**Population and Evolutionary Genetics**

**Unit 7. Population and Evolutionary Genetics (6 lectures)**

Allele frequencies, Genotype frequencies, Hardy-Weinberg Law, role of natural selection, mutation, genetic drift. Genetic variation and Speciation.

**Gene frequencies**

The proportions of different alleles of a gene present in a Mandalian population are known as gene frequency. Thus the proportions of gametes produced by a population carrying the different alleles of a gene are the frequencies of these alleles in the population, i.e. gene frequencies. Gene frequencies in a population can be readily estimated by classifying the individuals of a random sample from the population into different genotypic classes for a gene. The number of each of the alleles of the gene present in these individuals is then computed, and their ratio to the total number for all the alleles of the gene is estimated, these ratios is gene frequencies for that locus.

If a gene has two alleles, e.g. A and B, their frequencies may be represented by p and q for a generalized treatment. If p=0.75 and q=0.25, and p+q=(0.75+0.25)=1. This relationship between p and q viz, p+q =1, will always hold good due to their definition. Therefore, 1-q=p, 1-p=q and 1-(p+q)=0. Thus if the frequency of one allele of a gene is known, the frequency of the other allele is readily estimable (=1-p or1-q).

**Genotype frequencies**

The proportions of different genotypes for a gene in a population are known as genotypic frequencies for that gene; often they are also called zygotic frequencies. Genotypic frequencies, like gene frequencies are estimated from samples drawn from random mating populations. The proportion of a genotype in a sample will be the ratio of the number of individuals having that genotype to the total number of individuals in the sample. Thus the genotypic frequencies for AB will be as follows: AB, 0.60(=60/100); AB, 0.30(=30/100); and AB, 0.10(=10/100). Gene frequencies are readily estimated from genotypic frequencies using them in place of the numbers of individuals in different genotypic classes.

**Factors affecting gene frequencies**

The equilibrium in random mating population is affected by

1. **Migration** – migration is the movement of individual into a population from a different population. Migration may introduce new alleles into the population or may change the frequencies of existing alleles. Migration may lead to change in gene frequency depending upon the following
2. The proportion migrant individuals to the total (after migration) of individual in the population, this is called migration.
3. The difference between the gene frequencies of the migrant individuals and the population into which migration has occurred.
4. **Mutation**- Mutation will destroy the existing alleles and creates new ones. This process is too slow to have an effect in the single generation needed to achieve Hardy-Weinberg equilibrium. Therefore in breeding population such effect may be ignored.
5. **Random drift** – genetic drift means loss of genes from a small population due to close breeding. In a small population repeated interbreeding involves in the same gene pole. There is no chance of in flow of new genes. Moreover there is sufficient chance of none of the individuals carrying particular allele may mat and reproduce successfully. The progenies of this individual will be totally devoid of that particular allele. In this way many alleles may be lost gradually from the population or particular alleles may proportionately increase. Such variation in gene frequency is called genetic drift. Due to genetic drift one of alleles becomes zero and that of the other allele become 1. The allele with the frequency of 1 is said to be fixed in the population because there would be no further change in its frequency. It may be expected that in a small population all the genes would become homozygous or would be fixed in due course of time. Breeding population are generally small, hence a certain amount of drift except to use very large population which is often not practicable.

Drift is important in 3 cases-

1. It is important for removing or promoting very rare alleles
2. Drift is responsible for changing the frequency of neutral mutation.
3. Drift in small population can produce unrepresentative allele frequencies which may be occur in a large population.
4. **Selection** – selection is defined as differential rate of reproduction of different genotype in a population. In crop improvement, selection is very important because it allows the selected genotypes to reproduce, while the undesirable genotypes are eliminated. Thus selection differentiate survival of different alleles during gametogenesis or during embryogenesis and development. The relative strength varies with the amount of advantage available. The possibility that a particular phenotype will survive is a measure of its fitness. Thus fitness refers to total reproductive potential or efficiencies.

**Hardy-Weinberg Law**

English mathematician, G.H. Hardy and a German physician W. Weinberg in 1908 discovered a principle concerning the gene frequencies in a population and this principle come to be known as **Hardy Weinberg law**. According to Mendalian inheritance in controlled mating in plants and animals, is found the Mendalian ratio 3:1. But random among a population showed different gene frequencies resulting in different phenotypes. Hardy and Weinberg showed that there is equilibrium between frequencies of gene in a population. The relative frequencies of each allele tend to remain constant generation after generation. This mathematical relation is now called **Hardy-Weinberg theorem or law**. The law states that the relative frequencies of different genes in a large panmictic population, after random mating tend to remain constant, generation after generation in the absence of mutation, selection, genetic drift and gene flow.

The following conditions must be considered in Hardy Weinberg law

1. The population must be infinitely large or at least large enough that sampling error is negligible.
2. Mating with the population occurs at random.
3. There must not be any selective advantage for any genotype so that all genotypes produced by random mating are equally viable and fertile
4. There should not be mutation, migration, selection and genetic drift.

To explain the law let us presumed that ‘p’ represent the frequency of a dominant allele (A) and ‘q’ the frequency of a recessive allele (a). As the sum of these frequencies represent 100% of the population, p+q=1. If in a population, 70% of the alleles are ‘A’ and 30% are ‘a’ than p=0.7 and q=0.3. In a checker board, the distribution of genotypes produced by random mating in the next generation can be expressed as

**P2+2pq+q2=1**

**This formula is known as Hardy-Weinberg formula or binomial expression.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gamete or allele frequency | | |
| Gamete or allele frequency |  | A (p=0.7) | a (q=0.3) |
| A (p=0.7) | AA P2=(0.49) | Aa pq=(0.21) |
| a (q=0.3) | Aa pq=(0.21) | aa q2=(0.09) |

In the new generation, the frequency of ‘A’ is P2+1/2(2pq) and the frequency of ‘a’ is q2+1/2(2pq).

For ‘a’ the frequency is q2+1/2(2pq)= 0.09+1/2(0.42)=0.3

For ‘A’ the frequency is p2+1/2(2pq)= 0.49+1/2(0.42)=0.7

Applying Hardy-Weinberg law= P2+2pq+q2=(0.7) 2+2 (0.7)( 0.3)+ (0.3) 2=0.49+0.42+0.09=1

A population in which a gene frequency remain constant generation after generation is said to be in a **genetic equilibrium** for that gene. In the above case as the frequencies of gene ‘A’ and ‘a’ remain constant, this condition listed Hardy Weinberg law is held true in the above population.

**Hardy Weinberg law and multiple allele**

When more than 2 alleles are present at a locus, the equilibrium condition is described by multinomial expression, (p+q+r+….)2 in which each later represents the gene frequencies p, q, r would we expect 6 genotypes at Hardy Weinberg equilibrium in the proportion p2A1A1, 2pqA1A2, 2prA1A3, q2A2A2, 2qrA2A3 and r2A3A3. If all the alleles are codominant with each other, each genotype has its own distinct phenotype and the gene frequencies are easily scored.

**Hardy-Weinberg law and sex linked alleles**

Sex linked characters show frequencies different from autosomal characters because the genes are located in the X-chromosomes. Since females have two X chromosomes and males one, it is evident that there can be never be a gene frequency of 11/2 as can exist for autosomal trends. The gene frequency can not only be one of the following, one, 2/3, 1/3 or none.

**Role of natural selection**

It is a process by which the forms of organisms in a population that are best adapted to the environment increase in frequency relative to less adapted forms over a number of generations. Darwin was the first to popularize the idea that speciation could readily occur through the prolonged action of natural selection, the process by which genetic variants that are better suited to the natural environment increase in frequency while variants that are not well suited decrease. It is predictable (deterministic), non-random process.

The fate of natural selection may be positive (or directional), negative (or purifying) and balancing selection. The genetic changes that confer a higher fitness tend to increase in frequency over time in the population. This evolutionary process is called positive selection. Conversely, genetic changes that decrease the organism’s fitness tend to disappear from populations through a process known as negative selection. Finally, it may happen that a mutation is advantageous only in heterozygotes but not in homozygotes. Such alleles tend to be maintained at an intermediated frequency in populations by way of the process known as balancing selection.

Variations neither useful nor injurious (i.e. do not affect the fitness of individuals) would not be affected by natural selection. The fate of these variations is essentially driven by genetic drift.

**Mutation**

This theory was put forth by Hugo de vries (1840-1935), a Dutch Botanist, who described a large number of discrete variations in *Oenothera lamarckiana*. He coined the term mutation for the phenomenon of appearance of sudden heritable changes, the various mutations were called new and separate species by him. He suggested that new species could arise in a single step due to mutation and natural selection was not important for such an evolution. Further, since mutations are random, it was assumed that evolution was also random, i.e., it did not proceed in a direction.

One of the important consequences of natural selection is that populations are expected to become progressively more and more adapted to their environment. Mutationsts considered that populations were preadapted and that adaptations did not arise due to natural selection.

It is now universally accepted that mutations are the ultimate source of all genetic variation present in the biological world, but that natural selection is required to bring about the evolution of new living forms.

Spontaneous mutation is the primary source of all genetic variations. Mutation is considered as a random phenomenon. The randomness of mutations was first demonstrated by Luria and Delbruck in 1943. The randomness of mutations is an important concept in biology because it is a requirement of the Darwinian view of evolution which holds that changes in the characteristics of an organism occur by chance and are not influenced by the environment in which the organism is placed. In contrast, the Lamarckian theory of evolution, which biologists rejects well over a century ago, states that organisms acquire changes that enable them to adapt to their environment. The Darwinian view requires that mutations occur at random, whereas Lamarckian evolution demands that adaptive or directed mutations occur in response to the environment.

**Genetic drift**

The Hardy-Weinberg principle states that in a very large populations, the allele frequencies remain constant from one generation to the next unless the equilibrium is distributed by migration, mutations or natural selection. However, in finite populations, random sampling can cause an existing allele to lost (not replaced). This causes of random change in allele frequencies. The random sampling change in allele frequencies simply as a result of chance from one generation to next in a finite population is called genetic drift. The concept of genetic drift was first introduced by Sewall Wright.

Because all populations are finite, alleles at all loci are potentially subject to random genetic drift. While it occurs in all populations, drift can be a major driving force for changing allele frequencies in small populations. The smaller the population, the more extreme it tend to be. The actual change in allele frequency caused by genetic drift is random for any given generation; however, the effects of drift will accumulate over time. Like selection, drift is a process of differential reproductive success; however, the key element of genetic drift is that which individuals survive and reproduce is random (unrelated to phenotype and genotype).

The consequences of genetic drift are numerous. It leads to random changes in allele frequencies. Drift causes fixation of alleles. When an allele reaches a frequency of 1, it is said to be ‘fixed’ in the population and when an allele reaches a frequency 0, it is lost. Once an allele becomes fixed, genetic drift comes to a halt, and the allele frequency cannot change unless a new allele is introduced in the population viz mutation or gene flow. Thus even while genetic drift is a random, directionless process, it acts to eliminate genetic variation over time. Drift leads to an increase in homozygosity for diploid organisms and causes as increase in the inbreeding coefficient. Genetic drift also has significant long term evolutionary consequences. It can facilitate speciation (creation of new species).

Although both genetic drift and natural selection affect evolution, genetic drift operates randomly while natural selection functions non-randomly. While natural selection has a direction, guiding evolution towards in that direction that increase survival and reproductive fitness, genetic drift has no direction and is guided only by chance. As a result, drift acts upon the genotypic frequencies within a population without regard to their phenotypic effects. In contrast, selection favours the spread of alleles whose phenotypic effects increase survival and reproductive fitness of their carriers, lowers the frequencies of alleles that cause unfavourable traits, and ignores those that are neutral. Certain circumstances can result in genetic drift having a significant impact on a population.

**Genetic bottlenecks**

A population bottleneck is an evolutionary event when a population drastically reduces to a significantly smaller size over a short period of time. The bottleneck may be caused by various event, such as an environmental disaster, the hunting of a species to the point of extinction or habitat destruction that results in the deaths of organisms. The bottleneck can result in radical changes in allele frequencies, completely independent of selection. When a population will be reduced to a very small number, the genetic diversity of the small population may be much less than the previous population. By chance alone, certain alleles may be overrepresented among the survivors, others may be under represented and some may be absent altogether which means that the population as a whole has few genetic characteristics. Following a population bottleneck, the remaining population faces a higher level of genetic drift, which describes random fluctuations in the presence of alleles in a population. In small populations, infrequently occurring alleles face a greater chance of being lost, which can further decrease the gene pool. Due to the loss of genetic variation, the new population can become genetically distinct from the original population, which has led to the hypothesis that population bottelenecks can lead to the evolution of new species.

**Founder effect**

A founder effect occurs when a new population (new colony) develops by a few members (known as founder population) of the original large population. This is essentially a sampling difference; the larger the founder population the more likely it will resemble the original (parent) population. The small number of individuals that act as founders might not possess all the alleles present in the parent population, and might display different frequency of those that they do possess. The founder effect can result in radical changes in allele frequencies. The founder population may have reduced genetic variation from the original population.

A classic example is the human population founded on Pitcairn Island by several of Bounty mutineers and some Polynesians. The unique combination of Caucasian and Polynesian traits that characterizes today’s Pitcairn Island population resulted from the small number of founders of the population. Similarly, the Afrikaner population has an unusually high frequency og the gene that causes Huntington’s disease, because those original Dutch colonists happened to carry that gene with unusually high frequency. This effect is easy to recognize in genetic diseases, but of course, the frequencies allsorts of genes are affected by founder events.

**Genetic variation**

The extent of genetic variation present in a population can be estimated at several levels, some of which are as follows: 1) phenotype, 2) chromosomes (generally rearrangements), 3) proteins and enzymes (electrophoretic variation), 4) amino acid sequences of specific proteins, 5) nucleotide sequences of a specific genes and 6) nucleotide sequences of the entire genomes.

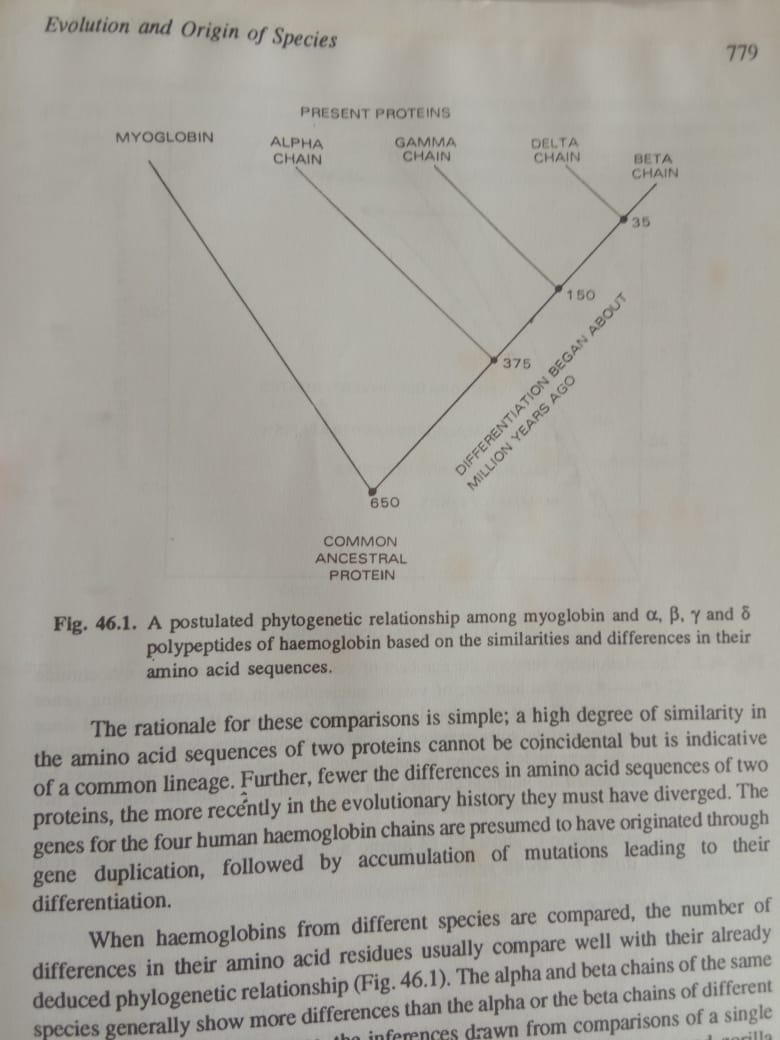
**Phenotypic variation**

Although members of a species show remarkable similarity in their phenotypes, they carry a large variety of gene mutations. For example, it was estimated that only 17% of the flies in natural populations of *Drossophila persimilis* are free of lethal mutations. Another indication on concealed genetic variability in natural populations is the existence of various chromosome rearrangements; these are widespread in many species, e.g. grasshoppers, snails, *Oneothera*, *Datura*, *Drosophila* etc. a much larger amount of genetic variablity is indicated from the studies on electrophoretic patterns of various proteins and enzymes. It has been estimated that various species ranging from man to Drosophila may exhibit polymorphism from 28% (man) to 81% (*Drosophila willistoni*) of their proteins.

The remarkable uniformity for various characters of the members of a species is proposed to be due to the canalization of development of those characters that are associated with fitness, canalization is the phenomenon by which the development of a character is normally unaffected by common variations in genotype as well as environment. Canalization is disrupted by environmental shocks and by some mutant genes. Further genotypes with increased or decreased canalization may be selected, such a selection is called canalization selection. Thus the canalization of development of characters will conceal a considerable amount of genetic variability for the trait since this variability will not be expressed at the phenotypic level.

**Amino acid sequences of specific proteins**

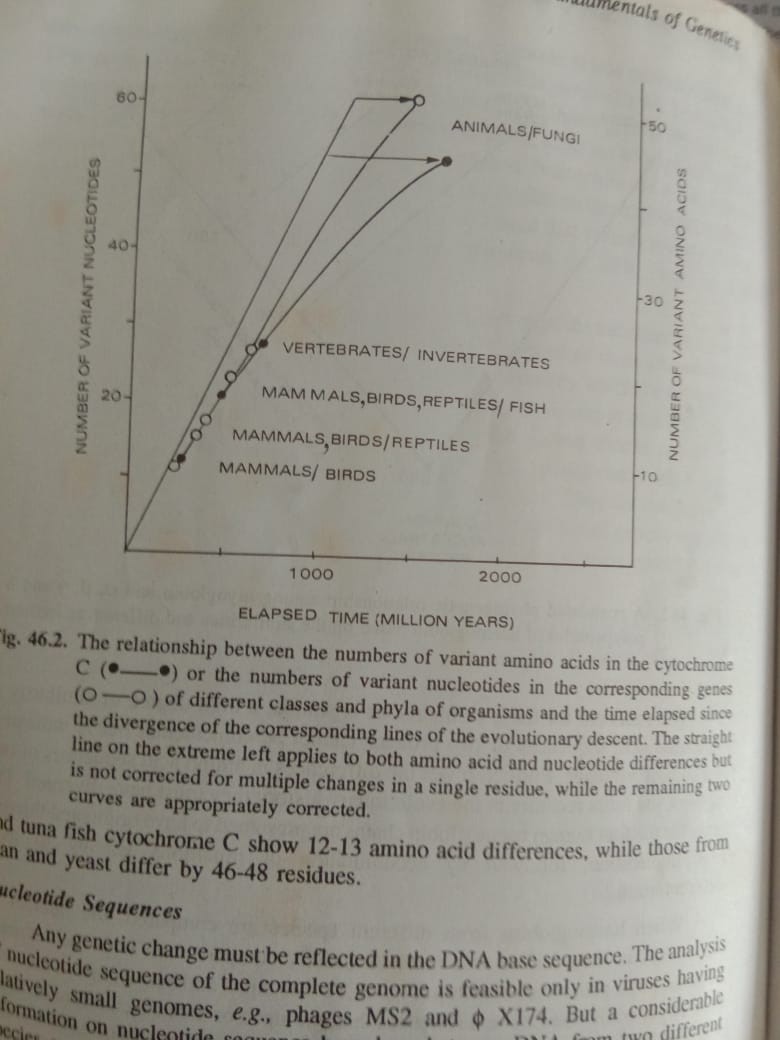
Proteins are basic to the development of various characters. Therefore phenotypic divergence among species must be based in the diversity of their proteins. Amino acid sequences of several different proteins from different taxa have been compared to asses their relatedness. One such protein is the haemoglobin, which is proposed to have originated from the muscle protein myoglobin. A comparison of the amino acid sequences of α, β, γ and δ polypeptides of human haemoglobin indicates that they all originated from the same ancestral protein. The α and β chains appear to have differentiated more recently.



The rationale for these comparisons is simple; a high degree of similarity to the amino acid sequences of two proteins cannot be coincidental but is indicative of a common lineage. Further, fewer the differences in amino acid sequences of two proteins, the more recently in the evolutionary history they must have diverged. The genes for the four human haemoglobin chains are presumed to have originated through gene duplication, followed by accumulation of mutations leading to their differentiation.

When haemoglobin from different species are compared, the number of differences in their amino acid residues usually compare well with their already deduced phylogentic relationship. The alpha and beta chains of the same species generally show more differences than the alpha or beta chains of different species. But in many instances, the inferences drawn from comparisons of a single protein may be highly misleading. For example, haemoglobin of man and gorilla differ by a single amino acid, which is the extent of the difference exhibited by many mutant human haemoglobins.

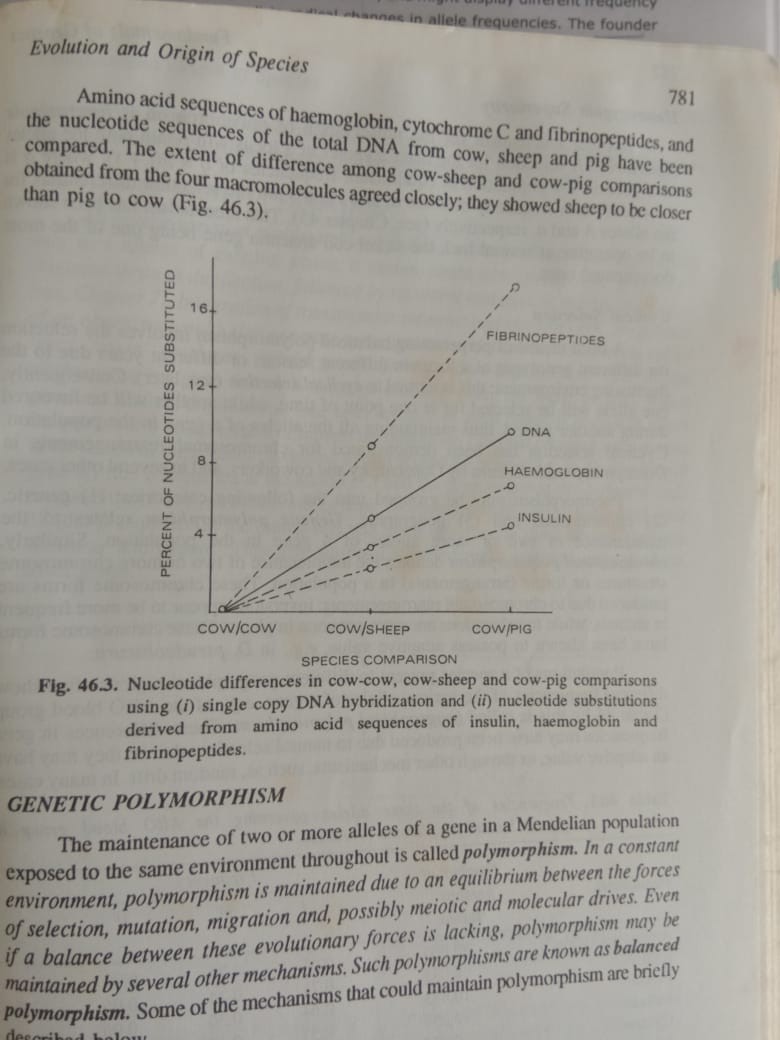
Cytochrome C is another extensively investigated protein. Cytochrome C has amino acid residues. A comparison of amino acid sequences of cytochrome C from different organisms shows that, as in the case of haemoglobin, the degree of their relatedness so estimated agrees well with their phylogenetic relationships. For example, human and tuna fish cytochrome C show 12-13 amino acid differences, while those from man and yeast differ by 46-48 residues.

****

**Nucleotide sequences**

Any genetic change must be reflected in the DNA base sequence. The analysis of nucleotide sequence of the complete genome is feasible only in viruses having relatively small genomes, e.g., phages MS2 and фX174. But a considerable information on nucleotide sequence homology between DNA from two different species, say X and Y can be obtained through DNA hybridization. In such studies, the total DNA from one species, say X, is radioactively labelled. The two DNA samples (designated as Xand Y) are denatured, mixed and allowed to renature. Preferential formation of labelled X-Y heteroduplexes is ensured by using a small amount of X relative to Y. the X-Y hybrid DNA is easily identified since it is radioactive; these heteroduplexes are isolated and their melting point is determined. It has been found that for each1% difference in nucleotide sequence of X and Y DNA, the melting point of their hybrid (X:Y DNA) molecules is reduced by 1.6°C.

Amino acid sequences of haemoglobin, cytochrome C and fibrinopeptides and the nucleotide sequences of the total DNA from cow, sheep and pig have been compared. The extent of difference among cow-sheep and cow-pig comparisons obtained from the four macromolecules agreed closely, they showed sheep to be closes than pig to cow.



**Speciation.**

Speciation is a lineage-splitting evolutionary process that produces two or more separate species. It refers to the creation of new and distinct biological species by branching off from the ancestral population. It occurs when gene flow is reduced sufficiently between sister populations to allow each to become irrevocably committed to different evolutionary lineages.

Reduced gene flow probably plays critical role in speciation. Various modes have been presented where by a single evolutionary lineage splits into two or more genetically independent lineages due to reduced gene flow. Modes of speciation are often classified as allopatric, parapatric and sympatric speciation.

**Allopatric speciation**

In 1963, Ernst Mayr contended that speciation takes place through **allopatry** (or geographic isolation). **Allpatric speciation** (**allo=other and patric= place**) is genetic divergence permitted by geographic isolation. Tthis process is dependent on random mutation, which accumulates steadily after a population has been subdivided in space. This subdivision can take place because of geographic barriers such as mountain ranges and water bodies. Because the two sub-populations are then reproductively isolated, each of the populations accumulates in different mutations and the two populations diverge.

This divergence is the strongest where the population size of one or both the sub-populations is small because sampling error is high in smaller populations. These differences will eventually lead to reproductive incompatabilities that will be keeping the populations distinct. One well-known example of allopatric speciation involves divergence of *Drosophila* populations in the Hawalian archipelago.

Reproductive isolating mechanisms usually originate incidentally in the speciation process. That is, they arise incidentally during the processs of evolution in isolated populations rather than being selected for. When isolated populations come together again, incomplete isolating mechanisms may allow hybrids to form. If the hybrids are normal, viable and can freely interbreed with individuals of each parent population, then no speciation has taken place. However, if the hybrids are at a disadvantage, natural selection may favour stronger isolating mechanisms. In this case, organisms that mate with individuals from the other population leave fewer offspring. The result is a more effective barrier to hybridization. Regions in which previously isolated populations come into contact and produce hybrids are called hybrid zones.

A simplified flow chart of allopatric speciation



**Sympatric speciation**

Sympatric speciation (sym- same and patric- place), conversely, does not require geographic isolation; instead, it relies on the development of reproductive isolation mechanisms to allow divergence of the two sub-populations. This reproductive isolation can be a result of a number of factors, including competition for resources, disruptive selection and sexual selection. In sympatric selection, selection acts against individuals of an intermediate type, either through decreased viability, or decreased fecundity. This creates the evolution of mating preferences and other mechanisms that result in decreased hybridization. In sympatric speciation, selection enhances traits that promote the divergence of the populations. This is fundamentally different from allopatric speciation, where mutation passively leads to differences that allow divergence of the populations.

Probably the best known example of sympatric speciation is the divergence of *Rhagoletis* *pomonella*, the maggot fly. This species has recently diverged into two subspecies due to the introduction of apple trees in the northeastern United States, where hawthorn trees were native. The flies used the hawthorn fruits to reproduce and to lay their eggs. The introduction of the apple trees provided a more nutritious food source for developing maggots, as well as an escape from parasitic wasps. The difference between the two type of fruits, such as maturation timing, allowed for the evolution of isolation mechanisms and subsequent divergence of the two subspecies.

A simplified flow chart of sympatric speciation



**Parapatric speciation**

Parapatric speciation (para=beside, patric=place) is a form of speciation in which the evolution of reproductive isolating mechanisms occurs when a population enters a new niche or habitat within the range of parent species. Generally, this occurs when there has been a drastic change in the environment within the original species habitat.

In parapatric speciation, there is no specific extrinsic barrier to gene flow. The population is continuous, but nonetheless, the population does not mate randomly. Individuals are more likely to mate with their geographic neighbours than with individuals in a different part of the population’s range. In this mode divergence may happen because of reduced gene flow within the population and varying selection pressures across the population range.

