Glossary

continuous variation: a variation in a characteristic in which individuals show a range of traits with small differences between them

discontinuous variation: a variation in a characteristic in which individuals show two, or a few, traits with large differences between them

dominant: describes a trait that masks the expression of another trait when both versions of the gene are present in an individual

F_1: the first filial generation in a cross; the offspring of the parental generation

F_2: the second filial generation produced when F_1 individuals are self-crossed or fertilized with each other

hybridization: the process of mating two individuals that differ, with the goal of achieving a certain characteristic in their offspring

model system: a species or biological system used to study a specific biological phenomenon to gain understanding that will be applied to other species

P: the parental generation in a cross

recessive: describes a trait whose expression is masked by another trait when the alleles for both traits are present in an individual

reciprocal cross: a paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross

trait: a variation in an inherited characteristic

Gene Interactions

The genes of an individual do not operate isolated from one another, but obviously are functioning in a common cellular environment. Thus, it is expected interactions between genes would occur. Bateson and Punnett performed a classical experiment that demonstrated genetic interactions. They analyzed the three comb types of chicken known to exist at that time:

Chicken Varieties	Phenotype
Wyandotte	Rose Comb
Brahmas	Pea Comb
Leghorns	Single Comb

Rose  **Pea**

Single  **Walnut**

Parental Cross	Rose Comb (Wyandotte) x Pea Comb (Brahama) (Both parents were pure lines.)
F1 Result	↓ All Walnut Combs (This is a new phenotype not previously seen.)
F2 Result	↓ 9 Walnut : 3 Rose : 3 Pea : 1 Single

Result: The F_1 differed from both parents and two new phenotypes not seen in the parents appeared in the F_2. How can this result be explained? The first clue is the F_2 ratio. We have seen this ratio before when the F_1 from a dihybrid cross is selfed (or intermated). This observation suggests that two genes may control the phenotype of the comb. The gene interactions and genotypes were determined by performing the appropriate testcrosses.

A series of experiments demonstrated that the genotypes controlling the various comb phenotypes are as follows.

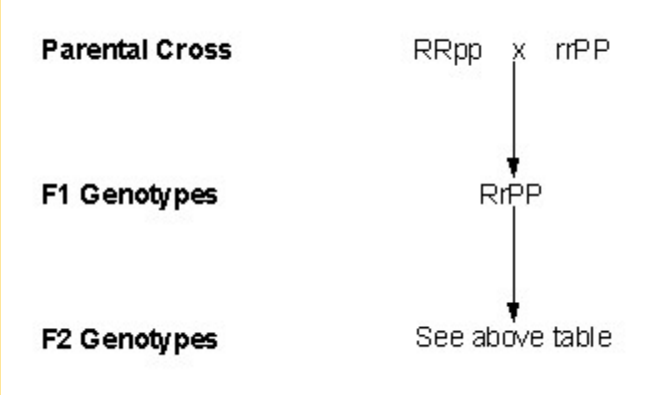
Phenotypes	**Genotypes**	**Frequency**
Walnut	*R_P_*	9/16
Rose	*R_pp*	3/16
Pea	*rrP_*	3/16
Single	*rrpp*	1/16

It was later shown that the genotypes of the initial parents were:

Rose = *RRpp*

Pea = *rrPP*

Therefore, genotypically the cross was:



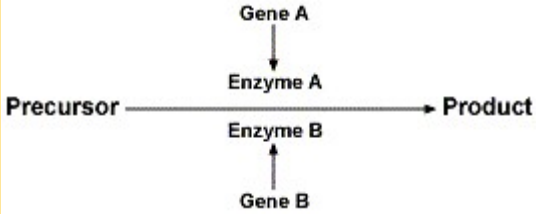
The development of any individual is obviously the expression of all the genes that are a part of its genetic makeup. Therefore, it is not an unexpected conclusion that more than one gene could be responsible for the expression of a single phenotype. We will now discuss this situation. First let's give a definition.

Epistasis - the interaction between two or more genes to control a single phenotype

The interactions of the two genes which control comb type was revealed because we could identify and recognize the 9:3:3:1. Other genetic interactions were identified because the results of crossing two dihybrids produced a modified Mendelian ratio. All of the results are modifications of the 9:3:3:1 ratio.

Example 1: 15:1 Ratio

Phenotypes: Kernel Color in Wheat



For this type of pathway a functional enzyme A or B can produce a product from a common precursor. The product gives color to the wheat kernel. Therefore, only one dominant allele at either of the two loci is required to generate the product.

Thus, if a pure line wheat plant with a colored kernel (genotype = *AABB*) is crossed to plant with white kernels (genotype = *aabb*) and the resulting F_1 plants are selfed, a modification of the dihybrid 9:3:3:1 ratio will be produced. The following table provides a biochemical explanation for the 15:1 ratio.

Genotype	**Kernel Phenotype**	**Enzymatic Activities**
9 *A_B_*	colored kernels	functional enzymes from both genes
3 *A_bb*	colored kernels	functional enzyme from the *A* gene pair
3 *aaB_*	colored kernels	functional enzyme from the *B* gene pair
1 *aabb*	colorless kernels	non-functional enzymes produced at both genes

If we sum the three different genotypes that will produce a colored kernel we can see that we can achieve a 15:1 ratio. Because either of the genes can provide the wild type phenotype, this interaction is called duplicate gene action.

Example **2:** 9:7 Ratio

Example: Flower color in sweet pea

If two genes are involved in a specific pathway and functional products from both are required for expression, then one recessive allelic pair at either allelic pair would result in the mutant phenotype. This is graphically shown in the following diagram.

If a pure line pea plant with colored flowers (genotype = *CCPP*) is crossed to pure line, homozygous recessive plant with white flowers, the F_1 plant will have colored flowers and a *CcPp* genotype. The normal ratio from selfing dihybrid is 9:3:3:1, but epistatic interactions of the *C* and *P* genes will give a modified 9:7 ratio. The following table describes the interactions for each genotype and how the ratio occurs.

Genotype	**Flower Color**	**Enzyme Activities/TH>**
9 *C_P_*	Flowers colored; anthocyanin produced	Functional enzymes from both genes
3 *C_pp*	Flowers white; no anthocyanin produced	p enzyme non-functional
3 *ccP_*	Flowers white; no anthocyanin produced	c enzyme non-functional
1 *ccpp*	Flowers white; no anthocyanin produced	c and p enzymes non-functional

Because both genes are required for the correct phenotype, this epistatic interaction is called **complementary gene action**.

Example 3: 12:3:1 Ratio
Phenotype: Fruit Color in Squash

With this interaction, color is recessive to no color at one allelic pair. This recessive allele must be expressed before the specific color allele at a second locus is expressed. At the first gene white colored squash is dominant to colored squash, and the gene symbols are *W*=white and *w*=colored. At the second gene yellow is dominant to green, and the symbols used are *G*=yellow, *g*=green. If the dihybrid is selfed, three phenotypes are produced in a 12:3:1 ratio. The following table explains how this ratio is obtained.

Shapes of Squash Fruit

Genotype	Fruit Color	Gene Actions
9 *W_G_*	White	Dominant white allele negates effect of *G* allele
3 *W_gg*	White	Dominant white allele negates effect of *G* allele
3 *wwG_*	Yellow	Recessive color allele allows yellow allele expression
1 wwgg	Green	Recessive color allele allows green allele expression

Because the presence of the dominant *W* allele masks the effects of either the *G* or *g* allele, this type of interaction is called **dominant epistasis**.

Example 4: 13:3 ratio

Phenotype: Malvidin production in *Primula*

Certain genes have the ability to suppress the expression of a gene at a second locus. The production of the chemical malvidin in the plant *Primula* is an example. Both the synthesis of the chemical (controlled by the *K* gene) and the suppression of synthesis at the *K* gene (controlled by the *D* gene) are dominant traits. The F_1 plant with the genotype *KkDd* will not produce malvidin because of the presence of the dominant *D* allele. What will be the distribution of the F_2 phenotypes after the F_1 was crossed?

Genotype	Phenotype and genetic explanation
9 *K_D_*	no malvidin because dominant *D* allele is present
3 *K_dd*	malvidin productions because dominant *K* allele present

3 *kkD_* no malvidin because recessive *k* and dominant D alleles present

1 *kkdd* no malvidin because recessive *k* allele present

The ratio from the above table is 13 no malvidin production to 3 malvidin production. Because the action of the dominant *D* allele masks the genes at the *K* locus, this interaction is termed **dominant suppression epistasis**.

Suppressor - a genetic factor that prevents the expression of alleles at a second locus; this is an example of epistatic interaction

Remember that epistasis is the interaction between different genes. If one allele or allelic pair masks the expression of an allele at the second gene, that allele or allelic pair is epistatic to the second gene. Therefore, the following table summarizes the four epistatic interactions discussed above.

Example	**Allelic Interactions**	**Type of Epistasis**
Wheat kernel color	*A* epistatic to *B*, *b* *B* epistatic to *A*, *a*	Duplicate genes
Sweet pea flower color	*cc* epistatic to *P*, *p* *pp* epistatic to *C*, *c*	Complementary gene action
Squash Fruit Color	*W* epistatic to *G*, *g*	Dominant epistasis
Primula malvidin production	*D* epistatic to *K*, *k*	Dominant suppression

Interaction of Non-allelic genes

In some cases there is a trait made up of more genes. This causes a **discrepancy in segregation ratios**. Usually, many genetic textbooks propose interaction between two genes each with two alleles with complete dominance. Segregation ratios in F2 generations of all types of interactions are summarized in the table below.

Dominant epistasis (12:3:1) sets up when dominant gene realize its potential regardless of recessive gene. Only when dominant gene has both alleles recessive then recessive gene can realize its phenotype.

On the contrary **recessive epistasis (9:3:4)** occurs when both recessive alleles of one gene produce uniform phenotype regardless of genotype of the second gene.

In **complementary factor (9:7)** when either of both genes is recessive homozygous, then it leads to an identical phenotype regardless of genotype of the other gene. To produce other phenotype, both genes have to have at least one dominant allele.

Polymorphic gene (9:6:1) sets up when both genes are responsible for producing same trait. Then genotype aabb does not produce anything, genotypes A-bb and aaB- produce a half of what A-B- produce. All in all, there is a cumulatory effect.

Duplicate gene (15:1) is similar to polymorphic gene but there is no cumulation when both genes have a dominant allele.

Inhibitory factor (13:3) happens when dominant genotype of one gene and recessive genotype of the other have same phenotype. In F2 generation we then receive two phenotype classes. One with genotypes A-B-, A-bb, aabb and second with genotypes aaB-.

	A-B-	A-bb	aaB-	aabb
Simple interaction	9	3	3	1
Dominant epistasis	12		3	1
Recessive epistasis	9	3	4	
Complementary factor	9	7		
Duplicate gene	15			1
Polymorphic gene	9	6		1
Inhibitory factor	13		3	

Multiple Alleles

Alleles are the pairs of genes occupying a specific spot called locus on a chromosome. Typically, there are only two alleles for a gene in a diploid organism. When there is a gene existing in more than two allelic forms, this condition is referred to as multiple allelism. Allelism refers to any of the several forms of a gene. These genetic variations arise usually through mutation and therefore are responsible for hereditary variations.

Gregor Mendel suggested that each gene would have only two alleles. Alleles are described as a *variant* of a gene that exists in two or more forms. Each gene is inherited in two alleles, i.e., *one from each parent*. Thus, this means there would also be having two different alleles for a trait.Offspring who have different alleles of a gene are described as *heterozygous* whereas those that have genotypes made of the same alleles (i.e. of a gene for a particular trait) are described as *homozygous*. Although humans (and all diploid organisms) can have only two alleles for any given gene in genetics, multiple allele traits may exist at the population level. Thus, multiple alleles are important in promoting variation within the same species.

In ***Mendelian inheritance***, a given chromosomal *locus* was occupied by two distinct types of gene alternatives – one *dominant* and one *recessive.* These alternatives are two alleles of the same gene. However, in the given population, there are instances wherein there are more than two alternatives (alleles) existing. Those instances are referred to as “multiple allelism” where different forms of the same gene exist within the population. These three or more variants for the same gene are called multiple alleles.

Examples of Multiple Alleles

Let us understand the concept further with the different *multiple alleles* examples presented below.

Coat color of cats

For thousands of years, domestic cats have been bred to achieve a variety of coat colors. The gene that determines the coat color of the cat appears to have multiple variants as the coats could

range from black, orange, brown, etc., to white. This means that there are multiple alleles responsible for the coat color.

The coat color gene has multiple alleles in the population and the pigment-producing protein will depend on the inheritance and the expression of these alleles. Other genes regulate curliness, shading, patterns, and even texture in a similar manner. The number of possible combinations and expressions of distinct genotypes from these alleles results in a wide variety of breeds. Even when only four alleles are shared between two parents for each gene, the variation can be astounding.

Take a look at Figure 1. The coat color of cats clearly indicates multiple allelism as it appears there are more than two alleles for the coat color phenotype — i.e. patches of black and orange (tortoiseshell pattern), black, grey, white, and patches of white fur (piebald spotting).Generally, the genotype of a cat can be determined by examining its coat coloration and pattern of coloration. If the phenotypes of the parental cats are known, it is usually possible to predict the coloration possibilities of kittens, though the calculations would be complex in most cases.

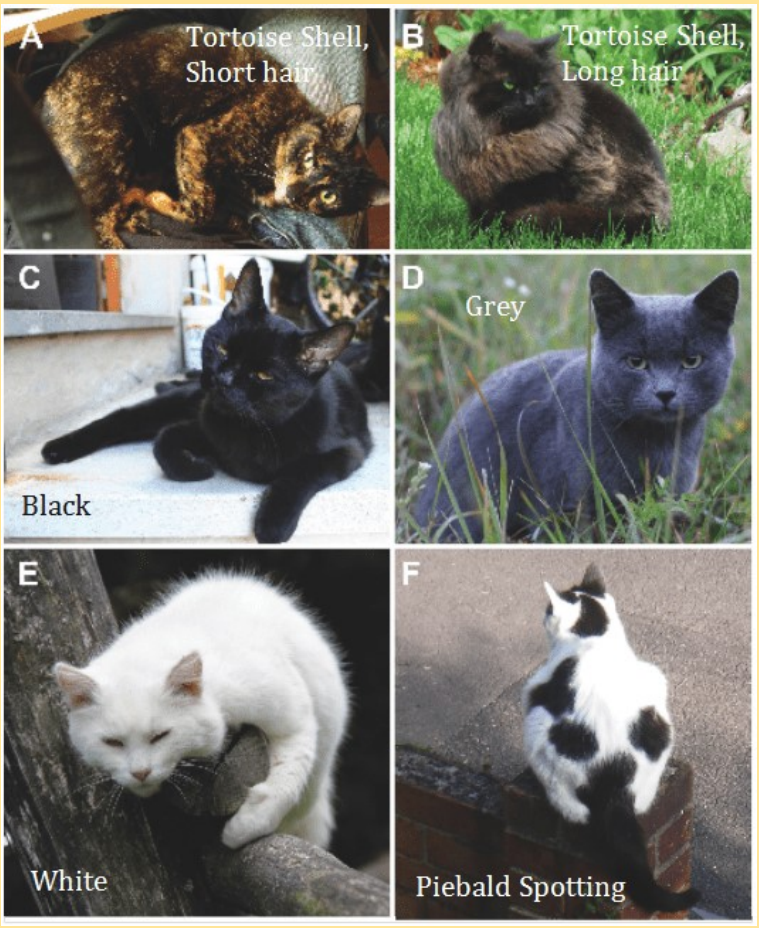


Figure 1: Images showing different coat colors among cats. Credit: Todd Nickle and Isabelle Barrette-Ng – LibreTexts, CC SA 3.0.

Multiple alleles in fruit flies

In 2000, scientists completed the complex genome mapping of the common fruit fly, *Drosophila melanogaster*. The fruit fly has been and continues to be a valuable laboratory animal due to its high reproduction rate and the ease with which large numbers of flies can be kept and analyzed. At approximately 165 million base pairs, a fruit fly's DNA is significantly smaller than that of a human. While humans have 23 chromosomes, fruit flies have only four. Nonetheless, there are approximately 17,000 genes on only four chromosomes. Each gene is responsible for a distinct aspect of the fly and is subject to mutation and the emergence of new alleles.

Normally, the wings of *Drosophila melanogaster* are quite long. Two mutations occurred at the same locus in different flies' one of which resulted in vestigial (reduced) wings and the other in antlered (less developed) wings.

When a fly with vestigial wings is crossed with another with antlered wings, the resulting F1 hybrids have intermediate wing lengths, indicating that none of the mutated genes is dominant. Occasionally referred to as the vestigial antlered compound, this hybrid contains two mutated genes at the same locus. Mendelian segregation and recombination are evident.

Apart from the vestigial and antlered wings, there are other phenotypes: *nicked wings, strap wings,* or *no wings* at all. The gene variants responsible for these phenotypes are the multiple alleles in this particular fruit fly population.

Multiple alleles among humans

There are traits in humans and other organisms that have three or more different types of alleles (genes). When a trait has three or more distinct alleles, we refer to it as having multiple alleles inheritance. The human ABO blood type alleles/trait is an example of a trait with multiple alleles. Three distinct alleles exist: allele A (I^A), allele B (I^B), and allele i (I^O or i).

If the allele A is present on the chromosome, protein A is produced, and the red blood cells of that individual contain protein A on their membrane. If the chromosome contains the allele B, protein B will be produced, and the red blood cells' membranes will contain protein B. Finally, if allele i is present on the chromosome, neither protein A nor protein B will be synthesized. These three alleles make up the ABO blood group trait.

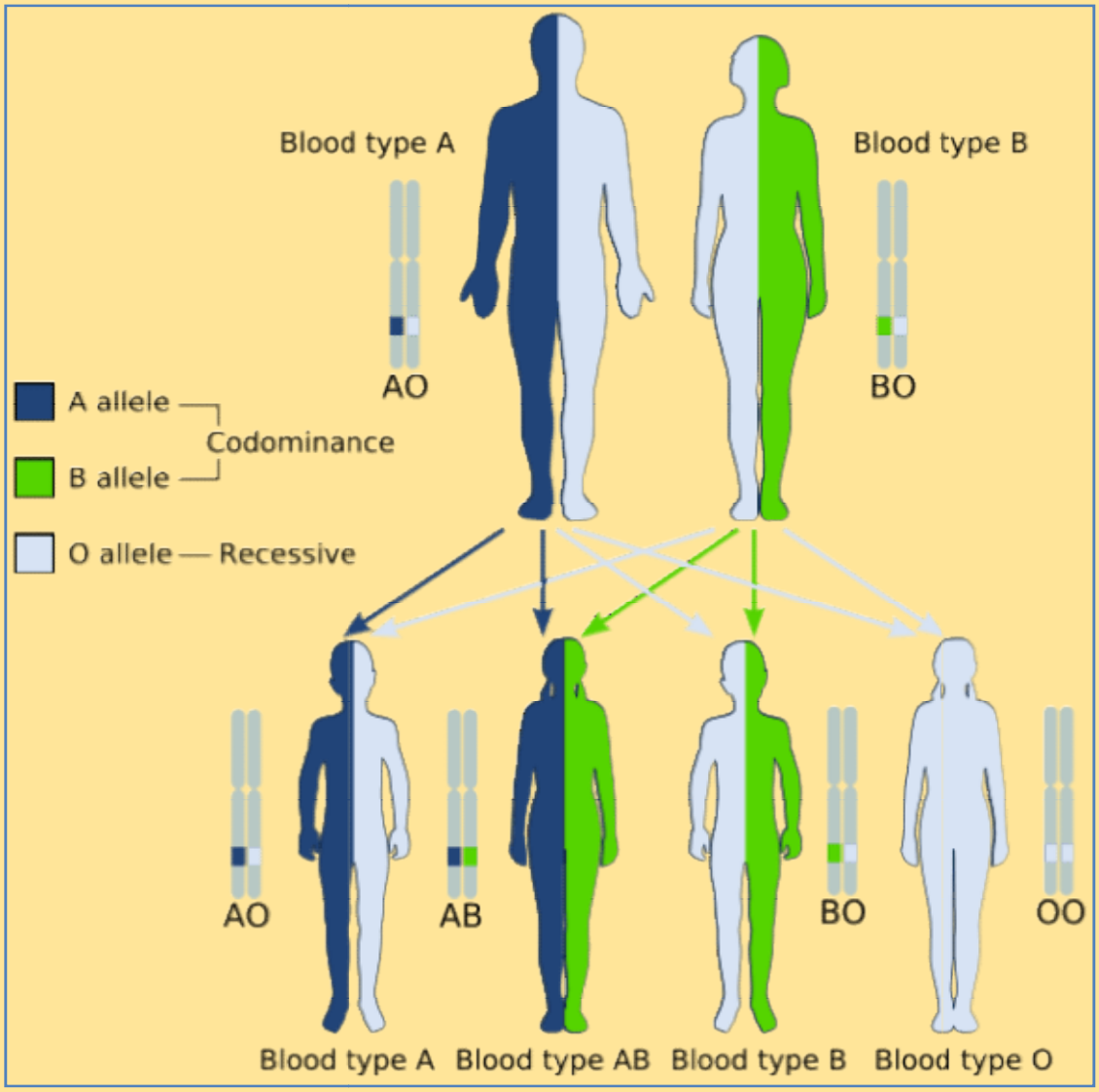


Figure 2: Possible blood types of offspring from a cross between male and female parent with blood type A and Blood type B, respectively.

The inheritance pattern of allele A and allele B shows codominance (co-dominance). Co-dominance occurs when neither allele is dominant over the other, and a heterozygous individual expresses both phenotypes. That is, if an individual possesses allele A on one chromosome and allele B on the second homologous chromosome, both proteins are expressed and red blood cells contain both protein A and protein B on their cell membranes.

READ: Non-Mendelian Inheritance – Multiple Alleles (ABO blood type)

The ABO blood type hereditary scheme in human beings is an illustration of multiple alleles blood types. Phenotypes are classified into four categories: blood group A, blood group B, blood group AB, and blood group O. In this case, the population has three alleles. The I^A allele when expressed results in the presence of A molecules on red blood cells, the I^B allele expression leads to the presence of B molecules on red blood cells, and the I^O allele expression results in the absence of such antigens on the red blood cells. The I^A and I^B alleles are not only codominant but both of them are dominant over the I^O allele. I^O allele is recessive and thus will be expressed if I^A or I^B is absent.

Although a population contains three alleles, each individual inherits only two of them from their parents. This results in the genotypes and phenotypes depicted below. Consider that when three alleles are present, there are six distinct genotypes. The number of phenotypes possible is determined by the dominance relationships between the three alleles. Watch the video below to understand multiple alleles in the form of ABO Blood types.

Multiple alleles in plants

While it is widely believed that potato tuber shape is continuous, visual phenotypes such as round or long tubers can be discerned at the diploid level. Although experimental evidence for the presence of multiple allele systems for a potato tuber is being reported for the first time, this study in potato can be compared to one in maize. The recessive allele for tuber shape can be regarded as the qualitatively recognized null or near-null allele. Variation between dominant alleles is quantitative. The view that the (most) recessive allele can be considered a null or near-null allele is consistent with how quantitative effects at a multiallelic locus are described. When additional metric traits are resolved into Mendelian factors using heterozygous parents in experimental designs, conclusions can be drawn about the relative importance of multiple alleles traits to multiple loci in explaining quantitative genetic alleles variation.

Multiple alleles in bacteria

Bacteria possess a large number of genes, several of which have multiple alleles. These various wild-type alleles are frequently associated with distinct types of virulence and can be used to

classify subspecies (e.g., housekeeping genes for Multi Locus Sequence Typing, MLST). As a result, it is critical to rapidly identify not only the desired gene but also the relevant allele.

Currently available sequencing-based methods are limited to mapping reads to each known allele reference, which is a lengthy process. Understanding and predicting the pathogenic impact and outbreak potential of a bacterial infection requires more than knowledge of the species responsible for the infection.

Bacterial virulence is frequently regulated on a sub-species level by a collection of specific genes or even alleles, necessitating the use of distinct treatment strategies for infections caused by the same bacterial species.

Antibiotic resistance, for example, is a well-known example of how minor variations in a gene can result in a diverse collection of antibiotic resistance profiles within a single taxonomic group.

Additionally, different alleles of the same gene may be responsible for distinct adhesion and invasion strategies, responses to the infected organism's immune response, and toxin production.

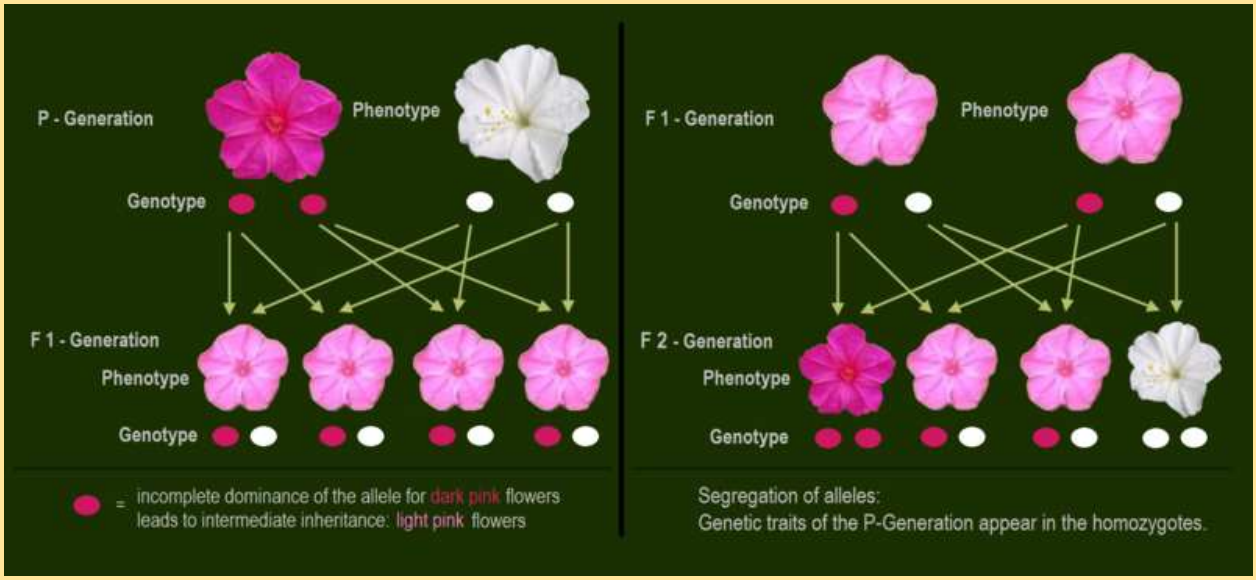
Apart from its importance for understanding virulence, identifying alleles of specific genes contributes to a more precise classification of bacteria.

Multiple Allelism, Pleiotropy and Epistasis

Multiple allelism may be misconstrued with other common terms in Genetics, such as pleiotropy and epistasis. Let us understand the difference between multiple allelism (defined as the condition of having multiple alleles) and other genetic conditions.

Pleiotropy refers to the condition wherein more than one gene can have multiple effects on the phenotype. While multiple allelism involves only a single gene with many variants (referred to as multiple alleles), pleiotropy entails more than one gene that determines phenotype. For example, a higher frequency of individuals with albino would have cross-eyes than the pigmented individuals. Thus, apart from having no sufficient pigment production in their skin and hair, another possible feature of albinoes is a cross-eyed trait. However, not all albinoes show this trait indicating that in such cases, the two traits are not linked.

Epistasis is when one gene has an effect on the expression of another gene. This occurs when genes interact to produce a particular trait. An example would be the determination of coat color in certain animals (e.g. horses) wherein the effect of one gene depends on the influence of another gene controlling the deposition of hair pigment.



Linkage and Crossing Over

Linkage and Crossing Over: Gregor John Mendel demonstrated in his experiment that characters are determined by certain factors. Such factors are stable and segregate independently at the time of gamete formation. He was, however, unaware of the location of these factors in the cell and thus could not identify the physical counterparts of these factors. The development of new and improved techniques subsequently helped in the discovery of chromosomes, chromosomal theory of inheritance, and the process of cell division. The significant workings of geneticists on the chromosomal theory of inheritance put light on the phenomenon of **Linkage Crossing Over**. Let's go deep into the article to study the characteristics, types, and significance of linkage and crossing.

Definition of Linkage: Since the number of genes in an organism is much larger than the number of pairs of chromosomes, therefore each chromosome pair must contain several genes. This means that during cell division, chromosomes move as a unit, and all the genes

of a chromosome move together and do not assort independently. Thus, linkage can be defined as the **tendency of certain genes of the same chromosome to be inherited together during chromosomal inheritance.**

William Bateson and **R.C. Punnett** discovered the phenomenon of Linkage. However, **Morgan** formulated the concept of linkage by his workings of inheritance in *Drosophila melanogaster*.

Key Terms Related to Linkage

Find below the key terms related to linkage:

1. **Linked genes:** The genes that do not show independent assortment and are inherited together with other genes as they are closely placed on the same chromosome are called linked genes.
2. **Non-linked genes:** The genes that are located farther apart from each other on a chromosome or the two chromosomes of a homologous pair are called non-linked genes. These genes show independent assortment.
3. **Linkage Group:** All the genes located on a homologous pair of chromosomes collectively form a group called the linkage group**.** The number of linkage groups in an organism is the same as the haploid number of chromosomes or the number of pairs of chromosomes in the organism.
4. **Linkage Value:** It is the degree of intensity with which the two genes are linked together. The linkage value depends on the distance between the linked genes on the same chromosome.
5. **Linkage Map:** It is a genetic map of genes located on a chromosome that helps to determine how often two gene loci are inherited together.

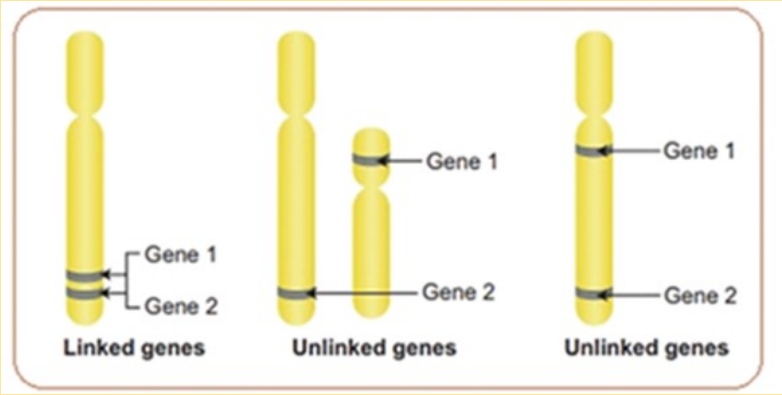


Fig: Arrangement of Linked and Non-linked Genes

Types of Linkage

Linkage can be classified on the basis of different parameters of inheritance. These are discussed as follows:

I. **Based on crossing over:** The linkage can be of two types:

1. **Complete linkage:** The two adjacent genes located on the same chromosome do not separate and are inherited together over generations due to the absence of crossing over. It is rare, but Morgan demonstrated the complete linkage in male *Drosophila* through his experiment.
2. **Incomplete linkage:** It is the phenomenon in which particular and well-tested characters, instead of appearing together for generations, give rise to a new combination in the F_2 generation. A dihybrid produces four types of gametes; however, the two parental types occur more frequently than the non-parental types. Incomplete linkage leads to the formation of new combinations due to the exchange of chromosome segments during crossing over. Incomplete linkage is common in organisms.

II. **Based on the chromosome involved:** The linkage can be classified into two types based on the types of chromosomes:

1. **Autosomal linkage:** It is the linkage of genes located on the autosome other than the sex chromosome. Humans have 22 autosomal linkage groups.
2. **Allosomal linkage:** It is the linkage of genes located on a sex chromosome. There are two types of sex chromosomes found in humans; X and Y chromosomes.

III. **Based on Genes Involved:** Depending on dominant and recessive alleles in a pair of genes, the linkage can be categorised into two phases:

1. **Coupling or Cis phase:** It is the phase or linkage in which either dominant alleles or recessive alleles of both the genes are present together on the same chromosome and are inherited together.
2. **Repulsion or Trans phase:** It is the phase when the dominant allele of a trait or gene is linked or paired with the recessive allele of another gene of the same chromosome.

Examples of Linkage

Bateson and **Punnett** perform a cross between the dominant pea plant with blue flowers and long pollen with a recessive pea plant having red flowers and round pollen to exhibit

the **incomplete linkage** during the cross. Besides maize plants, most organisms show the phenomenon of incomplete linkage.

A. **Incomplete Linkage in Maize:** This experiment of chromosomal inheritance is demonstrated by Hutchinson and can be described as follows:

1. A cross is made between the maize plant having coloured and full seeds with another maize plant having colourless and shrunken seeds.
2. Coloured and full seeds are dominant over colourless and shrunken seeds as all the plants of the F_1 generation have coloured and full seeds.
3. A test cross between recessive plant (colourless and shrunken seeds) and F_1 hybrid (heterozygous dominant) leads to the formation of the following four different types of seeds with a new combination:

- Coloured and full seeds
- Colourless and shrunken seeds
- Coloured and shrunken seeds
- Colourless and full seeds

However, the parental combinations are about 96.4% and new combinations about 3.6%.

This shows that paternal characters located on the same chromosome are linked together except for a small percentage of new combinations that arise due to the recombination (exchange between the non-sister chromatids of a homologous pair of chromosomes) during crossing over.

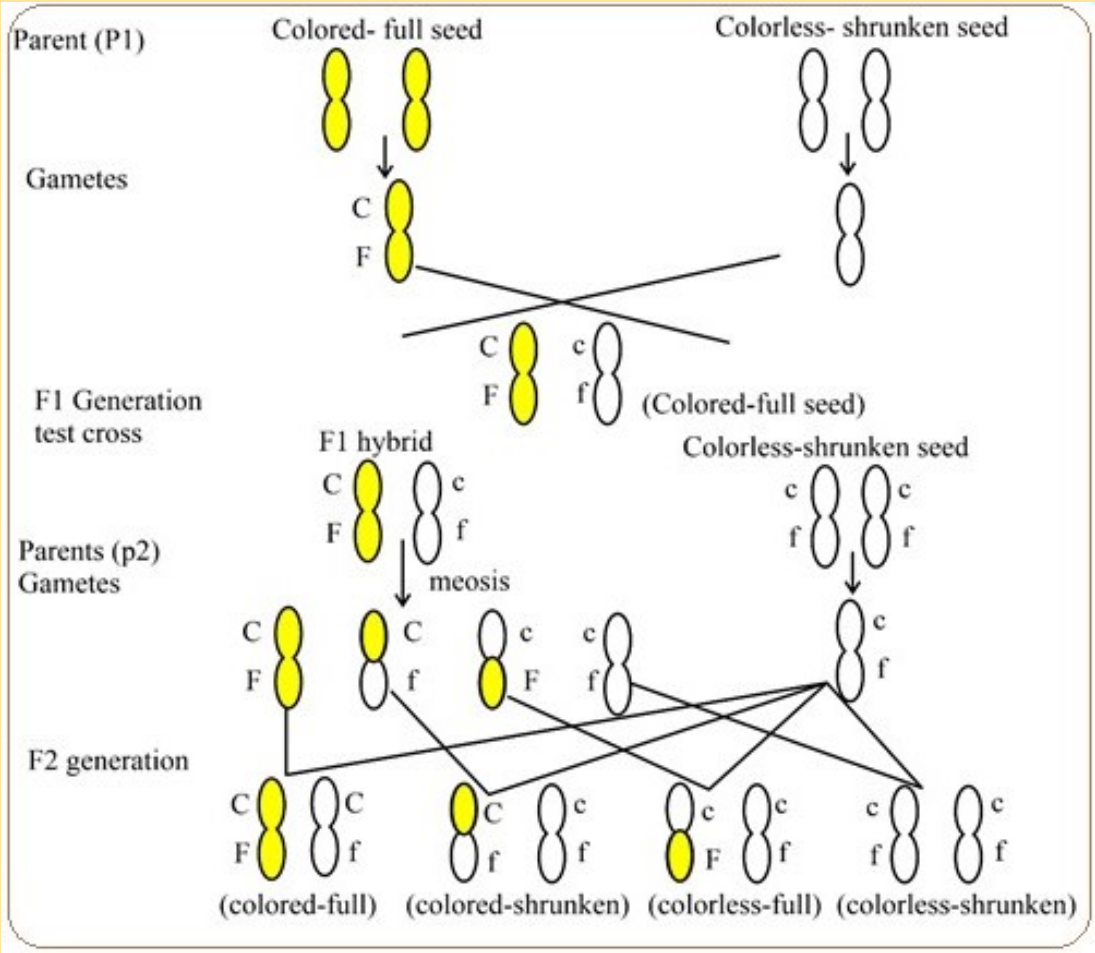


Fig: Incomplete linkage in Maize

B. **Complete linkage in male** *Drosophila***:** This experiment is demonstrated by Morgan. It can be described in the below-listed points:

1. A dominant *Drosophila* with a grey body and long-winged is crossed with a recessive *Drosophila* with a black body and vestigial wings.
2. The hybrids of the F_1 generation will be like those of dominant *Drosophila*, exhibiting the feature of grey body and long-winged.
3. A cross is then performed between the hybrids of the F_1 generation and recessive parent.
4. Four kinds of offspring in equal numbers were expected as a result of independent assortment. However, the offspring formed were phenotypically similar to the original parents used in the experiment and did not show any recombinant forms.

5. The results of the cross justify that the two closely placed genes of a chromosome are linked together. Hence, the grey body and long-winged characters appear together. On the other hand, the black body character is linked with vestigial wings.

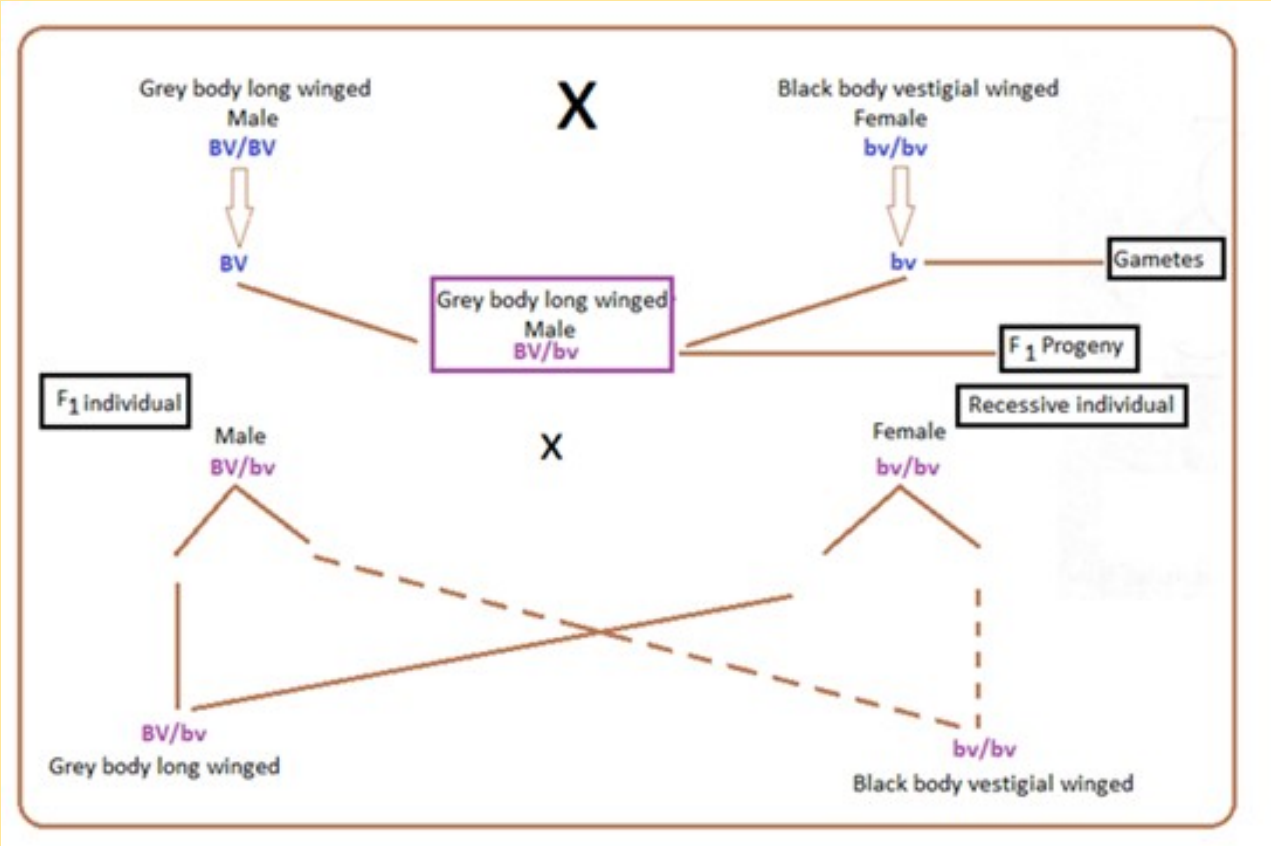


Fig: Complete linkage in Drosophila

Crossing-over

Crossing-over is the process by which homologous chromosomes exchange segments with each other. It occurs most often during the first meiotic division. When prometaphase begins, each chromosome has been duplicated to form two identical sister chromatids. Crossing-over takes place between homologous pairs of chromosomes, with sister chromatids of each homolog swapping segments at various places along the length of the chromosome. Crossing over also occurs between sister chromatids, but because they are genetically identical, such crossing over will not result in genetic recombination.The DNA strands must be broken to exchange their segments. The portions of the chromosome that undergo crossing over contain the same gene

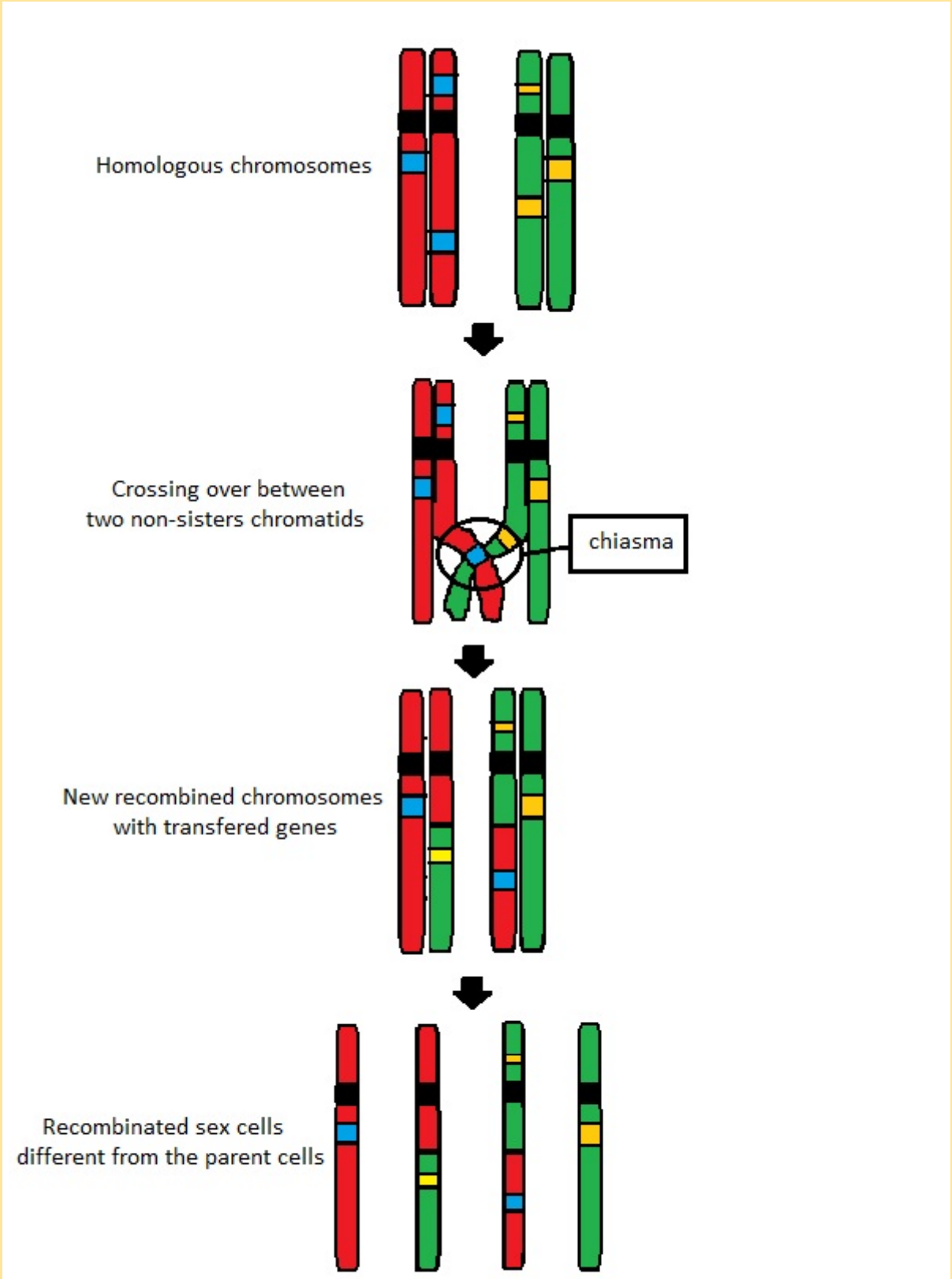
loci. Crossing-over leads to the recombination of the genes on the chromosomes. The point of crossover is visible as a cross-shaped chiasma.

The exchange is usually reciprocal – the exchanging segments of the two chromosomes are of similar size--, but crossing over can sometimes be unequal. Crossing over can occur at several locations between a synapsing pair of homologous chromosomes. **Crossing-over may be**:

1. *single crossing-over*;
2. *double crossing-over*;
3. *multiple crossing-over*.

The importance of crossing-over is in the generation of **genetic diversity** in the resulting gametes. *New, genetically recombined sex cells* are created. This leads to variation among offspring. Any given genetic combination can be advantageous or disadvantageous, but genetic variability of offspring increases the likelihood that at least some of the offspring will survive in an ever-changing environment.

The frequency of crossing-over between two gene loci depends on the *distance between the linked genes*. **Linked genes** are those located on the same chromosome. Unless crossing over occurs between them, the linked genes are transferred together every time meiosis occurs. If crossing-over occurs between linked genes, alleles that are usually inherited together can be recombined and passed on separately in two different sex cells. The likelihood of crossing over between two gene loci can be used to construct gene maps.



Crossing-over

There

Crossing Over and its mechanism

- **The phenomenon** of separation of linked genes is termed as **crossing over.**
- Linked genes separate during inheritance and produce **new or non-parental** combinations in case of **incomplete linkage.**
- The parental combinations are more than **the non-parental** combinations.
- It is also a **physical process** of exchange of genetic materials or segments between non-sister **chromatids** of two homologous chromosomes during meiosis.
- It results in **production of new** or non-parental combinations.
- It occurs during **prophase I** of **first meiotic division.**
- Generally, the **two non-sister chromatids** cross each other at one or more points.
- Such points of crossing over are known as **chiasma.**
- There is a break at the point of crossing over which is followed by a **re-union of corresponding** segments of chromatids.

Mechanism of crossing over

Mechanism of crossing over can be described under following headings:

1.Synapsis

- The **paternal and maternal chromosome** of a homologous pair come close together during **the prophase I of meiosis cell division.**
- They pair at the **zygotene sub-stage** which is called **synapsis.**
- During synapsis, **the homologous chromosomes** lie side by side and coil around each other to form a **bivalent.**

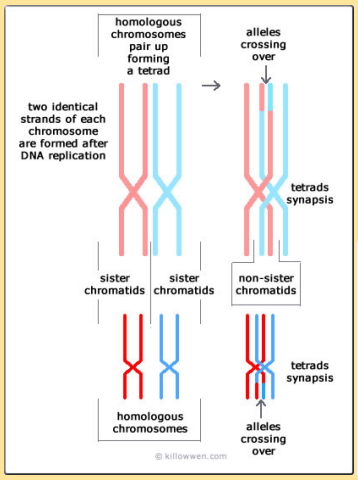
2.Duplication of chromosomes

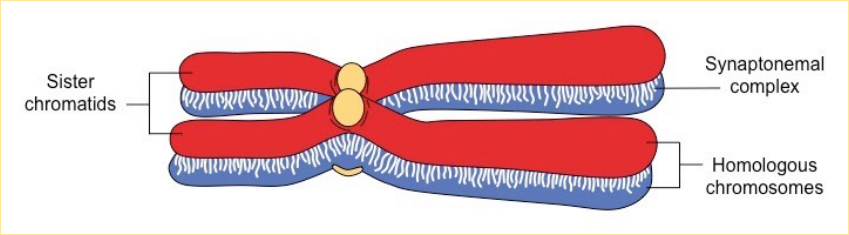
- **Synapsis** is followed by duplication of chromosomes.
- Here, each of the homologous chromosome in a **bivalent splits** longitudinally into **two sister chromatids** during **diplotene stage**.
- Thus the bivalent now consists of **four chromatids** which is called **tetrad.**

3.Crossing over

- During **pachytene stage**, when the paired chromosomes start separating, **the chromatids** remain attach at one or more points and thus establish one or more exchange per bivalent.

- This points of attachment are known as **chiasma.**
- **At each chiasma**, two non-sister chromatids of the bivalent break at the corresponding points and then rejoin with the exchange of segments forming new chromatids.





4.Terminalization

- The non-sister chromatids start repelling each other during **diakinesis sub stage** after completion of crossing over.
- The **chromatids separate** from the centromere.
- The chiasma shifts towards the **terminal ends.**
- The movement of chiasma towards ends is known as **terminalization.**
- Gradually **the chromatids get** condensed and thus shortened and the homologous chromosomes become separated.

Kinds of crossing over

- The **number of chiasma** usually depends upon the length of chromosome.
- Crossing over is of **three type**s depending upon the **number of chiasma**. They are:

1.Single crossing over

- In single crossing over there is **only one chiasma** in the chromosome i.e. the chromatids of homologous chromosome contact and break only at one point along their entire length.

2. Double crossing over

- In double crossing over **chiasma are formed at two points** i.e. chromatids break and rejoin at two points.
- It means **two chiasmata** are formed along the entire length of the chromosome.

3.Multiple crossing over

- When crossing over occurs at **more than two places** in the same chromosome pair and more than two **chiasmata** are formed, the crossing over is known as multiple crossing over.

Frequency of crossing over

- The frequency of crossing over depends upon **the distance of genes** between which crossing over occurs.
- The frequency of crossing over between any two linked genes increases with the distance between them.
- Similarly, **the closer the two genes are linked**, the lesser the chance for a chiasma occurring between them.

Factors influencing Crossing Over

a) Sex

- In male *Drosophila*, the crossing over is completely suppressed.
- However, in the male of mammals, **there is a tendency** of reduction of crossing over.

b)Temperature

- **Plough** demonstrated that when female *Drosophila* are subjected to high or low temperature, the percentage of the **crossing over is increased.**
- Age
- **Bridges** demonstrated in *Drosophila* that as the female **becomes older** the crossing over tends to increase.

c) X-radiations: The irradiations by x-rays and by radium **increase** the crossing over.

d) Nutritional effect

- **Levine** fed young *Drosophila* on a high calcium diet and found a **decrease** in crossing over.
- **Neel** found that the **larval starvation**s at certain age **increases** crossing over.

e) Distance between two genes

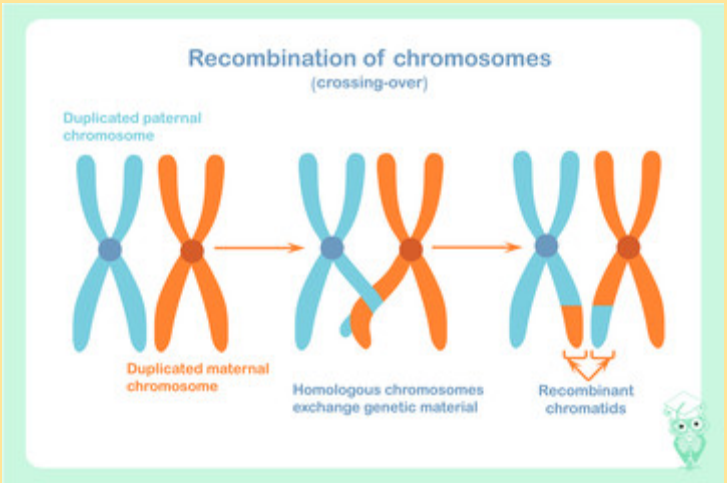
- The distance between genes is **the main factor** in crossing over.
- If two genes are distance apart, they have **greater chances** of crossing over than those which are closely set together.

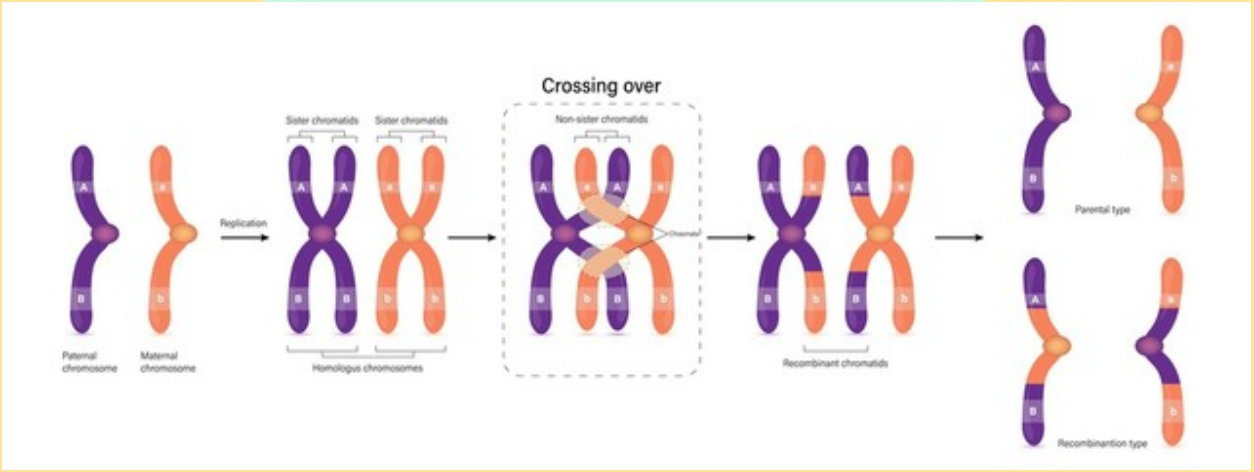
f) Mutant genes: Certain mutant genes increase the frequency of somatic crossing over.

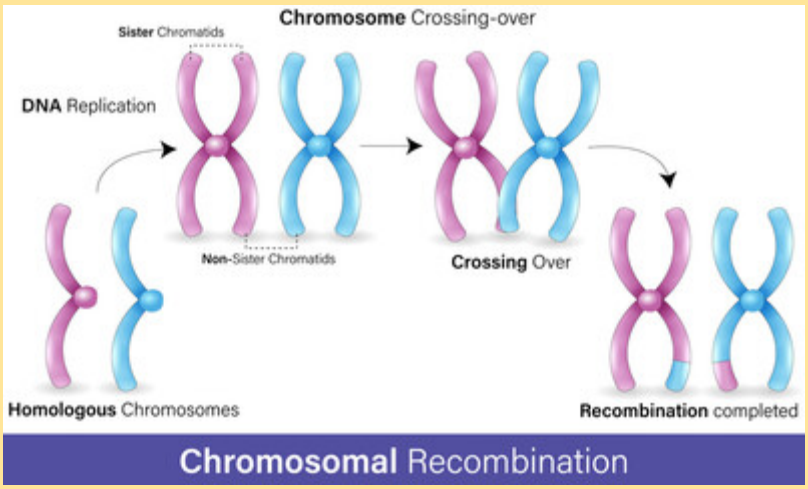
g) Near centromere and tips of chromosomes, crossing over is less frequent.

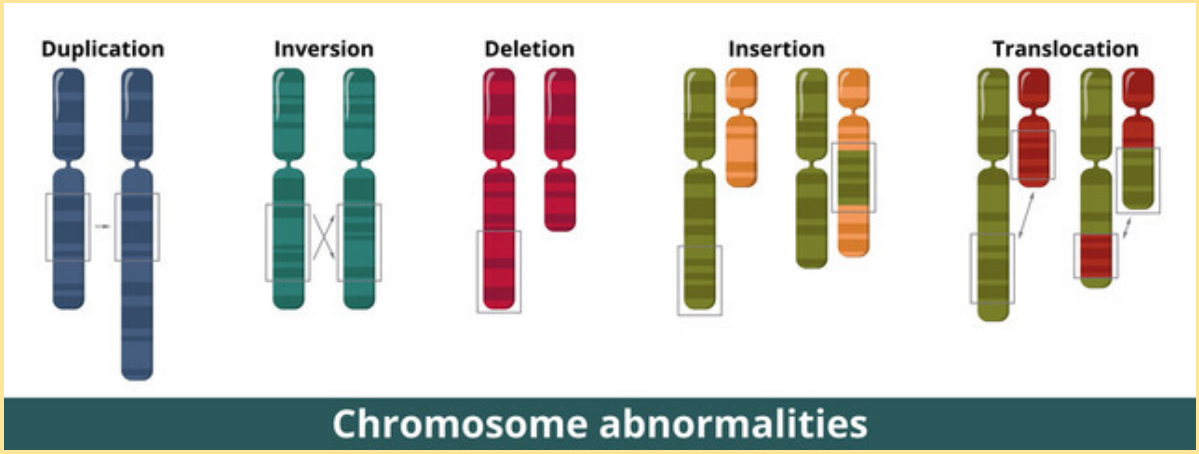
Significance of Crossing Over

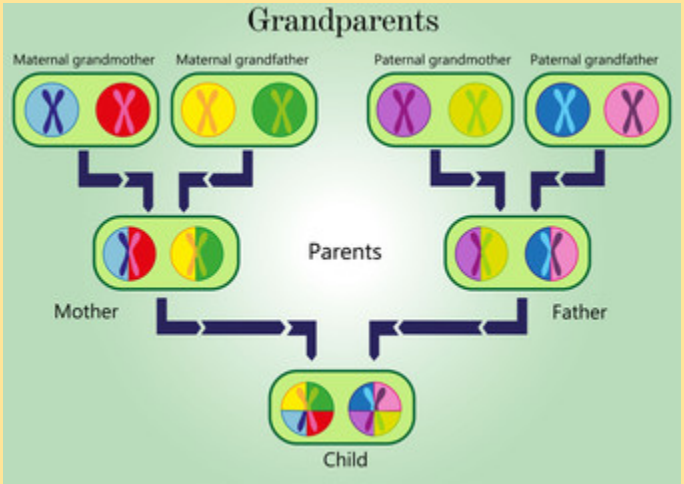
- It affords a proof for the **linear arrangement** of genes in the chromosome.
- As a result of crossing over, **new gene combinations** are produced, which play an important role in the **process of evolution**.
- It is necessary **for the natural selection**, because, due to this, the chances of **variation** increase.
- Due to crossing over, **useful recombinations can be formed** which might be used in breeding programmes.






Chromosome Crossing-over
Sister Chromatids
DNA Replication
Non-Sister Chromatids
Crossing Over
Homologous Chromosomes
Recombination completed
Chromosomal Recombination


Duplication
Inversion
Deletion
Insertion
Translocation
Chromosome abnormalities


Grandparents
Maternal grandmother
Maternal grandfather
Paternal grandmother
Paternal grandfather
Parents
Mother
Father
Child

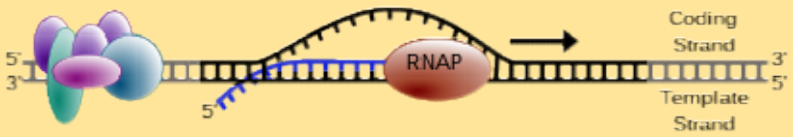
Gene expression

Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product that enables it to produce end products, proteins or non-coding RNA, and ultimately affect a phenotype. These products are often proteins, but in non-protein-coding genes such as transfer RNA (tRNA) and small nuclear RNA (snRNA), the product is a functional non-coding RNA. Gene expression is summarized in the central dogma of molecular biology first formulated by Francis Crick in 1958,[1] further developed in his 1970 article,[2] and expanded by the subsequent discoveries of reverse transcription[3][4][5] and RNA replication.

Mechanism

The process of gene expression is used by all known life—eukaryotes (including multicellular organisms), prokaryotes (bacteria and archaea), and utilized by viruses—to generate the macromolecular machinery for life.In genetics, gene expression is the most fundamental level at which the genotype gives rise to the phenotype, *i.e.* observable trait. The genetic information stored in DNA represents the genotype, whereas the phenotype results from the "interpretation" of that information. Such phenotypes are often displayed by the synthesis of proteins that control the organism's structure and development, or that act as enzymes catalyzing specific metabolic pathways.All steps in the gene expression process may be modulated (regulated), including the transcription, RNA splicing, translation, and post-translational modification of a protein. Regulation of gene expression gives control over the timing, location, and amount of a given gene product (protein or ncRNA) present in a cell and can have a profound effect on the cellular structure and function. Regulation of gene expression is the basis for cellular differentiation, development, morphogenesis and the versatility and adaptability of any organism. Gene regulation may therefore serve as a substrate for evolutionary change.

Transcription



The process of transcription is carried out by RNA polymerase (RNAP), which uses DNA (black) as a template and produces RNA (blue).

The production of a RNA copy from a DNA strand is called transcription, and is performed by RNA polymerases, which add one ribonucleotide at a time to a growing RNA strand as per the complementarity law of the nucleotide bases. This RNA is complementary to the template 3′ → 5′ DNA strand, with the exception that thymines (T) are replaced with uracils (U) in the RNA and possible errors.

In bacteria, transcription is carried out by a single type of RNA polymerase, which needs to bind a DNA sequence called a Pribnow box with the help of the sigma factor protein (σ factor) to start transcription. In eukaryotes, transcription is performed in the nucleus by three types of RNA polymerases, each of which needs a special DNA sequence called the promoter and a set of DNA-binding proteins—transcription factors—to initiate the process (see regulation of transcription below). RNA polymerase I is responsible for transcription of ribosomal RNA (rRNA) genes. RNA polymerase II (Pol II) transcribes all protein-coding genes but also some non-coding RNAs (*e.g.*, snRNAs, snoRNAs or long non-coding RNAs). RNA polymerase III transcribes 5S rRNA, transfer RNA (tRNA) genes, and some small non-coding RNAs (*e.g.*, 7SK). Transcription ends when the polymerase encounters a sequence called the terminator.

mRNA processing

While transcription of prokaryotic protein-coding genes creates messenger RNA (mRNA) that is ready for translation into protein, transcription of eukaryotic genes leaves a primary transcript of RNA (pre-RNA), which first has to undergo a series of modifications to become a mature RNA. Types and steps involved in the maturation processes vary between coding and non-coding preRNAs; *i.e.* even though preRNA molecules for both mRNA and tRNA undergo splicing, the steps and machinery involved are different. The processing of non-coding RNA is described below (non-coding RNA maturation).

The processing of pre-mRNA include 5′ *capping*, which is set of enzymatic reactions that add 7-methylguanosine (m^7G) to the 5′ end of pre-mRNA and thus protect the RNA from degradation by exonucleases. The m^7G cap is then bound by cap binding complex heterodimer (CBC20/CBC80), which aids in mRNA export to cytoplasm and also protect the RNA from decapping.

Another modification is 3′ *cleavage and polyadenylation*. They occur if polyadenylation signal sequence (5′- AAUAAA-3′) is present in pre-mRNA, which is usually between protein-coding sequence and terminator. The pre-mRNA is first cleaved and then a series of ~200 adenines (A) are added to form poly(A) tail, which protects the RNA from degradation. The poly(A) tail is bound by multiple poly(A)-binding proteins (PABPs) necessary for mRNA export and translation re-initiation. In the inverse process of deadenylation, poly(A) tails are shortened by the CCR4-Not 3′-5′ exonuclease, which often leads to full transcript decay.

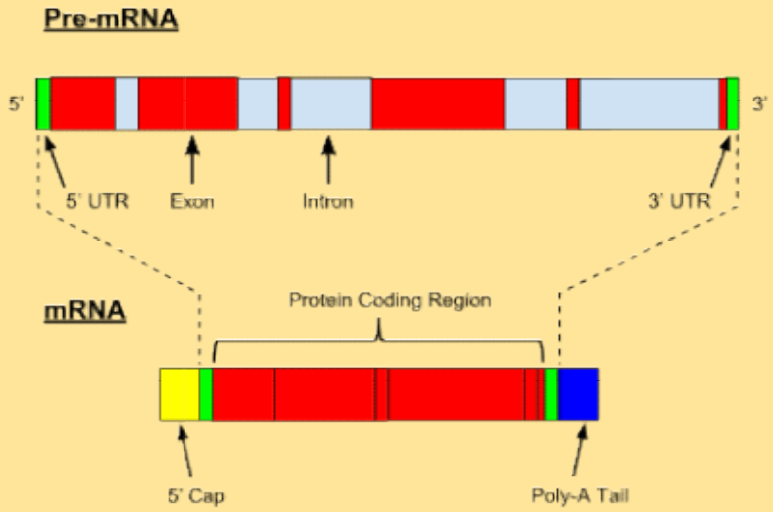


Illustration of exons and introns in pre-mRNA and the formation of mature mRNA by splicing. The UTRs (in green) are non-coding parts of exons at the ends of the mRNA.

A very important modification of eukaryotic pre-mRNA is *RNA splicing*. The majority of eukaryotic pre-mRNAs consist of alternating segments called exons and introns. During the process of splicing, an RNA-protein catalytical complex known as spliceosome catalyzes two transesterification reactions, which remove an intron and release it in form of lariat structure, and then splice neighbouring exons together. In certain cases, some introns or exons can be either removed or retained in mature mRNA. This so-called alternative splicing creates series of different transcripts originating from a single gene. Because these transcripts can be potentially translated into different proteins, splicing extends the complexity of eukaryotic gene expression and the size of a species proteome.

Extensive RNA processing may be an evolutionary advantage made possible by the nucleus of eukaryotes. In prokaryotes, transcription and translation happen together, whilst in eukaryotes, the nuclear membrane separates the two processes, giving time for RNA processing to occur.

Non-coding RNA maturation

In most organisms non-coding genes (ncRNA) are transcribed as precursors that undergo further processing. In the case of ribosomal RNAs (rRNA), they are often transcribed as a pre-rRNA that contains one or more rRNAs. The pre-rRNA is cleaved and modified (2′-*O*-methylation and pseudouridine formation) at specific sites by approximately 150 different small nucleolus-restricted RNA species, called snoRNAs. SnoRNAs associate with proteins, forming snoRNPs. While snoRNA part basepair with the target RNA and thus position the modification at a precise site, the protein part performs the catalytical reaction. In eukaryotes, in particular a snoRNP called RNase, MRP cleaves the 45S pre-rRNA into the 28S, 5.8S, and 18S rRNAs. The rRNA and RNA processing factors form large aggregates called the nucleolus.

In the case of transfer RNA (tRNA), for example, the 5′ sequence is removed by RNase P, whereas the 3′ end is removed by the tRNase Z enzyme and the non-templated 3′ CCA tail is added by a nucleotidyl transferase. In the case of micro RNA (miRNA), miRNAs are first transcribed as primary transcripts or pri-miRNA with a cap and poly-A tail and processed to short, 70-nucleotide stem-loop structures known as pre-miRNA in the cell nucleus by the enzymes Drosha and Pasha. After being exported, it is then processed to mature miRNAs in the cytoplasm by interaction with the endonuclease Dicer, which also initiates the formation of the RNA-induced silencing complex (RISC), composed of the Argonaute protein.

Even snRNAs and snoRNAs themselves undergo series of modification before they become part of functional RNP complex. This is done either in the nucleoplasm or in the specialized compartments called Cajal bodies. Their bases are methylated or pseudouridinilated by a group of small Cajal body-specific RNAs (scaRNAs), which are structurally similar to snoRNAs.

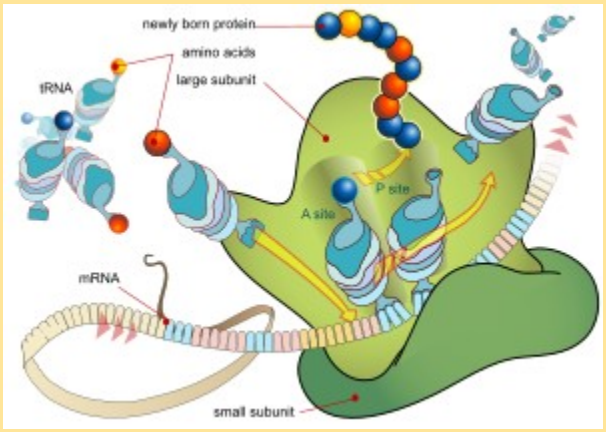
RNA export

In eukaryotes most mature RNA must be exported to the cytoplasm from the nucleus. While some RNAs function in the nucleus, many RNAs are transported through the nuclear pores and into the cytosol. Export of RNAs requires association with specific proteins known as exportins. Specific exportin molecules are responsible for the export of a given RNA type. mRNA transport

also requires the correct association with Exon Junction Complex (EJC), which ensures that correct processing of the mRNA is completed before export. In some cases RNAs are additionally transported to a specific part of the cytoplasm, such as a synapse; they are then towed by motor proteins that bind through linker proteins to specific sequences (called "zipcodes") on the RNA.

Translation

For some non-coding RNA, the mature RNA is the final gene product. In the case of messenger RNA (mRNA) the RNA is an information carrier coding for the synthesis of one or more proteins. mRNA carrying a single protein sequence (common in eukaryotes) is monocistronic whilst mRNA carrying multiple protein sequences (common in prokaryotes) is known as polycistronic.



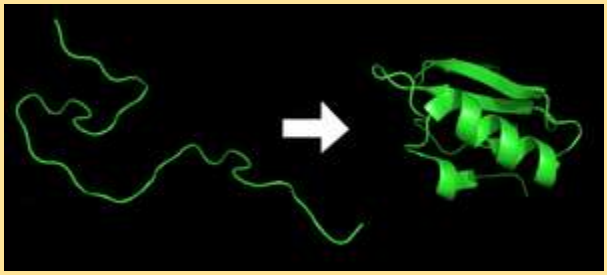
During the translation, tRNA charged with amino acid enters the ribosome and aligns with the correct mRNA triplet. Ribosome then adds amino acid to growing protein chain.

Every mRNA consists of three parts: a 5′ untranslated region (5′UTR), a protein-coding region or open reading frame (ORF), and a 3′ untranslated region (3′UTR). The coding region carries information for protein synthesis encoded by the genetic code to form triplets. Each triplet of nucleotides of the coding region is called a codon and corresponds to a binding site complementary to an anticodon triplet in transfer RNA. Transfer RNAs with the same anticodon sequence always carry an identical type of amino acid. Amino acids are then chained together by the ribosome according to the order of triplets in the coding region. The ribosome helps transfer RNA to bind to messenger RNA and takes the amino acid from each transfer RNA and makes a

structure-less protein out of it. Each mRNA molecule is translated into many protein molecules, on average ~2800 in mammals.

In prokaryotes translation generally occurs at the point of transcription (co-transcriptionally), often using a messenger RNA that is still in the process of being created. In eukaryotes translation can occur in a variety of regions of the cell depending on where the protein being written is supposed to be. Major locations are the cytoplasm for soluble cytoplasmic proteins and the membrane of the endoplasmic reticulum for proteins that are for export from the cell or insertion into a cell membrane. Proteins that are supposed to be produced at the endoplasmic reticulum are recognised part-way through the translation process. This is governed by the signal recognition particle—a protein that binds to the ribosome and directs it to the endoplasmic reticulum when it finds a signal peptide on the growing (nascent) amino acid chain.

Folding



Protein before (left) and after (right) folding

Each protein exists as an unfolded polypeptide or random coil when translated from a sequence of mRNA into a linear chain of amino acids. This polypeptide lacks any developed three-dimensional structure (the left hand side of the neighboring figure). The polypeptide then folds into its characteristic and functional three-dimensional structure from a random coil. Amino acids interact with each other to produce a well-defined three-dimensional structure, the folded protein (the right hand side of the figure) known as the native state. The resulting three-dimensional structure is determined by the amino acid sequence (Anfinsen's dogma).

The correct three-dimensional structure is essential to function, although some parts of functional proteins may remain unfolded. Failure to fold into the intended shape usually produces inactive proteins with different properties including toxic prions. Several neurodegenerative and other diseases are believed to result from the accumulation

of *misfolded* proteins. Many allergies are caused by the folding of the proteins, for the immune system does not produce antibodies for certain protein structures.

Enzymes called chaperones assist the newly formed protein to attain (fold into) the 3-dimensional structure it needs to function. Similarly, RNA chaperones help RNAs attain their functional shapes. Assisting protein folding is one of the main roles of the endoplasmic reticulum in eukaryotes.

Translocation

Secretory proteins of eukaryotes or prokaryotes must be translocated to enter the secretory pathway. Newly synthesized proteins are directed to the eukaryotic Sec61 or prokaryotic SecYEG translocation channel by signal peptides. The efficiency of protein secretion in eukaryotes is very dependent on the signal peptide which has been used.

Protein transport

Many proteins are destined for other parts of the cell than the cytosol and a wide range of signalling sequences or (signal peptides) are used to direct proteins to where they are supposed to be. In prokaryotes this is normally a simple process due to limited compartmentalisation of the cell. However, in eukaryotes there is a great variety of different targeting processes to ensure the protein arrives at the correct organelle.

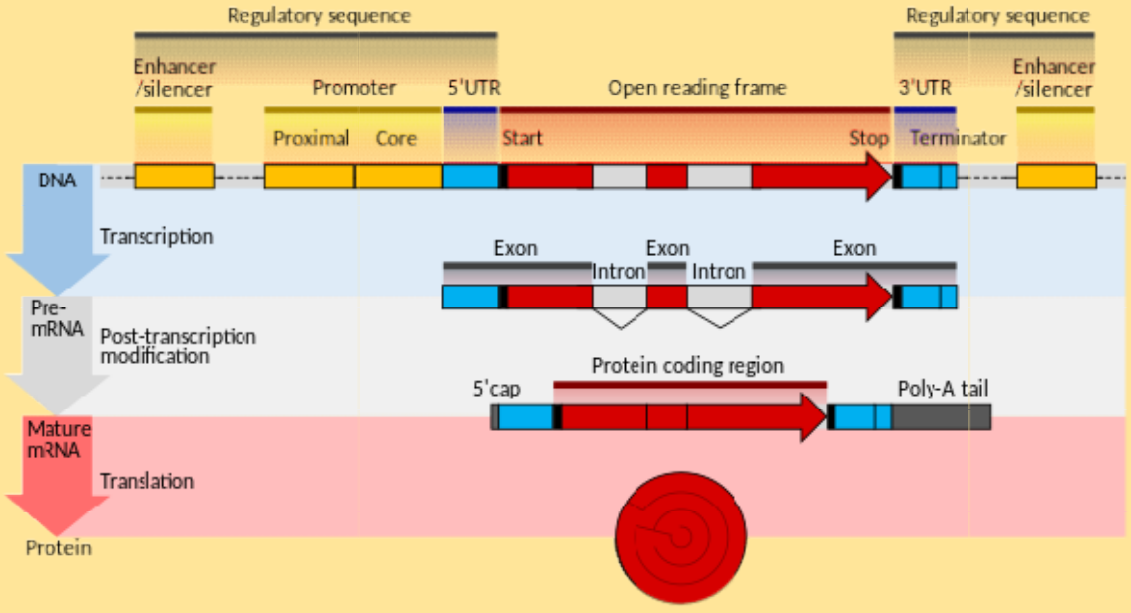
Not all proteins remain within the cell and many are exported, for example, digestive enzymes, hormones and extracellular matrix proteins. In eukaryotes the export pathway is well developed and the main mechanism for the export of these proteins is translocation to the endoplasmic reticulum, followed by transport via the Golgi apparatus.

Regulation of gene expression

Regulation of gene expression, or **gene regulation**, includes a wide range of mechanisms that are used by cells to increase or decrease the production of specific gene products (protein or RNA). Sophisticated programs of gene expression are widely observed in biology, for example to trigger developmental pathways, respond to environmental stimuli, or adapt to new food sources. Virtually any step of gene expression can be modulated, from transcriptional initiation, to RNA processing, and to the post-translational modification of a protein. Often, one gene regulator controls another, and so on, in a gene regulatory network.

Gene regulation is essential for viruses, prokaryotes and eukaryotes as it increases the versatility and adaptability of an organism by allowing the cell to express protein when needed. Although as early as 1951, Barbara McClintock showed interaction between two genetic loci, Activator (*Ac*) and Dissociator (*Ds*), in the color formation of maize seeds, the first discovery of a gene regulation system is widely considered to be the identification in 1961 of the *lac* operon, discovered by François Jacob and Jacques Monod, in which some enzymes involved in lactose metabolism are expressed by *E. coli* only in the presence of lactose and absence of glucose.

In multicellular organisms, gene regulation drives cellular differentiation and morphogenesis in the embryo, leading to the creation of different cell types that possess different gene expression profiles from the same genome sequence. Although this does not explain how gene regulation originated, evolutionary biologists include it as a partial explanation of how evolution works at a molecular level, and it is central to the science of evolutionary developmental biology ("evo-devo").



Lac Operon

Early insights into mechanisms of transcriptional regulation came from studies of *E. coli* by researchers Francois Jacob & Jacques Monod. In *E. coli,* and many other bacteria, genes encoding several different proteins may be located on a single transcription unit called an **operon**. The genes in an operon share the same transcriptional regulation, but are translated individually. Eukaryotes generally do not group genes together as operons (exception is *C. elegans* and a few other species).

Basic *lac* Operon structure

E. coli encounters many different sugars in its environment. These sugars, such as **lactose** and **glucose**, require different enzymes for their metabolism. Three of the enzymes for lactose metabolism are grouped in the ***lac* operon**: ***lacZ***, ***lacY***, and ***lacA*** (Figure 12.1.112.1.1). *LacZ* encodes an enzyme called **β-galactosidase**, which digests lactose into its two constituent sugars: glucose and galactose. *lacY* is a **permease** that helps to transfer lactose into the cell. Finally, *lacA* is a **trans-acetylase**; the relevance of which in lactose metabolism is not entirely clear. Transcription of the *lac* operon normally occurs only when lactose is available for it to digest. Presumably, this avoids wasting energy in the synthesis of enzymes for which no substrate is present. A single mRNA transcript includes all three enzyme-coding sequences and is called polycistronic. A cistron is equivalent to a gene.

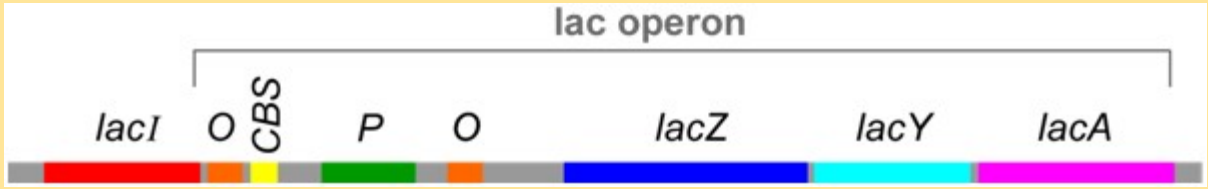


Figure 12.1.112.1.1: Diagram of a segment of an E. coli chromosome containing the lac operon, as well as the lacI coding region. The various genes and cis-elements are not drawn to scale. (Originall-Deyholos-CC:AN)

***cis*- and *trans*Regulators**

In addition to the three protein-coding genes, the *lac* operon contains short DNA sequences that do not encode proteins, but are instead binding sites for proteins involved in transcriptional regulation of the operon. In the *lac* operon, these sequences are called **P (promoter)**, **O (operator)**, and **CBS (CAP-binding site)**. Collectively, sequence elements such as these are called ***cis*-elements** because they must be located on the same piece of DNA as the genes they regulate. On the other hand, the proteins that bind to these *cis*-elements are called ***trans*-regulators** because (as diffusible molecules) they do not necessarily need to be encoded on the same piece of DNA as the genes they regulate.

***lacI* is an allosterically regulated repressor**

One of the major *trans*-regulators of the *lac* operon is encoded by ***lacI***. Four identical molecules of *lacI* proteins assemble together to form a **homotetramer** called a **repressor** (Figure 12.1.212.1.2). This repressor binds to two operator sequences adjacent to the promoter of the *lac* operon. Binding of the repressor prevents RNA polymerase from binding to the promoter (Figure 12.1.312.1.3). Therefore, the operon will not be transcribed when the operator is occupied by a repressor.

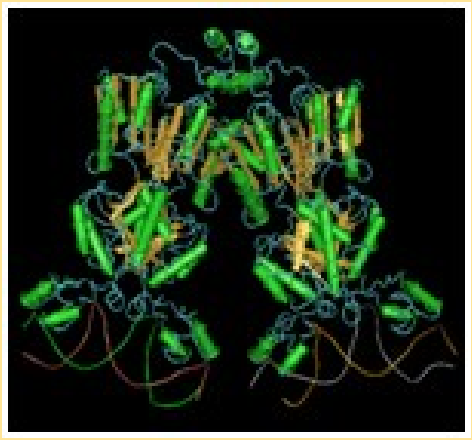


Figure 12.1.212.1.2: Structure of lacI homotetramer bound to DNA (Origianl-Deyholos-CC:AN)Besides its ability to bind to specific DNA sequences at the operator, another importantproperty of the *lacI* protein is its ability to bind to lactose. When lactose is bound to *lacI*, the shape of the protein changes in a way that prevents it from binding to the operator. Therefore, in the presence of lactose, RNA polymerase is able to bind to the promoter and transcribe the *lac* operon, leading to a moderate level of expression of the *lacZ*, *lacY*,

and *lacA* genes. Proteins such as *lacI* that change their shape and functional properties after binding to a ligand are said to be regulated through an **allosteric** mechanism. The role of *lacI* in regulating the *lac* operon is summarized in Figure 12.1.412.1.4

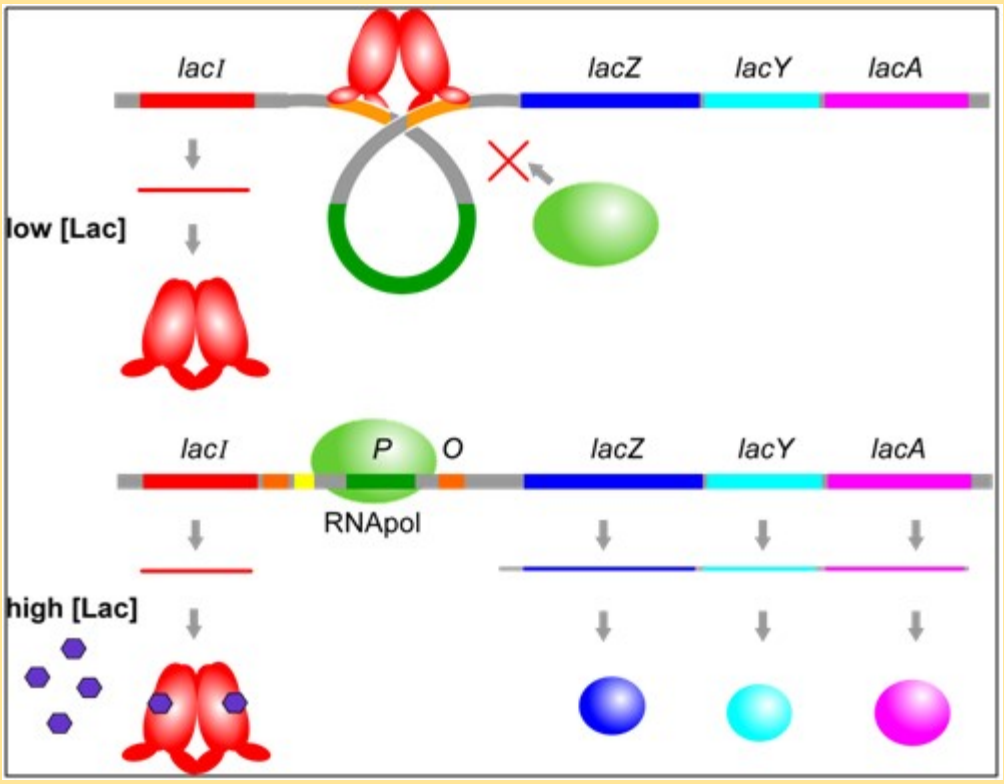


Figure 12.1.312.1.3: When the concentration of lactose [Lac] is low, lacI tetramers bind to operator sequences (O), thereby blocking binding of RNApol (green) to the promoter (P). Alternatively, when [Lac] is high, lactose binds to lacI, preventing the repressor from binding to O, and allowing transcription by RNApol. (Origianl-Deyholos-CC:AN)

CAP is an allosteric activator of the *lac* operon

A second aspect of *lac* operon regulation is conferred by a *trans*-factor called **cAMP binding protein** (**CAP**, Figure 12.1.412.1.4). CAP is another example of an allosterically regulated *trans*-factor. Only when the CAP protein is bound to cAMP can another part of the protein bind to a specific *cis*-element within the *lac* promoter called the **CAP binding sequence (CBS)**. CBS is located very close to the promoter (P). When CAP is bound to at CBS, RNA polymerase is better able to bind to the promoter and initiate transcription. Thus, the presence of cAMP ultimately leads to a further increase in *lac* operon transcription.

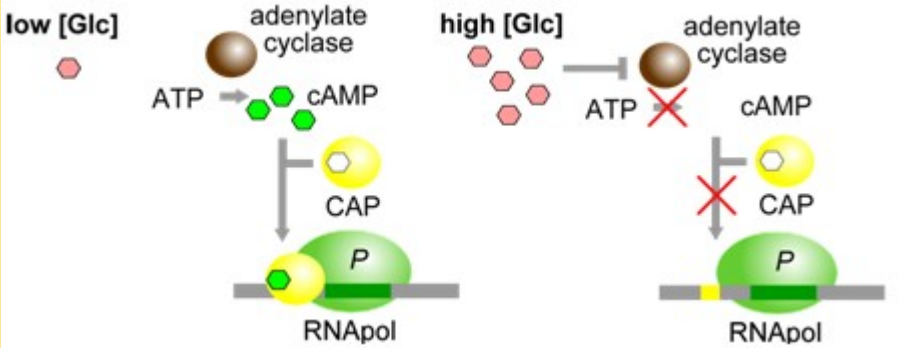
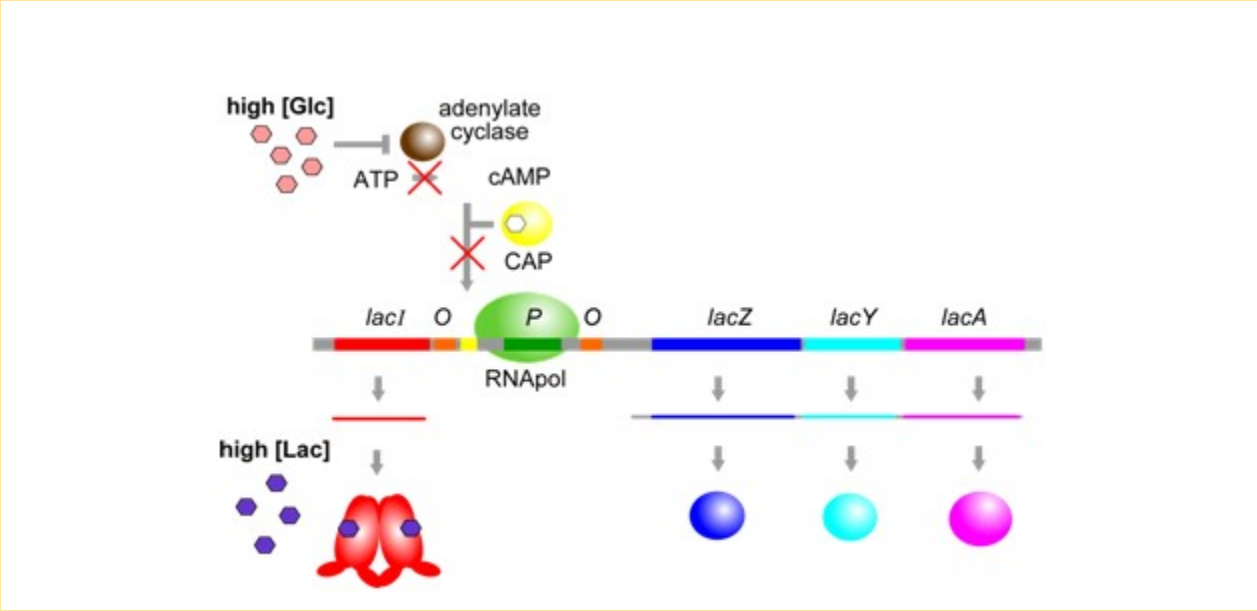


Figure 12.1.412.1.4: CAP, when bound to cAMP, helps RNApol to bind to the lac operon. cAMP is produced only when glucose [Glc] is low. (Origianl-Deyholos-CC:AN)

The physiological significance of regulation by cAMP becomes more obvious in the context of the following information. The concentration of cAMP is inversely proportional to the abundance of glucose: when glucose concentrations are low, an enzyme called **adenylate cyclase** is able to produce cAMP from ATP. Evidently, *E. coli* prefers glucose over lactose, and so expresses the *lac* operon at high levels only when glucose is absent and lactose is present. This provides another layer of logical control of *lac* operon expression: only in the presence of lactose, and in the absence of glucose is the operon expressed at its highest levels.



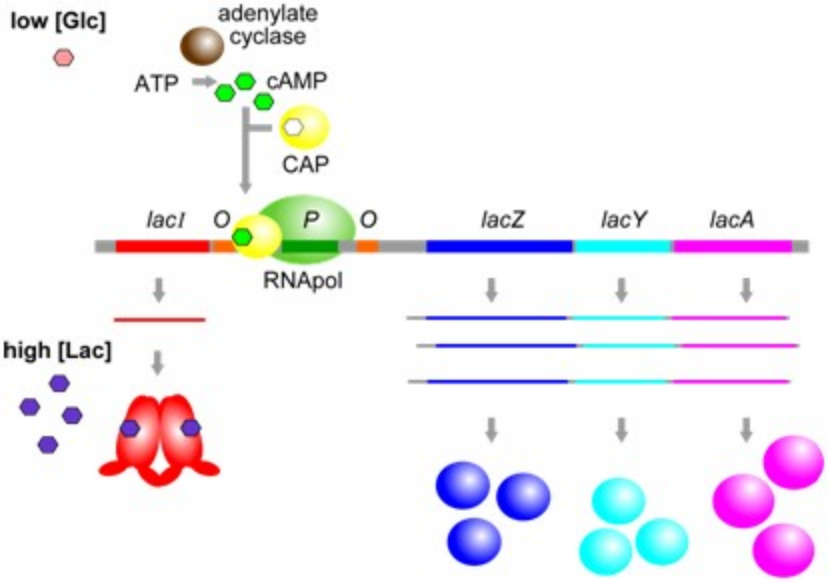


Figure 12.1.612.1.6: When glucose [Glc] and lactose [Lac] are both high, the lac operon is transcribed at a moderate level, because CAP (in the absence of cAMP) is unable to bind to its corresponding cis-element (yellow) and therefore cannot help to stabilize binding of RNApol at the promoter. Alternatively, when [Glc] is low, and [Lac] is high, CAP and cAMP can bind near the promoter and increase further the transcription of the lac operon. (Origianl-Deyholos-CC:AN)