NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF SCIENCE AND TECHNOLOGY

COURSE CODE: BIO 301

COURSE TITLE: GENETICS II

BIO 301 GENETICS II (2 UNITS)

Course Outline:

1. Population genetics
2. Cytogenetics
3. Variation in plants and animals
4. Microbial genetics
5. Biochemical and biomedical genetics
6. Human genetics.
7. Further consideration of various deviations from basic principles
8. Pedigree analysis and
9. Gene interactions.

Course Objective:

At the end of this course, students should be familiar with an introduction to population genetics, cytogenetics, microbial genetics, variation in plants and animals, biochemical and biomedical genetics, human genetics, deviations from basic principles, pedigree analysis and gene interactions.

1.0. Population Genetics

1.1. Introduction

Population genetics is the study of the frequency and interaction of alleles and genes in populations under the influence of the four main evolutionary processes: natural selection, genetic drift, mutation and gene flow. It also takes into account the factors of recombination, population subdivision and population structure. A sexual population is a set of organisms in which any pair of members can breed together. This implies that all members belong to the same species and live near each other. Population genetics attempts to explain such phenomena as adaptation and speciation.

For example, all of the moths of the same species living in an isolated forest are a population. A gene in this population may have several alternate forms, which account for variations between the phenotypes of the organisms. An example might be a gene

for coloration in moths that has two alleles: black and white. A gene pool is the complete set of alleles for a gene in a single population; the allele frequency for an allele is the fraction of the genes in the pool that is composed of that allele (for example, what fraction of moth coloration genes are the black allele). Evolution occurs when there are changes in the frequencies of alleles within a population; for example, the allele for black color in a population of moths becoming more common.

Population genetics began as a reconciliation of the Mendelian and biometrician models. A key step was the work of the British biologist and statistician R.A. Fisher. In a series of papers starting in 1918 and culminating in his 1930 book *The Genetical Theory of Natural Selection*, Fisher showed that the continuous variation measured by the biometricians could be produced by the combined action of many discrete genes, and that natural selection could change allele frequencies in a population, resulting in evolution. In a series of papers beginning in 1924, another British geneticist, J.B.S. Haldane worked out the mathematics of allele frequency change at a single gene locus under a broad range of conditions. Haldane also applied statistical analysis to real-world examples of natural selection, such as the evolution of industrial melanism in peppered moths, and showed that selection coefficients could be larger than Fisher assumed, leading to more rapid adaptive evolution.

The American biologist Sewall Wright, who had a background in animal breeding experiments, focused on combinations of interacting genes, and the effects of inbreeding on small, relatively isolated populations that exhibited genetic drift. In 1932, Wright introduced the concept of an adaptive landscape and argued that genetic drift and inbreeding could drive a small, isolated sub-population away from an adaptive peak, allowing natural selection to drive it towards different adaptive peaks.

The work of Fisher, Haldane and Wright founded the discipline of *population genetics*. This integrated natural selection with Mendelian genetics, which was the critical first step in developing a unified theory of how evolution worked. John Maynard Smith was Haldane's pupil, whilst W.D. Hamilton was heavily influenced by the writings of Fisher.

The American George R. Price worked with both Hamilton and Maynard Smith. American Richard Lewontin and Japanese Motoo Kimura were heavily influenced by Wright.

1.2. Goals of Population Genetics

1. To describe how the frequency of an allele which controls a trait changes over time
2. To analyze the factors that lead to changes in gene (allele) frequencies
3. To determine how changes in gene (allele) frequencies affect evolution and speciation

1.3. Why Study Populations and Gene Frequencies

- Genetic variability is necessary for evolutionary success
- Measuring genetic variability at many loci can characterize a population
- Variability of phenotypic and molecular traits are analyzed

1.4. Variability and Gene (or Allelic) Frequencies

- Genetic data for a population can be expressed as gene or allelic frequencies
- All genes have at least two alleles
- Summation of all the allelic frequencies for a population can be considered a description of the population
- Frequencies can vary widely among the alleles in a population
- Two populations of the same species do not have to have the same allelic frequencies

1.5. Deriving Genotypic Frequencies

Genotypic frequencies - describes the distribution of genotypes in a population. **Example**: blood type locus with two alleles , *M* or *N*, and three *MM*, *MN*, *NN* genotypes are possible (the following data was collected from a single human population)

Genotye	No. of Individuals	Genotypic Frequencies
MM	1787	*MM*=1787/6129=0.289
MN	3039	*MN*=3039/6129=0.50
NN	1303	*NN*=1303/6129=0.21
Total	6129	

1.6. Deriving Gene (or Allelic) Frequencies

To determine the allelic frequencies we simply count the number of M or N alleles and divide by the total number of alleles.

$f(M)$ = [(2 x 1787) + 3039]/12,258 = 0.5395

$f(N)$ = [(2 x 1303) + 3039]/12,258 = 0.4605

By convention one of the alleles is given the designation p and the other q. Also $p + q = 1$.

p=0.5395 and q=0.4605

Furthermore, a population is considered by population geneticists to be polymorphic if two alleles are segregating and the frequency of the most frequent allele is less than 0.99.

1.7. Deriving allelic frequencies from genotypic frequencies

The following example will illustrate how to calculate allelic frequencies from genotypic frequencies. It will also demonstrate that two different populations from the same species do not have to have the same allelic frequencies.

	Percent			**Allelic Frequencies**	
Location	***MM***	***MN***	***NN***	***p***	***q***
Ibadan	83.5	15.6	0.90	0.913	0.087
Gombe	31.2	51.5	17.30	0.569	0.431

Let $p=f(M)$ and $q=f(N)$

Thus, $p=f(MM) + ½\ f(MN)$ and $q=f(NN) + ½\ f(MN)$.

So the results of the above data are:

Ibadan: p=0.835 + ½ (0.156)=0.913 and q=0.009 + ½ (0.156)=0.087

Gombe: p=0.312 + ½ (0.515)=0.569 and q=0.173 + ½ (0.515)=0.431.

Clearly the allelic frequencies vary between these populations.

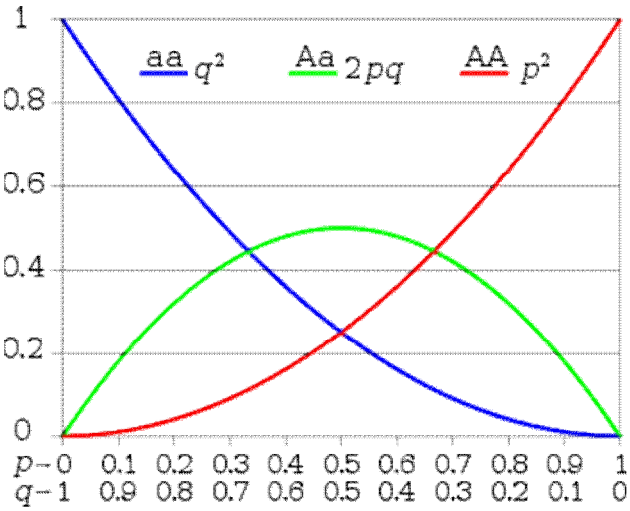


Fig. 1: Hardy–Weinberg genotype frequencies for two alleles: the horizontal axis shows the two allele frequencies *p* and *q* and the vertical axis shows the genotype frequencies. Each curve shows one of the three possible genotypes.

Population genetics was a vital ingredient in the emergence of the modern evolutionary synthesis. Its primary founders were Sewall Wright, J. B. S. Haldane and R. A. Fisher, who also laid the foundations for the related discipline of quantitative genetics.

1.8. Hardy–Weinberg principle

Natural selection will only cause evolution if there is enough genetic variation in a population. Before the discovery of Mendelian genetics, one common hypothesis was blending inheritance. But with blending inheritance, genetic variance would be rapidly lost, making evolution by natural selection unlikely. The *Hardy-Weinberg principle* provides the solution to how variation is maintained in a population with Mendelian inheritance. According to this principle, the frequencies of alleles (variations in a gene) will remain constant in the absence of selection, mutation, migration and genetic drift. The Hardy-Weinberg "equilibrium" refers to this stability of allele frequencies over time.

A second component of the Hardy-Weinberg principle concerns the effects of a single generation of random mating. In this case, the genotype frequencies can be predicted from the allele frequencies. For example, in the simplest case of a single locus with two

alleles: the dominant allele is denoted **A** and the recessive **a** and their frequencies are denoted by p and q; freq(**A**) = p; freq(**a**) = q; $p + q = 1$. If the genotype frequencies are in Hardy-Weinberg proportions resulting from random mating, then we will have freq(**AA**) = p^2 for the **AA** homozygotes in the population, freq(**aa**) = q^2 for the **aa** homozygotes, and freq(**Aa**) = $2pq$ for the heterozygotes.

1.8.1. The four factors affecting the Hardy Weinberg principle

1. **Natural selection**

Natural selection is the fact that some traits make it more likely for an organism to survive and reproduce. Population genetics describes natural selection by defining fitness as a propensity or probability of survival and reproduction in a particular environment. The fitness is normally given by the symbol **w**=1+**s** where **s** is the selection coefficient. Natural selection acts on phenotypes, or the observable characteristics of organisms, but the genetically heritable basis of any phenotype which gives a reproductive advantage will become more common in a population. In this way, natural selection converts differences in fitness into changes in allele frequency in a population over successive generations.

Before the advent of population genetics, many biologists doubted that small difference in fitness were sufficient to make a large difference to evolution. Population geneticists addressed this concern in part by comparing selection to genetic drift. Selection can overcome genetic drift when **s** is greater than 1 divided by the effective population size. When this criterion is met, the probability that a new advantageous mutant becomes fixed is approximately equal to **s**. The time until fixation of such an allele depends little on genetic drift, and is approximately proportional to log(sN)/s.

2. **Genetic drift**

Genetic drift is a change in allele frequencies caused by random sampling. That is, the alleles in the offspring are a random sample of those in the parents. Genetic drift may

cause gene variants to disappear completely, and thereby reduce genetic variability. In contrast to natural selection, which makes gene variants more common or less common depending on their reproductive success, the changes due to genetic drift are not driven by environmental or adaptive pressures, and may be beneficial, neutral, or detrimental to reproductive success.

The effect of genetic drift is larger for alleles present in a smaller number of copies, and smaller when an allele is present in many copies. Vigorous debates wage among scientists over the relative importance of genetic drift compared with natural selection. Ronald Fisher held the view that genetic drift plays at the most a minor role in evolution, and this remained the dominant view for several decades. In 1968 Motoo Kimura rekindled the debate with his neutral theory of molecular evolution which claims that most of the changes in the genetic material are caused by neutral mutations and genetic drift. The role of genetic drift by means of sampling error in evolution has been criticized by John H Gillespie and Will Provine, who argue that selection on linked sites is a more important stochastic force.

The population genetics of genetic drift are described using either branching processes or a diffusion equation describing changes in allele frequency. These approaches are usually applied to the Wright-Fisher and Moran models of population genetics. Assuming genetic drift is the only evolutionary force acting on an allele, after t generations in many replicated populations, starting with allele frequencies of p and q, the variance in allele frequency across those populations is

$$V_t \approx pq\left(1-\exp\left\{-\frac{t}{2N_e}\right\}\right).$$

3. **Mutation**

Mutation is the ultimate source of genetic variation in the form of new alleles. Mutation can result in several different types of change in DNA sequences; these can either have no effect, alter the product of a gene, or prevent the gene from functioning. Studies in the fly *Drosophila melanogaster* suggest that if a mutation changes a protein produced

by a gene, this will probably be harmful, with about 70 percent of these mutations having damaging effects, and the remainder being either neutral or weakly beneficial.

Mutations can involve large sections of DNA becoming duplicated, usually through genetic recombination. These duplications are a major source of raw material for evolving new genes, with tens to hundreds of genes duplicated in animal genomes every million years. Most genes belong to larger families of genes of shared ancestry. Novel genes are produced by several methods, commonly through the duplication and mutation of an ancestral gene, or by recombining parts of different genes to form new combinations with new functions. Here, domains act as modules, each with a particular and independent function, that can be mixed together to produce genes encoding new proteins with novel properties. For example, the human eye uses four genes to make structures that sense light: three for color vision and one for night vision; all four arose from a single ancestral gene. Another advantage of duplicating a gene (or even an entire genome) is that this increases redundancy; this allows one gene in the pair to acquire a new function while the other copy performs the original function. Other types of mutation occasionally create new genes from previously noncoding DNA.

In addition to being a major source of variation, mutation may also function as a mechanism of evolution when there are different probabilities at the molecular level for different mutations to occur, a process known as mutation bias. If two genotypes, for example one with the nucleotide G and another with the nucleotide A in the same position, have the same fitness, but mutation from G to A happens more often than mutation from A to G, then genotypes with A will tend to evolve. Different insertion vs. deletion mutation biases in different taxa can lead to the evolution of different genome sizes. Developmental or mutational biases have also been observed in morphological evolution. For example, according to the phenotype-first theory of evolution, mutations can eventually cause the genetic assimilation of traits that were previously induced by the environment.

Mutation bias effects are superimposed on other processes. If selection would favour either one out of two mutations, but there is no extra advantage to having both, then the mutation that occurs the most frequently is the one that is most likely to become fixed in a population. Mutations leading to the loss of function of a gene are much more common than mutations that produce a new, fully functional gene. Most loss of function mutations are selected against. But when selection is weak, mutation bias towards loss of function can affect evolution. For example, pigments are no longer useful when animals live in the darkness of caves, and tend to be lost. This kind of loss of function can occur because of mutation bias, and/or because the function had a cost, and once the benefit of the function disappeared, natural selection leads to the loss. Loss of sporulation ability in a bacterium during laboratory evolution appears to have been caused by mutation bias, rather than natural selection against the cost of maintaining sporulation ability. When there is no selection for loss of function, the speed at which loss evolves depends more on the mutation rate than it does on the effective population size, indicating that it is driven more by mutation bias than by genetic drift. Due to the damaging effects that mutations can have on cells, organisms have evolved mechanisms such as DNA repair to remove mutations. Therefore, the optimal mutation rate for a species is a trade-off between costs of a high mutation rate, such as deleterious mutations, and the metabolic costs of maintaining systems to reduce the mutation rate, such as DNA repair enzymes. Viruses that use RNA as their genetic material have rapid mutation rates, which can be an advantage since these viruses will evolve constantly and rapidly and thus evade the defensive responses of e.g. the human immune system.

4. Gene Flow & Transfer

Gene flow is the exchange of genes between populations, which are usually of the same species. Examples of gene flow within a species include the migration and then breeding of organisms, or the exchange of pollen. Gene transfer between species includes the formation of hybrid organisms and horizontal gene transfer.

Migration into or out of a population can change allele frequencies, as well as introducing genetic variation into a population. Immigration may add new genetic material to the established gene pool of a population. Conversely, emigration may remove genetic material. Other gene transfer mechanisms may include:

a) Reproductive isolation

As barriers to reproduction between two diverging populations are required for the populations to become new species, gene flow may slow this process by spreading genetic differences between the populations. Gene flow is hindered by mountain ranges, oceans and deserts or even man-made structures such as the Great Wall of China, which has hindered the flow of plant genes.

Depending on how far two species have diverged since their most recent common ancestor, it may still be possible for them to produce offspring, as with horses and donkeys mating to produce mules. Such hybrids are generally infertile, due to the two different sets of chromosomes being unable to pair up during meiosis. In this case, closely related species may regularly interbreed, but hybrids will be selected against and the species will remain distinct. However, viable hybrids are occasionally formed and these new species can either have properties intermediate between their parent species, or possess a totally new phenotype. The importance of hybridization in creating new species of animals is unclear, although cases have been seen in many types of animals, with the gray tree frog being a particularly well-studied example.

Hybridization is, however, an important means of speciation in plants, since polyploidy (having more than two copies of each chromosome) is tolerated in plants more readily than in animals. Polyploidy is important in hybrids as it allows reproduction, with the two different sets of chromosomes each being able to pair with an identical partner during meiosis. Polyploids also have more genetic diversity, which allows them to avoid inbreeding depression in small populations.

b) Genetic structure

Because of physical barriers to migration, along with limited tendency for individuals to move or spread (vagility), and tendency to remain or come back to natal place (philopatry), natural populations rarely all interbreed as convenient in theoretical random models (panmixy). There is usually a geographic range within which individuals are more closely related to one another than those randomly selected from the general population. This is described as the extent to which a population is genetically structured. Genetic structuring can be caused by migration due to historical climate change, species range expansion or current availability of habitat.

c) Horizontal Gene Transfer

Horizontal gene transfer is the transfer of genetic material from one organism to another organism that is not its offspring; this is most common among bacteria. In medicine, this contributes to the spread of antibiotic resistance, as when one bacteria acquires resistance genes it can rapidly transfer them to other species. Horizontal transfer of genes from bacteria to eukaryotes such as the yeast *Saccharomyces cerevisiae* and the adzuki bean beetle *Callosobruchus chinensis* may also have occurred. An example of larger-scale transfers are the eukaryotic bdelloid rotifers, which appear to have received a range of genes from bacteria, fungi, and plants. Viruses can also carry DNA between organisms, allowing transfer of genes even across biological domains. Large-scale gene transfer has also occurred between the ancestors of eukaryotic cells and prokaryotes, during the acquisition of chloroplasts and mitochondria. The basic models of population genetics consider only one gene locus at a time. In practice, epistatic and linkage relationships between loci may also be important.

Epistasis

Because of epistasis, the phenotypic effect of an allele at one locus may depend on which alleles are present at many other loci. Selection does not act on a single locus, but on a phenotype that arises through development from a complete genotype.

According to Lewontin (1974), the theoretical task for population genetics is a process in two spaces: a "genotypic space" and a "phenotypic space". The challenge of a *complete* theory of population genetics is to provide a set of laws that predictably map a population of genotypes (G_1) to a phenotype space (P_1), where selection takes place, and another set of laws that map the resulting population (P_2) back to genotype space (G_2) where Mendelian genetics can predict the next generation of genotypes, thus completing the cycle. Even leaving aside for the moment the non-Mendelian aspects of molecular genetics, this is clearly a gargantuan task. Visualizing this transformation schematically:

$$G_1 \xrightarrow{T_1} P_1 \xrightarrow{T_2} P_2 \xrightarrow{T_3} G_2 \xrightarrow{T_4} G_1' \rightarrow \cdots$$

(adapted from Lewontin 1974, p. 12). XD

T_1 represents the genetic and epigenetic laws, the aspects of functional biology, or development, that transform a genotype into phenotype. We will refer to this as the "genotype-phenotype map". T_2 is the transformation due to natural selection, T_3 are epigenetic relations that predict genotypes based on the selected phenotypes and finally T_4 the rules of Mendelian genetics.

In practice, there are two bodies of evolutionary theory that exist in parallel, traditional population genetics operating in the genotype space and the biometric theory used in plant and animal breeding, operating in phenotype space. The missing part is the mapping between the genotype and phenotype space. This leads to a "sleight of hand" (as Lewontin terms it) whereby variables in the equations of one domain, are considered parameters or *constants*, where, in a full-treatment they would be transformed themselves by the evolutionary process and are in reality *functions* of the state variables in the other domain. The "sleight of hand" is assuming that we know this mapping. Proceeding as if we do understand it is enough to analyze many cases of interest. For example, if the phenotype is almost one-to-one with genotype (sickle-cell

disease) or the time-scale is sufficiently short, the "constants" can be treated as such; however, there are many situations where it is inaccurate.

Linkage

If all genes are in linkage equilibrium, the effect of an allele at one locus can be averaged across the gene pool at other loci. In reality, one allele is frequently found in linkage disequilibrium with genes at other loci, especially with genes located nearby on the same chromosome. Recombination breaks up this linkage disequilibrium too slowly to avoid genetic hitchhiking, where an allele at one locus rises to high frequency because it is linked to an allele under selection at a nearby locus. This is a problem for population genetic models that treat one gene locus at a time. It can, however, be exploited as a method for detecting the action of natural selection via selective sweeps.

In the extreme case of primarily asexual populations, linkage is complete, and different population genetic equations can be derived and solved, which behave quite differently to the sexual case. Most microbes, such as bacteria, are asexual. The population genetics of microorganisms lays the foundations for tracking the origin and evolution of antibiotic resistance and deadly infectious pathogens. Population genetics of microorganisms is also an essential factor for devising strategies for the conservation and better utilization of beneficial microbes.

1.9. Modern evolutionary synthesis

The mathematics of population genetics were originally developed as the beginning of the modern evolutionary synthesis. According to Beatty (1986), population genetics defines the core of the modern synthesis. In the first few decades of the 20th century, most field naturalists continued to believe that Lamarckian and orthogenic mechanisms of evolution provided the best explanation for the complexity they observed in the living world. However, as the field of genetics continued to develop, those views became less tenable. During the modern evolutionary synthesis, these ideas were purged, and only evolutionary causes that could be expressed in the mathematical framework of

population genetics were retained. Consensus was reached as to which evolutionary factors might influence evolution, but not as to the relative importance of the various factors.

Theodosius Dobzhansky, a postdoctoral worker in T. H. Morgan's lab, had been influenced by the work on genetic diversity by Russian geneticists such as Sergei Chetverikov. He helped to bridge the divide between the foundations of microevolution developed by the population geneticists and the patterns of macroevolution observed by field biologists, with his 1937 book *Genetics and the Origin of Species*. Dobzhansky examined the genetic diversity of wild populations and showed that, contrary to the assumptions of the population geneticists, these populations had large amounts of genetic diversity, with marked differences between sub-populations. The book also took the highly mathematical work of the population geneticists and put it into a more accessible form. Many more biologists were influenced by population genetics via Dobzhansky than were able to read the highly mathematical works in the original.

1.10. Selection vs. genetic drift

Fisher and Wright had some fundamental disagreements and a controversy about the relative roles of selection and drift continued for much of the century between the Americans and the British.

In Great Britain E.B. Ford, the pioneer of ecological genetics, continued throughout the 1930s and 1940s to demonstrate the power of selection due to ecological factors including the ability to maintain genetic diversity through genetic polymorphisms such as human blood types. Ford's work, in collaboration with Fisher, contributed to a shift in emphasis during the course of the modern synthesis towards natural selection over genetic drift.

Recent studies of eukaryotic transposable elements, and of their impact on speciation, point again to a major role of nonadaptive processes such as mutation and genetic drift.

Mutation and genetic drift are also viewed as major factors in the evolution of genome complexity.

1.11. The Hardy-Weinberg Law

The HW law is the unifying concept of population genetics. It was named after the two scientists who simultaneously discovered the law. The law predicts how gene frequencies will be transmitted from generation to generation given a specific set of assumptions. Specifically,

If an infinitely large, random mating population is free from outside evolutionary forces (i.e. mutation, migration and natural selection), **then** the gene frequencies will not change over time, and the frequencies in the next generation will be:

p^2 for the *AA* genotype
***2pq* for the *Aa* genotype, and**
q^2 for the *aa* genotype.

Let us examine the assumptions and conclusions in more detail starting first with the assumptions.

1. Infinitely large population

No such population actually exists in nature. The effect that is of concern is genetic drift which is a problem in small populations.

Genetic drift - is a change in gene frequency that is the result of chance deviation from expected genotypic frequencies.

2. Random mating

These are matings in a population that occur in proportion to their allelic frequencies. For example, if the allelic frequencies in a population are:

$f(M) = 0.91$
$f(N) = 0.09$

then the probability of *MM* individuals occurring is 0.91 x 0.91 =0.828.

If a significant deviation occurred, then random mating did not happen in this population.

Point to remember about random mating: Within a population, random mating can be occurring at some loci but not at others.

Examples of random mating loci: Blood type, RFLP patterns.Examples of non-random mating loci: intelligence, physical stature

3. No evolutionary forces affecting the population

The principal forces are, Mutation, Migration and Selection. Some loci in a population may be affected by these forces, and others may not; those loci not affected by the forces can by analyzed as a Hardy-Weinberg population.

Mathematical Derivation of the Hardy-Weinberg Law

If *p* equals the frequency of allele *A* in a population and *q* is the frequency of allele *a* in the same population, union of gametes would occur with the following genotypic frequencies:

		Female Gametes*	
		p (*A*)	*q* (*a*)
Male Gametes	*p* (*A*)	p^2 (*AA*)	*pq* (*Aa*)
	q (*a*)	*pq* (*Aa*)	q^2 (*aa*)

The gamete and offspring genotypes are in parentheses.

From the table, it is clear that the prediction regarding genotypic frequencies after one generation of random mating is correct. That is: Prediction regarding stability of gene frequencies. The following is a mathematical proof of the second prediction. To determine the allelic frequency, they can be derived from the genotypic frequencies as shown above.

$p = f(AA) + \frac{1}{2}f(Aa)$	(substitute from the table on previous page)
$p = p^2 + \frac{1}{2}(2pq)$	(factor out p and divide)
$p = p(p + q)$	($p + q = 1$; therefore $q = 1 - p$; make this substitution)
$p = p\,[p + (1 - p)]$	(subtract and multiply)
$p = p$	

Self Assessment Assignment

1. Define population genetics and mention the goals of population genetics
2. Differentiate between allelic and genotypic frequencies
3. Explain the concept of the Hardy Weinberg principle

Tutor Marked Assessment

1. Mention and discuss the principles affecting the Hardy – Weinberg principle.

2.0. CYTOGENETICS

CYTOGENETICS

Cytogenetics is branch of biology devoted to the study of chromosomes and their implications in genetics.

CHROMOSOMAL ABERRATIONS

Occasionally, spontaneous (without any known causal factor) variations in the structure and number of chromosomes have been observed in nature. These variations are called chromosomal aberrations and can be due to either (a) structural changes or (b) numerical changes

Origin of structural aberration

Chromosomes are structures with definite organization. However, through various means they may be broken and their normal structure disrupted. X-rays, atomic radiations and various chemicals are among the agents that can cause breaks in chromosomes. Breaks also sometimes occur under natural conditions, where in most instances the reason for breakage is not known. An initially single deviation from the normal can give rise to a whole series of unusual cytological events.

Breakage-fusion-bridge Cycle : In the gametophyte and endosperm of corn, ends of chromosomes that have recently been broken behave as though they were “sticky,” as is shown by their tendency to adhere to one another. Extensive studies of broken chromosomes in corn have been made by Barbara Mc Clintock. She found that following reduplication of a broken chromosome the two sister chromatids may adhere at the point of previous breakage. The fused sister chromatids would be unable to separate readily. In effect, they constitute a single chromatid with two centromeres, a dicentric chromatid. As the centromeres move to opposite poles at anaphase, the dicentric chromatid stretches out, forming a chromatin bridge from one pole toward the other. This bridge eventually breaks, but the break does not always occur at the point of previous fusion. Therefore, chromosomes may be formed that show duplications or deficiencies if compared with an original type.

Thus, if the original chromosome is **C B A** the type ***C B A A*** is a duplication type, since region *A* is represented twice. The type ***C B*** is a deficiency, since region *A* is absent. When a chromosome bridge breaks, perhaps as a the result of tension caused by the movement of two centromeres of the dicentric chromatid two new broken ends are formed. Each of these has the same qualities of adhesiveness that gave rise to the original fusion. This situation permits repetition of more similar events as described above, in cyclic series. Spontaneous production of chromosome aberrations through breakage -fusion -bridge cycles may occur in this manner for some time. But when a broken chromosome is introduced into the sporophytic generation such cycles cease, as the broken ends heal in the zygote.

Structural chromosomal aberrations: Structural chromosomal changes include those chromosomal aberrations which alter the chromosome structure i.e. the number of genes, the sequence or kind of genes present in the chromosome(s) and do not involve a change in chromosome number.

Types of structural chromosomal aberrations: These aberrations may be confined to a single chromosome or more than one chromosome. Hence they are of two types.

1. Intra-chromosomal aberrations
2. Inter-chromosomal aberrations

I. Intra -chromosomal aberrations: When aberrations remain confined to a single chromosome of a homologous pair, they are called intra -chromosomal aberrations.

II. Inter-chromosomal aberrations : When breaks occur in non-homologous chromosomes and the resulting fragments are inter-changed by both the nonhomologous chromosomes, they are known as inter-chromosomal aberrations.

I. Intra-chromosomal aberrations: They may be of the following types:

1. Deletions or deficiencies: Deletion is due to the loss of a part of a chromosome. In a deletion, a chromosome lacks either a terminal or an interstitial segment which may include only a single gene or a part of a gene. Hence it is of two types.

i) Terminal deletion: If a break occurs near the end of chromosome and a small piece of terminal chromosome is lost, it is called terminal deletion.

ii) Interstitial or intercalary deletion: Sometimes two breaks may occur at any two points and the broken ends of the original chromosome get fused or reunited and as a result, an interstitial deletion is formed. If the chromosome has a centromere, it will persist. Otherwise it will be lost during cell division. Both types of deletions can be observed during pachytene stage of meiosis or in the polytene chromosomes. Deficiencies can be artificially induced using radiations. However, in majority of the cases, deficiencies are intercalary because, when a terminal part of the chromosome is lost, it can not be repaired unless it unites with another broken end.

Genetic significance / e ffects of deletions:

1. Organisms with homozygous deletion do not survive to an adult stage because a complete set of genes is lacking (lethal effect).

2. Small deficiencies, if present in heterozygous condition (deficiency heterozygote) can be tolerated by the organism. In such individuals during pachytene stage of meiosis the unpaired segment of the normal chromosome of an intercalary deletion heterozygote produces a characteristic loop in a bivalent. In case of terminal deletion heterozygotes, the segment towards one end of the normal chromosome remains unpaired. Loops can be observed in salivary gland chromosomes of *Drosophila* or giant chromosomes, which are found in a permanent state of pairing. Therefore even small deficiencies could be detected in these chromosomes.

3. Deficiencies have an effect on the inheritance also. In presence of deficiency a recessive allele will behave like a dominant allele. When an organism heterozygous for a pair of alleles Aa loses a portion of the chromosome bearing the dominant allele ‘A’, the recessive allele being in the hemizygous condition will be expressed phenotypically. This phenomenon is known as pseudo-dominance.

4. This principle of pseudo -dominance has been utilized for location of genes on specific chromosomes in *Drosophila*, maize and other organisms. Thus, deletions are important cytological

tools for mapping genes. 5. Deletions play an important role in species formation and creating variability through chromosomal mutations. In *Drosophila* , deletions were recorded on X-chromosome in the regions containing genes-w (for white eye), fa (for facet eye) and v (for vermilion eye colour).

2. Duplications or Repeats: Duplication occurs when a segment of chromosome is represented two or more times in a chromosome of a homologous pair. The extra segment may be a free fragment with a centromere or a chromosomal segment of the normal complement. As a result, in one chromosome of the homologous pair, there will be deletion, while in other there will be a duplication. Duplication was first reported in *Drosophila* by C.B. Bridges in 1919. Duplications are of four types.

1. Tandem: The extra chromosome segment may be located immediately after the normal segment in precisely the same orientation (i.e. having the same gene sequence)

2. Reverse tandem: The gene sequence in the extra segment of a tandem duplication is in the reverse order i.e. is inverted. (eg . cb in place of bc)

3. Displaced: The extra segment may be located in the same chromosome but away from the normal segment.

4. Reverse displaced: The gene sequence in the extra segment of a displaced duplication is in the reverse order i.e. is inverted (eg. ed in place of de). If duplication is present in only one of the two homologous chromosomes, at pachytene stage of meiosis, cytological observations characteristic of deficiency will be obtained in duplication also.

Genetic significance / effects of duplications:

1. Duplications are not as harmful as deletions.
2. Some duplications are useful in the evolution of new genetic material.
3. Large duplications can reduce fertility.
4. The phenotype may be altered.

One of the classical examples of duplication in *Drosophila* is that of bar eye. Bar eye is a character where the eyes are narrower as compared to normal eye shape. This phenotypic character is due to the duplication for a part of chromosome. By the study of giant salivary gland chromosomes of *Drosophila melanogaster*, it could be demonstrated that 'Bar' character was due to duplication in the region 16A of x chromosome. Barred individuals (16A 16A) gave rise to ultra bar (16A 16A 16A) and normal wild type (16A) due to unequal crossing over.

3. Inversions: When a segment o f a chromosome is oriented in reverse direction, such a segment is said to be inverted and the phenomenon is termed as inversion. Gene sequence in an inverted segment is exactly the opposite of that in its normal homologous pair. Inversions would involve two breaks followed by reunion of interstitial segment in a reverse order i. e . the segment rotates by

180_0. (Let us imagine that a chromosome 1-2-3-4-5-6-7-8 gives rise to another chromosome having the order 1-2-6-5-4-3-7-8. the segment 3-4-5-6 has rotated here at 180_0 giving an inverted order of genes 6-5-4-3). Inversions can be of two types depending upon whether centromere is involved or not in inversions.

(a) Paracentric inversions and (b) Pericentric inversions.

(a) Paracentric inversions: Paracentric inversions are those inversions, where the inverted segment does not include centromere.

(b) Pericentric inversions: The inverted segment includes the centromere in pericentric inversions. (pericentric means surrounding the centromere or on the periphery of centromere).

Cytology of inversions: When both the members of a homologous pair have similar type of inversion, it is called inversion homozygote. Meiosis is normal in inversion homozygotes. When only one chromosome of a homologous pair has inversion it is called inversion heterozygote. Due to an inverted segment in one of the two homologous chromosomes, the normal kind of pairing is not possible in an inversion heterozygote. In order to enable pairing of homologous segments, a loop is formed by each of the two chromosomes. This kind of configuration will be observed in paracentric as well as pericentric inversions. However, the products of crossing over and the subsequent stages of meiosis will differ in these two kinds of inversions.

1. **Paracentric inversion:** A single crossing over or an odd number of crossovers in an inverted region will result into the formation of a dicentric chromosome (having two centromeres) and an acentric chromatid (without centromere) when two chromatids are involved in the crossing over. These dicentric chromatid and acentric chromatid will be observed at anaphase I in the form of a bridge and a fragment.

2. **Pericentric inversion:** In a pericentric inversion, although at pachytene, the configuration observed is similar to that described above for paracentric inversion, the products of crossing over and the configurations of subsequent stages of meiosis differ. In this case, two of the four chromatids resulting after meiosis will have deficiencies and duplications. Unlike paracentric inversion, no dicentric bridge or acentric fragment will be observed at anaphase I. However, in pericentric inversion, if the two breaks are not situated equidistant from the centromere, this will result in the change in shape of the chromosome. For instance, a metacentric chromosome may become sub-metacentric and vice -versa.

Genetic consequences of inversions

1. Simple inversions do not have primary effects other than change in chromosome shape.
2. A peculiar kind of position effect occurs due to suppression of the transcription of gene.
3. Normal linear pairing of homologous chromosomes is not possible.

4. Heterozygosity will be maintained from generation to generation.

5. Among the four chromatids resulting after crossing over, two chromatids would have deficiencies and duplications. The gametes having these chromosomes will not function normally and lead to high sterility. Therefore there should be considerable gametic or zygotic lethality. In plants there will be sufficient pollen sterility. However, since the products of single crossing over will not function and the only crossover products recovered will be double cross overs, the observed frequency of recombination between any two genes of interest will be considerably reduced. Due to this reason, inversions (especially paracentric inversions) are often called crossover suppressors. This reduction in crossing over is not the actual reduction in cytological crossing over, but is the result of lack of recovery of the products of single cross overs. This property of inversion has been utilized in the production of ClB stock used by Muller for the detection of sex-linked lethal mutations in *Drosophila melanogaster*.

4. Shifts: Shifts are altered forms of inversions. In this type of aberrations, the genes are in the right order but a segment is shifted either to the right or to the left.

5. Isochromosome: Isochromosome is the one in which both the arms are identical in both gene content and morphology. It arises when the centromere divides in wrong plane yielding two daughter chromosomes each of which carries the information of one arm only but present twice.

II. Interchromosomal aberrations : These are of the following types.

1. Translocation : Integration of a chromosome segment into a non –homologous chromosome is called translocation. It involves shifting of one part of chromosome to another non-homologous chromosome. The phenomenon of translocation was discovered by C.B. Bridges in 1923 in *Drosophila* and by Hugo de Vries in *Oenothera lamarckiana*.

Translocation is of two types

1. Simple translocation : In simple translocation, the terminal segment of chromosomes is integrated at one end of a non -homologous chromosome. However, they are rare.

2. Reciprocal translocation : If two non -homologous chromosome exchange segments which need not be of same size, it results in reciprocal translocation. Production of reciprocal translocation requires a single break in each of the two non –homologous chromosomes followed by reunion of the chromosome segments thus produced. An individual having reciprocal translocation may be either a translocation homozygote or a translocation heterozygote. When both the chromosomes from each pair are involved, it produces translocation homozygote and when only one chromosome from each pair of two homologues is involved, it gives rise to translocation heterozygote. In a translocation homozygote, the two homologues of each of the two translocated chromosomes are identical in their gene content. As a result, they form normal bivalent and there is no detectable cytogenetic aberration (peculiarity). In a translocation heterozygote, one member from each of two

homologous pairs is involved in reciprocal translocation, while the remaining chromosomes of the two concerned pairs are normal. Due to the pairing between homologous segments of chromosomes, a cross-shaped (+) figure involving four chromosomes will be observed at pachytene. These four chromosomes at metaphase I will form a quadrivalent, which may exhibit any one of the following three orientations. 1. Alternate 2. Adjacent I 3. Adjacent II

1. **Alternate :** In this case, the centromeres lying alternate to each other in the cross shaped figure move to the same pole. In otherwords, the adjacent chromosomes will orient towards opposite poles. As a result, the two normal ch romosomes move to one pole, while the two translocated chromosomes move to the opposite pole. Such a segregation can take place only when the cross shaped figure of four chromosomes is twisted to form a figure of 'oo'.

2. **Adjacent – I:** In adjacent I orientation, adjacent chromosomes having nonhomologous centromeres will orient towards the same pole. In other words, the chromosomes having homologous centromeres will orient towards the opposite poles. Thus a ring of four chromosomes will be observed.

3. **Adjacent – II:** In adjacent II orientation, the adjacent chromosomes having homologous centromeres will orient towards the same pole. In this case also a ring of four chromosomes will be observed. In both Adjacent-I and adjacent-II disjunctions, one normal and one translocated chromosome move to the opposite poles. Adjacent-I and adjacent-II disjuctions will form

gametes which would carry duplications or deficiencies and as a result would be non - functional or sterile. Therefore, in a plant having translocation in heterozygous condition, there will be considerable pollen sterility.

Genetic significance of translocation heterozygotes:

1. They produce semi sterile plants with low seed set.
2. Some genes which earlier assorted independently tend to exhibit linkage relationship.
3. The phenotypic expression of a gene may be modified when it is translocated to a new position in the genome. 4. The presence of translocation heterozygosity can be detected by the occurrence of semi-sterility and low seed set. This can then be confirmed at me iosis by quadrivalent formation. Functional gametes will be formed only from alternate disjunction, which will give rise to three kinds of progeny *viz*., normal, translocation heterozygotes and translocation homozygotes in 1:2:1 ratio.

Role of structural chromosomal aberrations in plant breeding

1. They are useful in the identification of chromosomes
2. Utilization of vigour as in case of duplication.
3. Useful in genome analysis.
4. Useful for the transfer of desirable characters through translocation.

5. They have evolutionary significance.

Lecture No.: 28, 29 & 30

NUMERICAL CHROMOSOMAL ABERRATIONS

Each species of micro -organisms, plants and animals is characterized by particular chromosome complement or set of genome, represented once in gametic (haploid) cell i.e. n and twice in somatic (diploid) cells i.e. 2n. The term genome refers to a complete set of chromosomes of a diploid species. All the members of a genome are distinct from each other in gene content and often in morphology. Members of a genome do not pair. Possession of such sets of chromosomes or genomes, gives a specific chromosome number to each species. But sometimes, some irregularities may occur during mitosis, meiosis or fertilization and may produce cells with variant chromosome number. A deviation from the diploid state represents a numerical chromosomal aberration which is often referred to as heteroploidy. Individuals possessing variant chromosome number are known as heteroploids. Variation in chromosome number (ploidy) may occur through the addition or loss of complete chromosome set or genome (euploidy) or of one or few chromosomes (aneuploidy).

Thus numerical changes in chromosomes (heteroploidy) can be mainly of two types: 1. Euploidy and 2. Aneuploidy.

Euploidy: (Greek word; Eu = true or even; ploidy = unit). The term euploidy designates a change in chromosome number which involves entire set of chromosomes. Euploids have one or more complete genomes, which may be identical with or distinct from each other. The somatic chromosome number of a euploid individual is exact multiple of basic chromosome number of that species. The basic chromosome number refers to the haploid or gametic chromosome number of a diploid species and in case of polyploidy species, the haploid chromosome number of parental diploid species; represented by x. Euploidy includes monoploids, diploids and polyploids.

Monoploids: Monoploids contain a single chromosome set and are characteristically sterile. In other words monoploids have the basic chromosome number (x) of a spec ies. Monoploids (x) differ from haploids (n) which carry half or gametic chromosome number. In a true diploid species, both monoploid and haploid chromosome number are same (i. e. x=n).

Haploid: Haploid is a general term used to designate the individuals or tissues with a gametic chromosome number i.e. n.

Differences between monoploids and haploids:

	Monoploids	**Haploids**
1	Represent gametic chromosome number of a diploid species	Represent gametic chromosome number of any species
2	Denoted by 'x' 2.	Denoted by 'n'
3	Monoploids are always haploids	Haploids cannot always be monoploids
4	Contain single set of genome	May contain one or more copies of genome.
5	Eg : Maize 2n = 20 x = 10 n = 10	Wheat 2n = 6x = 42 x = 7 n = 21

Haploids can be of two types:

1. Monohaploids: Individuals that arise from a normal diploid species. Eg. : Maize 2n = 20 and n = 10

2. Polyhaploids: Individuals that arise from any polyploid species. Eg : Wheat 2n = 6x = 42 and n=3x=21. Haploids can arise spontaneously or can be induced. Spontaneous haploids have been obtained in plants like tomato, cotton, barley, coffee, pearl millet and wheat. The first induced haploid was produced by Jorgensen in 1928 by crossing *Solanum nigrum* x *Solanum luteum* Guha and Maheswari in 1964 obtained haploid plants of *Datura innoxia* directly from the pollen by culturing the anthers. The main reason for plant breeders' interest to obtain the haploids has been to develop a new and rapid method of breeding homozygous diploids or polyploids through diploidization using colchicine. In a haploid, every gene is present only once. Doubling of chromosomes should theoretically result in complete homozygosity. Repeated inbreeding for homozygosity in plants requires many (8-9) generations. But doubling of the chromosome number of haploids results in immediate homozygosity. Diploids obtained through the chromosome doubling of haploids are known as dihaploids.

Characteristic features of haploid plants:

Haploids are smaller, less vigorous than their diploid phenotypes. Haploids are sterile, as the chromosomes have no regular pairing partner homologous chromosomes during

meiosis and they are found as univalents at metaphase I of meiosis. The meiotic products are deficient in one or more chromosomes. For instance, a haploid in maize, (2n=20) will have 10 chromosomes and the number of chromosomes in gametes can range from 0-10. Consequently considerable sterility will be found in a haploid maize. Haploids can be produced through anther culture, parthenocarpy, delayed pollination etc.

Diploidy: Diploidy is characterized by presence of two genomes in each somatic cell of the diploid organism. Most animals and plants are diploids. The diploidy is related with fertility, balanced growth, vigour, adaptability and survival of diploid organisms.

Polyploidy: The organisms with more than two genomes are called polyploids. Among plants, polyploidy occurs in multiple seri+es of 3, 4, 5, 6, 7, 8 etc. of the basic chromosome number and thus resulting in triploids, tetraploids, pentaploids, hexaploids, heptaploids, octaploids etc., respectively. Generally ploidy levels higher than tetraploid are not commonly encountered in the natural population. However there are some exceptions. Eg: hexaploid (6x) wheat, octaploids (8x) straw berries, many commercial fruits and ornamental plants, liver cells of man

etc.

Origin of polyploidy: Different degrees of ploidy are originated by different means. However, two basic irregular processes have been discovered by which polyploids may evolve from diploid plants and become established in nature.

1. Somatic doubling: Cells sometimes undergo irregularities in mitosis and give rise to meristematic cells that perpetuate these irregularities in new generations of plants.

2. Reproductive process: Reproductive cells may have an irregular reduction division in which the sets of chromosomes fail to separate completely to the poles at anaphase. Both the sets thus become incorporated in the same nucleus resulting in the doubling of chromosome number in the gamete. Thus a triploid originates by irregularities during meiosis (i.e. by union of diploid gamete with haploid gamete) Likewise a tetraploid may originate by the somatic doubling of the chromosome number or by union of unreduced diploid gametes.

Induction of polyploidy: Polyploidy can be induced by two methods. 1. By physical agents and 2. By chemical agents

1. By physical agents:

a) Temperature shocks: Extreme changes in temperature results in a higher frequency of polyploid cells.

b) Centrifugation: Centrifugation of seedlings or plants causes polyploidy in their cells.

c) X-rays: X- rays can also induce polyploidy

2. By chemical agents: Some chemicals like colchicine, chlorohydrate, mercuric chloride have been found to induce polyploidy in plants. Colchicine treatment is the most effective and most widely used treatment for chromosome doubling. The chromosome doubling effect of colchicine was first described by Blakeslee and Nebel independently. Colchicine interferes or disturbs the formation of spindle fibres during cell division and thus inhibits the movement of sister chromatids to the opposite poles. Colchicine is a poisonous chemical isolated from the seeds and bulbs of autumn crocus (*Colchicum autumnale*). Pure colchicine is $C_{22}H_{25}O_6N$.

Kinds of polyploids: Polyploids are distinguished on the basis of source of chromosomes into three main kinds. 1. Auto polyploids, 2. Allopolyploids and 3. Segmental allopolyploids

1. Autopolyploids: In a plant, when same set of chromosomes of a genome are increased in number, autopolyploids are obtained. The prefix “auto” indicates that the ploidy involves homologous chromosome sets. For example, if a diploid species has two similar sets of chromosomes / genomes designated as AA, an autotriploid will have three similar (AAA) genomes and autotetraploid will have four similar (AAAA) genomes. Genetical and morphological characters expressed by autopolyploid depend on the genetic constitution of parent plant. In general, expression is exaggerated either in positive / negative direction. For example: vegetative growth may be more vigourous. Leaves may be more broader or dark green in colour. The floral parts, fruits and seeds may be bigger. However, autopolyploids occur rarely in natural populations. Eg: Auto triploids – *Cynodon dactylon* (doob grass).

a) Auto triploids: Auto triploids have three complete sets of genomes of the same species in somatic cell. Triploids can arise in several ways. Generally, in nature they originate by the fusion of a haploid gamete with a diploid gamete (unreduced gamete). Diploid gametes occur sporadically as unreduced germ cell in a diploid organism. They are also produced by meiosis in tetraploid organism or in segments of otherwise diploid organisms where doubling of the somatic chromosome number has taken place. Triploids can be produced

artificially by crossing between autotetraploid and diploid species. Triploids are generally highly sterile. In an autotriploid, there are three sets of homologous chromosomes. If these sets are normally paired, trivalents (as observed in primary trisomics) will be observed. The trivalents can not disjoin normally and will either disjoin 2:1 chromosomes to two poles or will disjoin 1:1 leaving one chromosome as a laggard. The number of chromosomes in the gametes of a triploid organism therefore, will vary from n to 2n. Most of these gametes are unbalanced leading to high degree of sterility. Examples of triploidy in animals are rather rare. Triploids are useful only in those plant species which propagate asexually like banana, sugarcane, apple, sugarbeet, watermelon, tomato, *Cynodon dactylon* (doob grass). Seedless watermelons are grown commercially in Japan. They are produced by crossing tetraploid (4x, used as female) and diploid (2x, used as male) lines, since the reciprocal cross (diploid female X tetraploid male) is not successful. The triploid plants do not produce true seeds. Almost all the seeds are small, white rudimentary structures like cucumber seeds. For good fruit setting, diploid plants are planted in the ratio of 1 diploid : 5 triploid plants.

b) Auto tetraploids: In autotetraploids, four copies of the genome of same species (AAAA or BBBB) are present. They may arise spontaneously or can be induced artificially by doubling the chromosomes of a diploid species with colchicine treatment. In autotetraploids, since there are four sets of chromosomes, quadrivalents are formed, which disjoin in a normal 2:2 manner giving diploid gametes. Rarely, a quadrivalent may disjoin in 3:1 or may leave one or more chromosomes as laggards at anaphase I. Therefore autotetraploids also have a certain degree of sterility, although it will not be as high as in autotriploids. Autotetraploids are usually larger and more vigorous than the diploid species. Eg: Rye, grapes, groundnut, potato, coffee, *Oenothera lamarckiana*. In an autotetraploid, four chromosomes are homologous to each other, hence each gene has four copies. A simplex individual has one dominant and three recessive alleles (Aaaa), a duplex has two dominant and two recessive alleles (AAaa), a triplex has three dominant and one recessive alleles (AAAa), a quadruplex has all dominant alleles (AAAA), while a nulliplex has no dominant alleles (aaaa).

2. Allopolyploids: A polyploid containing genetically different chromosome sets from two or more species is known as allopolyploid. The prefix “allo” indicates the involvement of non-homologous sets of chromosomes.

Origin of allopolyploids: Natural allopolyploids most likely originate through chromosome doubling of F_1 hybrid produced by chance through natural hybridization between two distinct species of the same genus or from different genera. Experimental production of allopolyploids is achieved through chromosome doubling of F_1 hybrid with the help of colchicine. Such allopolyploids are often called synthetic allopolyploids. The synthesis of allopolyploids involves two steps. 1. Production of F_1 hybrids by crossing two distinct species and 2. Chromosome doubling of such F_1 hybrids. The man made cereal Triticale is an example of synthetic allopolyploid.

Amphidiploid: It is an allopolyploid (allotetraploid) which arises by combining genomes of two different species. The amphidiploids are fertile due to the presence of homologous chromosomes and behave as a diploid during meiosis. The term amphidiploids was proposed by Nawashin.

Natural allopolyploids: Inter-specific crossing followed by chromosome doubling in nature have resulted in origin of some natural allopolyploid crops like cotton, tobacco, mustard, wheat, etc. The origin of some natural allopolyploid crops is briefly presented below:

(i) Cotton: The new world cotton (*Gossypium hirsutum*) is an interesting example of allopolyploidy. J.O. Beasley crossed old world cotton (*Gossypium herbaceum*) with American cotton (*Gossypium raimondii*) and doubled the chromosome number in F_1 hybrids. The allopolyploid thus produced resembled the cultivated new world cotton (*Gossypium hirsutum*) and when crossed with it gave fertile F_1 hybrids. These results thus suggested that tetraploid cotton (*Gossypium hirsutum*) originated from two diploid species namely *Gossypium herbaceum* (2n = 26) and *Gossypium raimondii* (2n = 26).

(ii)Tobacco: There are two cultivated species of tobacco. i. e. *Nicotiana tabacum* and *Nicotiana rustica.* a) *Nicotiana tabacum* is an allotetraploid and available evidence suggests that it is derived from a cross between *Nicotiana sylvestris* x *Nicotiana tomentosa*

b) *Nicotiana rustica* is believed to be an amphidiploid obtained from a cross between *Nicotiana paniculata* and *Nicotiana undulate Nicotiana paniculata* X *Nicotiana undulate* 2n = 24(diploid) 2n = 24(diploid)

(ii) *Brassica*: Several of *Brassica* species like *Brassica juncea*, *Brassica napus* and *Brassica carinata* are allotetraploids (amphidiploids). It is believed that *Brassica juncea* is an

amphidiploid derived from a cross between *Brassica nigra* and *Brassicacampestris*; *Brassica napus* is an amphidiploid derived from a cross between *Brassica oleracea* and *Brassica campestris* and *Brassica carinata* is an amphidiploid derived from a cross between *Brassica nigra* and *Brassica oleracea*.

(iv) Wheat: The common or bread wheat, *Triticum aestivum* (formerly *Triticum spelta*) is an allohexaploid. It was artificially synthesized in 1946 by Mc Fadden and Sears. It has two copies each of the genomes A, B and D and its somatic complement is represented as AA BB DD. The sources of A and D genomes are more or less unanimously accepted as *Triticum monococcum* (AA) and *Triticum tauschii* (DD) (formerly *Aegilops squarrosa* –goat grass), respectively. There is considerable doubt about the source of B genome. According to one hypothesis, *Aegilops speltoides* may be the source of this genome. But recent evidences do not support this idea. Most likely, the donor of B genome is now extinct and its identity is still not clear. Most likely, the amphidiploid AABB was produced initially. This gave rise to a tetraploid wheat, *Triticum turgidum* (formerly, *Triticum dicoccum* – emmer wheat). This amphidiploid (AABB) was subsequently outcrossed with *Triticum tauschii* (formerly *Aegilops squarrosa* – goat grass) to ultimately yield the hexaploid wheat, *Triticum aestivum* (AABBDD) *Triticum monococcum* X Unknown diploid species 2n = 14(diploid)

Artificial allopolyploids: Artificial allopolyploids have been synthesized in some crops either to study the origin of naturally available allopolyploids or to explore the possibilities of creating new species. Some examples of artificial allopolyploids are given below:

i) Triticale: Triticale, a man made cereal, is first produced by Muntzing. Triticale is a new crop species synthesized by crossing wheat and rye (*Secale cereale)*. a) Some triticales are hexaploids and are developed from a cross between tetraploid wheat (*Triticum turgidum)* and rye. Wheat X Rye

Triticum turgidum X *Secale cereale*

2n = 28 (Tetraploid) 2n = 14 (diploid)

(iii) Raphanobrassica: Russian geneticist G.D. Karpechenko in 1927 synthesized Raphanobrassica, which is an allopolyploid resulting from a cross between Radish and cabbage. He wanted to develop a fertile hybrid between these two species with roots of radish and leaves of cabbage. However, the F_1 hybrids, he got, were diploid having roots of cabbage and shoots of radish. They were highly sterile because of failure of each set of chromosomes to provide sufficient genetic homology to effect pairing. Among these sterile

F_1 hybrids, he found certain fertile allotetraploids which contained 36 chromosomes due to spontaneous doubling and were named as Raphanobrassica. Radish X Cabbage *Raphanus sativus* X *Brassica oleracea* 2n = 18 (diploid) 2n = 18 (diploid)

3. Segmental allopolyploids: In some allopolyploids, different genomes present are not quite different from one another. Consequently, in these polyploids, chromosomes from different genomes do pair together to some extent and mulitvalents are formed. This means that certain segments of chromosomes and not the entire chromosomes are homologous (Homeologous chromosomes). Therefore such allopolyploids are called segmental allopolyploids according to Stebbins (1943 – 1950). These are intermediate between autopolyploids and allopolyploids and can be identified by their meiotic behaviour. The common hexaploid bread wheat is also regarded as a segmental allopolyploid, because the three diploid genomes i.e. A, B and D are related (homoeologous) to each other.

Effects of polyploidy:

1. Genetical effects: The polyploidy often results in sterility. For example, an extra set of chromosomes in case of triploids is distributed in various combinations resulting in genetically imbalanced gametes.

2. Phenotypic effects: Most usual phenotypic effect of polyploidy is gigantism in morphology of plants. Eg: The tetraploid plants may have large sized pollen grains, cells of leaves, stomata, xylem etc. than a normal diploid plant. They are also more vigorous. As a result, large sized fruits, seeds and flowers are obtained from economically important plants.

3. Physiological effects: The ascorbic acid content has been reported to be higher in tetraploid cabbage and tomato than in corresponding normal diploid species. Corn flour of tetraploid maize has been found to contain 40% more vitamin A than that of normal diploid species.

Aneuploidy

It is any deviation from a euploid condition. This condition can be expressed either as an addition of one or more entire chromosome or as a loss of such chromosomes to a genomic number. Aneuploidy can be due to 1. Loss of chromosomes in mitotic or meiotic cells due to laggards (lagging chromosomes), which are characterized by retarded movement during anaphase.

2. Irregularities of chromosome distribution during meiosis of polyploids with uneven number of basic genomes like triploids and pentaploids.

3. The occurrence of multipolar mitosis resulting in irregular chromosome distribution during anaphase. Aneuploids can be of the fo llowing types:

1. Monosomy: The diploid organism which lacks one chromosome of a single homologous pair is called monosomic with genomic formula 2n-1. A monosomic produces two types of gametes n and n-1 because single chromosome without a pairing partner may go to either of poles duringmeiosis. The monosomics are usually weaker than normal diploids. Monosomics are normally found in polyploids and the diploids cannot tolerate them. The polyploids have several chromosomes of same type and therefore, this loss can be easily balanced by homologous or partially homologous chromosomes from other genomes. The number of possible monosomics in an organism will be equal to the haploid chromosome number. In common wheat, since 21 pairs of chromosomes are present, 21 possible monosomics are known. These 21 monosomics in wheat were produced by Sears in 1954 in the variety Chinese spring and are being used for genetic studies all over the world. Monosomics have been used extensively in wheat breeding for the purpose of chromosome substitution. Double monosomics (2n-1-1) or triple monosomics (2n-1-1-1) could also be produced in polyploids like wheat. In double monosomics the missing chromosomes are non-homologous.

2. Nullisomy: Diploid organisms which have lost a pair of homologous chromosomes are called nullisomics with genomic formula 2n-2. In double monosomy and nullisomy, the chromosome number is same but the genomic formula differs. In nullisomy (2n-2), a complete homologous chromosome pair is missing. In double monosomy, (2n-1-1), two chromosomes of different chromosome pairs (one each from two different chromosome pairs) are missing. Nullisomics are not usually found in natural populations, but have to be obtained by intercrossing or selfing of monosomics (2n-1). These can occur by fusion of two gametes that are lacking in the same chromosome. Nullisomic series are not of great agronomic importance, but used for genetic studies. They exhibit reduced vigour, fertility and survival. Double nullisomy (2n-2-2) involves loss of two pairs of homologous chromosomes. Monosomics and nullisomics are together known as hypoploids, which refers to loss of one or two chromosomes from the normal diploid.

3. Trisomy: Trisomics are those organisms which have one extra chromosome (2n+1). Since the extra chromosome may belong to any one of the different chromosome pairs, the number of possible trisomics in an organism will be equal to the haploid chromosome number. For instance, in n n-1 n n-1 barley (2n = 14) the haploid chromosome number is n = 7. Consequently seven trisomics are possible, in a trisomic, one of the pairs of chromosomes has an extra member and forms a trivalent during anaphase I of meiosis. Two chromosomes will go to one pole and one chromosome will go to other pole. As a result, two types of gametes are formed i.e. n and n+1. This is very common in plants and has variable effects on phenotype. In plants, the first case of trisomy was investigated in Jimson weed i.e. *Datura stramonium* by A.F. Blakeslee and J.Belling in 1924. *Datura* (2n = 24) normally has 12 pairs of chromosomes in somatic cells. But in an individual, they discovered 25 chromosomes. The size, shape and spine characteristics of seed capsule of this trisomic plant were different from seed capsule of wild type species. Through experimental breeding, Blakeslee and his associates succeeded in producing all 12 possible trisomics. When these were grown, each was found to have a distinguishable phenotype that was attributed to extra set of genes present on the extra chromosome contained in each of the 12 pairs of chromosomes. An individual having two extra chromosomes each belonging to a different chromosome pair is called double trisomic (2n + 1 + 1). Depending on the nature of extra chromosome, simple trisomics are of three types.

a) Primary trisomics: The additional chromosome is normal one in primary trisomics.

b) Secondary trisomics: Trisomics having isochromosome as additional chromosome.

c) Tertiary trisomics: When additional chromosome in a trisomic is translocated one, it is known as tertiary trisomic. The first human trisomic syndrome discovered was the one involving ‘G’ group of chromosomes called Mongolism or Down’s syndrome.

4. Tetrasomy: Tetrasomics have a particular chromosome represented four times. Therefore the general chromosome formula for tetrasomics is 2n+2. All the 21 possible tetrasomics in wheat are viable. Tetrasomics often behave more regularly than the aneuploids with odd number of chromosomes. The four homologues tend to form a quadrivalent at meiosis and disjunction often proceeds fairly regularly, two by two. Trisomics and tetrasomics are together known as hyperploids or polysomics, which refers to addition of one or two chromosomes to a single or two different homologous pairs.

Applications of aneuploids: Aneuploids are useful in crop improvement and genetic studies as detailed below:

1) Aneuploids have been used to determine the phenotypic effects of loss or gain of different chromosomes.

2) They are used to produce chromosome substitution lines. Such lines provide information on the effect of different chromosomes of a variety in the same genetic background.

3) They are also used to produce alien addition and alien substitution lines.

4) Monosomics are also used in transferring chromosomes with desirable genes from one species to another.

5) Aneuploid analysis permits the location of a gene as well as of a linkage group on to a specific chromosome. Monosomics and nullisomics are used for this purpose.

6) Studies on nullisomic and tetrasomic combinations made it possible to establish homoeology among the chromosomes of A, B and D genomes of wheat.

7) Aneuploids are also useful in identifying the chromosomes involved in translocations (tertiary trisomics).

8) Aneuploids are also useful in the preparation of molecular maps.

9) They may be used for obtaining chromosome specific probes.

(Probe is a DNA sequence that is used to detect the presence of the same DNA sequence in a test DNA sample).

A summary of terms used to describe heteroploidy (variation in chromosome number):

Heteroploid A change from diploid

A. Euploid Number of genomes or copies of a genome is more or less than two

a) Monoploid One copy of a single genome x

b) Haploid Gametic chromosome complement n

i) Monohaploid Haploid individuals that arise from a normal diploid

ii) Polyhaploid Haploid individuals that arise from a polyploid

iii) Dihaploid Diploids obtained through the chromosome doubling of haploids

c) Diploid Two copies of genome 2x

d) Polyploidy More than two copies of one genome or two copies each of two or more genomes

1. Autoployploid Genomes are identical with each other

i. Autotriploid Three copies of one genome 3x

ii. Autotetraploid Four copies of one genome 4x

iii. Autopentaploid Five copies of one genome 5x

iv. Autohexaploid Six copies of one genome 6x

v. Autoheptaploid Seven copies of one gnome 7x

vi. Autooctaploid Eight copies of one genome 8x

2. Allopolyploid Two or more distinct genomes (Generally each genome has two copies)

i Allotetraploid Two copies each of two distinct genomes ($2x_1 + 2x_2$)** or (AA BB)

ii Allohexaploid Two copies each of three distinct genomes ($2x_1 + 2_2 + 2x_3$)*** or (AA BB CC)

iii Allooctaploid Two copies each of four distinct genomes ($2x_1 + 2_2 + 2x_3 + 2x_4$)*** or (AA BB CC DD)

Term Type of change Symbol

B. Aneuploid One or few chromosomes extra or missing from 2n 2n+ few

a) Monosomic One chromosome missing 2n-1

b) Double monosomic One chromosome from each of two different chromosome pairs missing 2n-1-1

c) Nullisomic One chromosome pair missing 2n-2

d) Trisomic One chromosome extra 2n+1

e) Double trisomic One chromosome from each of two different chromosome pairs extra 2n+1+1

f) Tetrasomic One chromosome pair extra 2n+2

2.1. Introduction

Everyone has 23 pairs of chromosomes, 22 pairs of autosomes and one pair of sex chromosomes. The science that relates to the study of these chromosomes is referred to as "cytogenetics." Persons who look at chromosome preparations on slides are cytogenetic technologists or cytogeneticists. A trained cytogeneticist examines the number, shape, and staining pattern of these structures using special technologies. In this way, extra chromosomes, missing chromosomes, or rearranged chromosomes can be detected.

Studies of chromosomes begin with the extraction of whole chromosomes from the nuclei of cells. These chromosomes are then placed on glass slides, stained with special stains, and examined under a microscope. Sometimes, pictures are taken of the chromosomes on the slides, and the picture is cut into pieces so the chromosome pairs can be matched. Each chromosome pair is assigned a special number (from 1 to 22, then X and Y) that is based on their staining pattern and size.

There are many disorders that can be diagnosed by examining a person's whole chromosomes. Down syndrome, in which an individual has an extra chromosome 21, can be determined by cytogenetic studies. When there are three chromosomes in one group instead of a pair, it is referred to as a "trisomy." Missing chromosomes can also be detected, as in the case of Turner's syndrome, in which a female has only a single X chromosome. When there is only one chromosome instead of a pair, it is referred to as a "monosomy."

Abnormalities in chromosome structure are also observed with cytogenetic staining techniques. The Fragile X syndrome, the most common inherited cause of mental retardation, takes its name from the appearance of the stained X chromosome under a microscope. There is a site near the end of this chromosome that does not stain, indicating its fragility. The gene in the fragile region is important in making a special protein needed by developing brain cells.

Sometimes, pieces of a chromosome will break off and attach to another chromosome somewhere in a person's genome. When this happens, it is referred to as a "translocation." An example of a disease caused by a translocation would be chronic myelogenous leukemia (CML), in which a part of chromosome 9 breaks off and attaches itself to chromosome 22. Another example would be Burkitt lymphoma, in which a piece of chromosome 8 attaches to chromosome 14. These chromosomal translocations cause disease because the broken piece usually attaches to the new chromosome near a special gene that then becomes activated and produces tumor cells. Translocations can sometimes be seen under the microscope if a special stain is used (via conventional cytogenetic analysis).

A special technique called fluorescence in situ hybridization (FISH) can be used to view changes in chromosomes that result from genetic variations. An aberrant gene segment in a chromosome can be made to "light up" or fluoresce when it is bound by a special probe. Genetic changes in some cancers can be detected using this method. For instance, FISH is one of the methods used to determine increased copy number (amplification) of the gene *ERBB2* (also known as HER2/*neu*) in breast cancer. There are many other applications of FISH technology as well, such as chromosome microdeletions, in which a certain part of a chromosome is completely missing. In this case, the chromosome segment will not fluoresce compared to a normal set of chromosomes.

2.2. Chromosome Abnormalities

Chromosome abnormalities include inversion, insertion, duplication, and deletion. These are types of mutations. Since DNA is information, and information typically has a beginning point, an inversion would produce an inactive or altered protein. Likewise deletion or duplication will alter the gene product.

2.3. DNA and RNA Structure

DNA and RNA belong to a class of macromolecules called nucleic acids. Nucleic acids are polynucleotides which means they contain many nucleotides joined together. A nucleotide consists of:

- One cyclic five-carbon sugar (The carbons found in this sugar are numbered 1' through to 5')
- One phosphate
- One nitrogenous base

The sugar is deoxyribose in DNA and ribose in RNA. The only difference between the sugars is that ribose has a hydroxyl group (OH) on the 2' carbon and deoxyribose does not. This makes deoxyribose more stable than ribose. The phosphate is linked to the 5'

carbon of the sugar in both RNA and DNA. The nitrogenous bases are adenine(A), guanine(G), cytosine(C), thymine(T), and uracil(U). Adenine and guanine are purines (contains a six membered ring of carbon and nitrogen fused to a five membered ring). Adenine and guanine are both found in DNA and RNA. Cytosine, thymine and uracil are pyrmidines (contains a six membered ring of carbon and nitrogen).Cytosine is found in both DNA and RNA, but thymine is only found in DNA and uracil is only found in RNA. A nucleotide is formed when a phosphate attahes to the 5' carbon of the sugar and one of the nitrogenous baseses attaches to the 1' carbon of the sugar. A strand of DNA or RNA consists of nucleotides linked together by phosphodiester bonds. A phosphodiester bond exists between the phosphate of one nucleotide and the sugar 3' carbon of the the next nucleotide. This forms a backbone of alternating sugar and phosphate molecules known as the "sugar-phosphate backbone". RNA, in most cases, consists of one strand of nucleic acids linked together by phosphodiester bonds. A DNA molecule consists of two strands of nucleotides twisted together to form a double helix. The sugar-phosphate backbone is found on the outside of this helix and the bases are found braching towards the middle. Hydrogen bonds join the the nitrogenous bases and hold the two strands together.

2.4. Complementation

The two strands of DNA are complementary to one another because of the properties of base pairing : A will only pair with T by two hydrogen bonds , G will only pair with C by three hydrogen bonds. For example: If one strand is ACGTA the other strand is TGCAT. Complementation is important for storage and transmittance of genetic information

2.5. Antiparallel Strands

The two strands of DNA are also antiparallel(run in opposite directions) to one another. A strand of DNA can have the direction 5'-3' or 3'- 5'. One strand in the DNA molecule is 5'-3' and the other strand is 3'-5'. A DNA strand is assigned direction based on what is found at the end of the strands. The end of the strand with a free phosphate is the 5'

end because phosphate attaches to the 5' carbon of the sugar. The end with a free OH group is the 3' end because the OH group is attached to the 3' carbon of the sugar.

Gene Definition: A gene is a sequence of nucleotides in DNA that codes for a functional produ**ct.**

2.6. Bacterial Chromosome

A chromosome is an association of genes and some protein. A bacteria usually has a single circular chromosome. The chromosome is about 1mm long which is about 1000 times longer than the typical bacterial cell. Therefore, it is looped, folded and packed tightly inside the cell.

2.7. Protein Structure

Proteins are macromolecules that play many functions in the cell. They are used for support, storage, transport of other substances, defense against invaders, and catalytic enzymes. This is just a brief overview of protein structure. Proteins are composed of repeating units called amino acids. An amino acid consists of a carbon atom bonded to a hydrogen, a carboxyl group and an R side chain. There are 20 different amino acids that can be found in proteins. The R side chain is the variable part that makes amino acids different from one another. It is also the properties of the R group that determines how amino acids will interact with each other. Amino acids are linked by peptide bonds to form polypeptide chains. The sequence of amino acids in this chain in known as the primary structure of the protein. A protein consists of one or more of these polypeptide chains folded and coiled together. The conformation depends on the R groups present and how they interact.

3.0: VARIATION IN PLANTS AND ANIMALS

Genetics is concerned with explaining how some characteristics are passed from generation to generation, i.e. heredity, or inheritance. However, it is important to firstly deal with some background information about the characteristics themselves. This may then be tied in with evolution and the development of different species of living organisms.

Like most living organisms, humans show variation. If you consider almost any characteristic, you will find differences between various people (or other animals or plants) in a population. There are two forms of variation: continuous and discontinuous variation.

Characteristics showing continuous variation vary in a general way, with a broad range, and many intermediate values between the extremes. As a matter of fact, if you consider a large enough sample from a population, perhaps plotting frequency as a histogram or as a frequency polygon, you will find that most of the values are close to the average (mean), and extreme values are actually rather rare. Height is an example of a continuously variable characteristic, as long as you consider a consistent sample, for example a large number of people of a particular age and sex.

As you will find out later, it is usually difficult to give a straightforward explanation of the genetic basis for these continuously variable characteristics because they result from a combination of genetic factors as well as environmental influences.

Some *other* examples of human characteristics which show continuous variation are weight, foot length or any measurable dimension. Characteristics showing discontinuous variation fall into a few very distinct classes. The ability to roll the tongue, and blood groups, are examples of discontinuous variation. These characteristics can be explained much more easily by simple rules of genetics and are less likely to be affected by other factors. For the tongue-rolling ability, there are two classes i.e rollers and non-rollers. For blood group, there are four classes i.e. A, B, O and B.

3.1: Causes of variation

Some of the characteristics possessed by an individual in a population can be said to be inherited - i.e. derived from the previous generation. These characteristics are passed on, in a fairly predictable way, as a result of sexual reproduction. Sexual reproduction also introduces an element of randomness, so that variation is brought about in a population. These two almost contradictory factors - *dependable inheritance of characteristics from parents,* and *variation within the population* - are essential to an understanding of the process of evolution. Some examples of characteristics which may

be inherited by a child from his or her parents:
1. hair colour
2. eye colour
3. skin colour etc.

There is a methodical consideration of the way in which these characteristics are passed on and it is certainly too simplistic to imply that characteristics like facial features of children can merely be attributed to parents by looking at them, without knowing the background of previous generations.

The examples often chosen give the impression that inheritance covers only trivial features, such as the shape of the human chin, or ear-lobes. However, an extremely wide range of characteristics are known to be passed on in this way. In fact, practically every aspect of normal human body functioning is under hereditary (genetic) control, because there are many examples of fairly rare "conditions" (diseases which cannot be transmitted from one person to the next, but which are caused by defective functioning of certain cells) which can be inherited in exactly the same way as hair or eye colour. These inherited forms of quot; disease" may also be called inborn errors of metabolism.

Other characteristics are said to be acquired during life (non-inherited). These may also be said to be caused by environmental effects. Some examples of acquired human characteristics.
1.scars
2.fillings
3.ability to speak other languages

4.ability to ride bicycle etc.

It must also be said that some characteristics probably have both an inherited and an environmental basis, such as (possibly) I.Q. - intelligence quotient. The balance between them is the answer to the "nature *versus* nurture" question.

Similar considerations apply in all living organisms; for example different plants grown in different conditions of light or temperature may show differences in growth rate and vigour, and understanding the causes of this variation is quite fundamental in controlling or increasing agricultural and horticultural productivity.

A Gene Mutation is a very rare event indeed. A mutation in a *single inheritable characteristic* (gene) is usually less likely than one in a million, but once it has

happened, it *may* be passed on to the next generation, along the same lines as other inherited characteristics.

However, not all individuals carrying mutations survive; most have been found to be harmful, so that the organisms carrying them are at a disadvantage. In the wild, such organisms are unlikely to survive.

Nevertheless, some (beneficial) mutations confer an advantage, and others (neutral?) cause neither advantage nor disadvantage - at least until there is some reason for selection of adapted types to occur. This may be another reason for variation within a population. In fact, the few different forms resulting from mutation which are beneficial can spread through a population by natural selection, and this may have the eventual effect of changing a population so much that it differs from its original form - resulting in the evolution of a new species.

Chromosome Mutations may also result in change in the *number of chromosomes* incorporated into sex cells. A child produced as a result may have, for instance, an extra chromosome, or an extra part of a chromosome attached to the normal set. Down's syndrome is caused by having 47 chromosomes instead of the normal 46 per cell.

The frequency of babies being born with Down's syndrome is 1 in 700. Mutations as described above may occur "naturally", but in the laboratory it has been shown that they may be caused (more efficiently) by other means. Similarly, various factors in the environment may increase the chance of mutations occurring. The risks associated with some lifestyle activities are nowadays known, and exposure to many of these are avoidable. For instance, chemicals in tobacco include mutagens, and several types of ionising radiation have mutagenic effects. Some examples of treatments which may cause mutations are X rays, gamma rays and ultra-violet rays. These are in fact similar to the causes of cancer, so the idea of "natural" causes for mutations (and cancer?) is probably rather dubious.

3.2: Genetic variation

We are all unique because each of us has a unique combination of genetic variants, including SNPs in our DNA. A SNP or “snip” can affect our inherited risk of disease. A SNP, pronounced “snip”, is a single variation in the nucleotide sequence of DNA and stands for Single Nucleotide Polymorphism that can affect our inherited risk and a multitude of other characteristics. Most of our features, internal and external are determined or influenced in some way by such variations in our DNA, i.e. by SNPs.

Each of us is different because we carry a unique combination of genetic variants, including SNPs.

SNPs are the result of alterations to DNA, usually called mutations. These mutations accumulate very gradually as DNA is passed on from parent to child, generation after generation. The slow rate of change to our DNA explains why children are so like their parents. However, the fact that our DNA can change, given enough time, explains why we are all different in size, shape, color and many other characteristics. Such differences are the result of the many SNPs that have arisen in the DNA of our species and its predecessors.

Even though the differences between people around us are often easy to see, it is nonetheless important to bear in mind that humans are on average 99.9% genetically identical. This means that if you were to compare your chromosomes with those of a random person, we would expect to find, on average, one SNP that differs every thousand DNA nucleotides. Our DNA is also surprisingly similar to that of other animals and organisms. Our blood group and eye color are determined by which alleles we inherit. Alleles are different versions of the same gene. An allele is the version of a particular SNP or chromosome segment that you inherited from either your mother or father. Your cells carry 23 pairs of chromosomes, where one was inherited from your mother and the other from your father. This means that for any nucleotide or SNP located on an autosomal chromosome you have inherited two versions (one maternal and the other paternal). These are usually referred to as your two alleles for that particular location in the genome. For the vast majority of chromosome pairs, you will have inherited the same allele from both parents (for example, two copies of the cytosine nucleotide, represented by the letter C). A systematic examination of the nucleotide sequence of your 23 chromosome pairs would reveal millions of locations where the nucleotides you inherited from your parents were different (for example, a C allele from your mother and a T allele from your father).

3.3: Genetic Reshuffling: Microscopic bodies that carry our genes

We inherit 46 chromosomes from our parents, which are divided into 23 pairs. Chromosomes are compact packages of DNA contained within single cells. Unraveled, the ultra-thin strands of DNA from a single human cell are about 3 meters long. The only way for the entire body to carry such vast amounts of DNA is by winding it into complex bundles, known as chromosomes, that take up less space inside the cell nuclei. We inherit 46 chromosomes from our parents, which are divided into 23 pairs. One chromosome from each pair comes from our father and the other from our mother. It is this mix of chromosomes from our parents that determine our characteristics, including our propensity to develop various diseases.

Recombination takes place when large segments of DNA are exchanged between each pair of chromosomes. The chromosomes you inherited from your parents were actually a mosaic of chromosomes they inherited from their parents, that is, your grandparents. This is due to a process called recombination. Recombination takes place when germ cells (egg or sperm) are produced, when large segments of DNA are exchanged between each pair of chromosomes. This kind of genetic shuffling means that any chromosome you inherited from your mother is in fact a mosaic of chromosomes she inherited from her parents, and so on. This reshuffling increases the possible number of combinations of genetic variants, which in turn ensures greater variability of characteristics among individuals.

X + Y = male
X + X = female

3.4: How genes determine sex

Of the 23 chromosome pairs, 22 are known as autosomal, where the paired chromosomes are almost identical in size and content. The remaining pair consists of sex chromosomes, known as such because they carry the genes responsible for sex determination. For this pair, if you inherited an X chromosome from both mother and father, you are female. Alternatively, if you inherited an X chromosome from your

mother and a Y chromosome from your father, then you are male. The X and Y chromosomes are very different in size and content.

Self Assessment Exercise

1. See if you can find some examples of these inherited human conditions which may be passed from one generation to the next.

2. Below is a list of characteristics which can be easily noted or measured. This includes a range of continuously variable and discontinuously variable characteristics, as well as features which are likely to be acquired rather than inherited.

Carry out a survey of the class to gauge the extent of variation in these characteristics. Use the space below to record your results. *As this is recorded in a grid format, you should be able to detect a pattern of correlation between some characteristics. Use the bottom row to summarise the information.*

Class survey

	CONTINUOUS VARIATION			*DISCONTINUOUS*		*ACQUIRED*	
Name/ initials	*Height /cm*	*Arm-span /cm*	*Weight /kg*	*Tongue roller? (Y/N)*	*Ear lobe? (free/joined)*	*Scars? (Y/N)*	*Fillings? (Y/N)*
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.

3. Why is it possible to calculate mean values for the first 3 characteristics, but not the others?

4.0. MICROBIAL GENETICS

4.1. Introduction

Microbial genetics is a subject area microbiology and genetic engineering. It studies the genetics of very small (micro) organisms. This involves the study of the genotype of microbial species and also the expression system in the form of phenotypes. It also involves the study of genetic processes taking place in these micro organisms i.e., recombination etc.The flow of genetic information in bacteria can be represented by:

There are two ways that genetic information can flow:

1. Information is transferred between bacterial generations. This occurs when DNA replicates and is distributed to two identical daughter cells through the process of binary fission.
2. Information is transferred within the bacterial cell. The result of this transfer are proteins needed for cell growth and metabolism. Because DNA and protein have different chemical components, DNA must first be transcribed into mRNA and then translated into a protein.

E. coli is a bacteria that lives in your colon. It has a metabolic pathway that allows for the synthesis of the amino acid tryptophan (Trp). This pathway starts with a precurser molecule and proceeds through five enzyme catalyzed steps before reaching the final Trp product. It is important that *E. coli* be able to control the rate of Trp synthesis because the amount of Trp available from the environment varies considerably. If you eat a meal with little or no Trp, the *E. coli* in your gut must compensate by making more. If you eat a meal rich in Trp, *E. coli* doesn't want to waste valuable resources or energy to produce the amino acid because it is readily available for use. Therefore, *E. coli* uses the amount of Trp present to regulate the pathway. If levels are not adequate, the rate of Trp synthesis is increased. If levels are adequate, the rate of Trp synthesis is inhibited. There are two ways of regulating the Trp pathway: The first method works to

decrease the synthesis of Trp by inhibiting the first enzyme in the pathway, preventing the rest of the pathway from proceeding. What inhibits the first enzyme? Trp does! The more Trp in the cell, the more that can bind to the first enzyme and prevent it from catalyzing the first step. This method of regulation is feedback inhibition in which the end product of a pathway acts as an inhibitor of an enzyme in that pathway. The other method of control stops the production of the enzymes in the pathway at the transcription level. Remember that enzymes are proteins that must be transcribed and translated from the genetic code. If the genes for the enzymes are not transcribed to mRNA, then translation to the enzymes cannot occur. Without enzymes, there is no Trp synthesis. This method of control is called regulation of gene expression because control is taking place at the genetic level. This method of control will now be examined in detail.

4.2. Cell Division

Introduction

Bacteria reproduce by binary fission. This process one cell divides to give two identical daughter cells. Most of the genetic information in bacteria is found on a single circular chromosome which consists of DNA and proteins. This information must be distributed equally to each of the daughter cells. This is only possible because DNA is a self-replicating molecule and is able to make an exact copy of itself before cell division occurs. One copy will go to each daughter cell. The structure of DNA plays an important role in the replication process. Two important features to keep in mind are:

- Complementation: The two strands of DNA are complementary to one another. The base adenine will only pair with the bast thymine and the base guanine will only pair with the base cytosine. The two strands of DNA are held together by hydrogen bonds that exist between complementary bases.
- Antiparallel: The two strands of DNA are also antiparallel to one another. One strand runs in the 5' to 3' direction, the other in the 3' to 5' direction.

4.3: **Replication**: Overview

Replication begins on the bacterial chromosome at a specific sequence of nucleotides called the origin. Enzymes called helicases recognize this sequence and bind to his site. Helicases unwind that two strands by breaking the hydrogen bonds that hold the strand together. This area where DNA separates and the bases are exposed is called the replication fork. Single-stranded binding proteins attach in chains along the separated strands for stablization and to prevent rewinding. Free nucleoside triphosphates in the cytoplasm are paired up with their complementary base on the exposed parental strand. A nucleoside triphosphate is just like a nucleotide except it has three phosphates instead of one. This makes it very reactive because of the cluster of negative charge. Once it is aligned properly with its complementary base , the nucleoside triphosphate is joined to the growing strand by an enzyme called DNA polymerase. This enzyme catalyses the hydrolysis of the phosphates as the nucleotide is added to the strand, forming a phosphodiester bond.

4.3.1. DNA Polymerase

DNA polymerase has two special properties that must be taken into account:

It can not initiate the synthesis of a chain all by itself. DNA polymerase can only add a nucleotide to the 3' end of an already growing chain. Therefore, it would not be able to join the very first two replicated nucleoside triphosphates. To start synthesis it adds the first nucleoside triphosphate to a primer. This is a short stretch of RNA made by the enzyme primase. The RNA primer is a several nucleotides long and is complementary to the parental strand. DNA polymerase can then join a nucleoside triphosphate to the 3' end of the primer. The primer will be replaced later by a DNA strand and joined to the replicated strand by DNA ligase.

The other special property that must be taken into account is that DNA polymerase can only synthesize in the 5' to 3' direction. This means it can only add bases to an exposed 3' end. This has important implications for antiparallel strands running in opposite directions: Remember that a replicated strand must be antiparallel to its parent strand. This is no problem for the parent strand that runs 3' to 5' because its replicate will be 5'

to 3'. Therefore, DNA polymerase can continuously join nucleoside triphophates to the replicating strand in the way described previously. This 5' to 3' replicated strand is called the leading strand. This is a problem for the parent strand running 5' to 3' for its replicated strand must be 3' to 5'. DNA polymerase cannot add nucleoside triphophates in this direction. Instead it must work in the opposite direction away from the replication fork. This newly replicated strand is called the lagging strand because it is discontinuously synthesized away from the replication fork. We will now look at how synthesis of the lagging strand occurs:

- To begin with, primase synthesizes short RNA primers that are complementary to the 5' to 3' parental strand.
- DNA polymerase then extends these primers by joining free nucleoside triphosphates to the growing strand in the 5' to 3' direction.
- The complex, the RNA primer with the added DNA nucleotides, is called an Okazaki fragment.
- The primer will be digested by DNA polymerase and will be replaced by DNA.
- All the Okazaki fragments are joined by DNA ligase. The result is a new complementary strand running 3' to 5'.

Semiconservative Replication

Each replicated strand winds with its template parental strand to form a double helix. Because each new DNA molecule contains one conserved parental strand and one new replicated strand, the process of replication is referred to as semiconsevative replication.

Bacterial Replication

The bacterial chromosome is circular. The replication process starts at a site called the origin. But instead of replicating in one direction around the chromosome, it replicates in two directions. This is called bidirectional replication as two replication forks move in opposite directions away from the origin. The forks will meet at the bottom of the

chromosome and replication is terminated as the two chromosomes separate.The bacterial chromosome is also capable of initiating multiple replication forks. The processes of bidirectional replication and the concept of multiple replication forks are discussed in more detail in the replication laboratory. Bacteria reproduce by a type of cell division called binary fission. In this process, one cell divides to give two identical daughter cells. Most of the genetic information in bacteria is found on a single circular chromosome which consists of DNA and proteins. This information must be distributed equally to each of the daughter cells. Therefore, the bacterial cell replicates its chromosome before the process of fission takes place. The two copies of the chromosome remain attached to the membrane and the membrane simply grows between the two attached sites. After the cell has grown to about twice its normal size, the membrane pinches inward and a cell wall develops. The parent has been divided into two identical daughter cells.

4.4. Cell Division

Bacteria reproduce by a type of cell division called binary fission. In this process, one cell divides to give two identical daughter cells. Most of the genetic information in bacteria is found on a single circular chromosome which consists of DNA and proteins. This information must be distributed equally to each of the daughter cells. Therefore, the bacterial cell replicates its chromosome before the process of fission takes place. The two copies of the chromosome remain attached to the membrane and the membrane simply grows between the two attached sites. After the cell has grown to about twice its normal size, the membrane pinches inward and a cell wall develops. The parent has been divided into two identical daughter cells.

4.5. Transcription

Transcription is the synthesis of an RNA strand from a DNA template. A gene's protein building instructions are transcribed to messenger RNA (mRNA). mRNA carries the code from DNA to the ribosomes where translation into a protein occurs. Transcription occurs in three stages:

1. Initiation:

RNA polymerase binds to DNA at a specific sequence of nucleotides called the promoter. The promoter contains an initiation site where transcription of the gene begins. RNA polymerase than unwinds DNA at the beginning of the gene.

2.Elongation:

Only one of the unmound DNA strands acts as a template for the RNA synthesis. RNA polymerase can only add nucleotids to the 3' end of the strand so like DNA, RNA must be synthesized in the 5' to 3' direction. Free ribonucleotides triphosphates from the cytoplasm are paired up with their complementary base on the exposed DNA template. RNA polymerase joins the ribonucleoside triphosphates to form an mRNA strand. As RNA polymerase advances, the process continues. The DNA that has been transcribed, re-winds to form a double helix.

3.0.Termination:

RNA polymerase continues to elongate until it reaches the terminator, a specific sequence of nucleotides that signals the end of transcription. Transcription stops and mRNA polymerase and the new mRNA transcript are released from DNA. The DNA double helix reforms.The termination sequence usually consists of a series of adjacent adenines preceded by a nucleotide palindrome. This gives an RNA molecule that assumes a stem-and loop configuration. This configuration stops RNA polymerase from transcribing any further.

4.5. Components of Translation

Translation is the process by which the nucleotide sequence of mRNA is converted to the amino acid sequence of a polypeptide. In bacteria, this process takes place in the cytoplasm. In the first step of the process, all the components needed for translation come together. These components include mRNA, tRNA and ribosomal units.

mRNA

mRNA is the product of transcription. It is a single-strand of ribonucleotides that is complementary to its gene template. The purpose of mRNA is to carry the genetic code from DNA to the ribosome for translation. mRNA is read in a series of triplets called codons. For example, the mRNA sequence AUGAAGCACUAC has four codons. Each codon corresponds to one amino acid. In the above code: AUG codes for the amino acid MET, AAG codes for Lys, CAC codes for His and UAC codes for Tyr.The dictionary of the genetic code tells us which of the 20 amino acids that a codon designates. This code is redundant or degenerate because most amino acids are signalled by several different codons. For example, the amino acid leu is coded for by six different codons. The genetic code contains 3 codons signal STOP AUG signals codes for Met but also signals START.

tRNA

Transfer RNA (tRNA) is a single strand of 80 ribonucleotides. It assumes a cloverleaf configuration because of interactions between the nitrogenous bases. It functions as an interpreter between nucleic acid and peptide sequences by picking up amino acids and matching them to the proper codons in mRNA. There are two important locations on a tRNA molecule that help it do this:

At the bottom of the loop are three ribonucleotides grouped together in an anticodon. An anticodon is complementary to an mRNA codon. An anticodon can recognize and bind to its complementary mRNA codon. Some tRNAs can recognize more than one codon because there is a relaxation of the complementation rule of base pairing between the anticodon and codon in the third position. This relaxation is called the Wobble Hypothesis.

At the 3’ end of the tRNA strand is where the amino acid attaches to the tRNA molecule. Each tRNA carries one amino acid that corresponds to an mRNA codon. The proper amino acid is joined to the tRNA by the enzyme aminoacyl-tRNA synthetase. There is

one type of this enzyme for each amino acid and the active site of each fits only the specific combination of the proper amino acid and tRNA.

Ribosomes

A ribosome is made of rRNA and proteins. A ribosome is composed of two subunits, a large subunit and a small subunit. These subunits join to form a functional ribosome when they attach to mRNA. Each ribosome has two binding sites for mRNA:

1. The p-site (peptidyl site)

2. The a-site (aminoacyl site)

Translation

Now that you know the parts and functions of the components needed for translation, we will discuss the process step by step:

The small ribosomal subunit binds to the mRNA molecule. At the same time, the initiator tRNA with the anticodon UAC and the amino acid Met binds to the start codon AUG (Remember AUG is the start codon).The large subunit binds to this complex. The initiator tRNA is in the p-site and there is no tRNA in the a-site. A tRNA carrying a second amino acid approaches and enters the empty a site. The anticodon of the tRNA binds to the second codon of mRNA in the a-site.The first amino acid is joined to the second amino acid by a peptide bond. The tRNA that was in the p-site is released and the ribosome moves along the mRNA until the second tRNA is in the p-site. This means the ribosome advances only one codon. The a-site is empty again.A third tRNA approaches and enters the empty a-site. Its anticodon binds to the mRNA codon in the a-site. A peptide bond is formed between the new amino acid in the a-site and the growing peptide chain. The tRNA is released from the p-site and the ribosome advances one codon. This process continues until the ribosome reaches a STOP codon. The STOP codons are UAA, UAG and UGA. When a ribosome reaches a STOP codon, the a-site accepts a protein called a release factor instead of a tRNA. The

release factor breaks the bond between the tRNA and the polypeptide. The polypeptide and tRNA are released. The released polypeptide forms a protein.

4.6. Gene Regulation

Introduction

E. coli is a bacteria that lives in your colon. It has a metabolic pathway that allows for the synthesis of the amino acid tryptophan (Trp). This pathway starts with a precurser molecule and proceeds through five enzyme catalyzed steps before reaching the final Trp product. It is important that *E. coli* be able to control the rate of Trp synthesis because the amount of Trp available from the environment varies considerably. If you eat a meal with little or no Trp, the *E. coli* in your gut must compensate by making more. If you eat a meal rich in Trp, *E. coli* doesn't want to waste valuable resources or energy to produce the amino acid because it is readily available for use. Therefore, *E. coli* uses the amount of Trp present to regulate the pathway. If levels are not adequate, the rate of Trp synthesis is increased. If levels are adequate, the rate of Trp synthesis is inhibited. There are two ways of regulating the Trp pathway:

The first method works to decrease the synthesis of Trp by inhibiting the first enzyme in the pathway, preventing the rest of the pathway from proceeding. What inhibits the first enzyme? Trp does! The more Trp in the cell, the more that can bind to the first enzyme and prevent it from catalyzing the first step. This method of regulation is feedback inhibition in which the end product of a pathway acts as an inhibitor of an enzyme in that pathway.

The other method of control stops the production of the enzymes in the pathway at the transcription level. Remember that enzymes are proteins that must be transcribed and translated from the genetic code. If the genes for the enzymes are not transcribed to mRNA, then translation to the enzymes cannot occur. Without enzymes, there is no Trp synthesis. This method of control is called regulation of gene expression because

control is taking place at the genetic level. This method of control will now be examined in detail.

Trp Operon

The genes for the five enzymes in the Trp synthesis pathway are clustered on the same chromosome in what is called the Trp operon. The Trp operon has three components:

Five Structural Genes: These genes contain the genetic code for the five enzymes in the Trp synthesis pathway

One Promoter:

DNA segment where RNA polymerase binds and starts transcription

One Operator:

DNA segment found between the promoter and structural genes. It determines if transcription will take place. If the operator is turned "on", transcription will occur. When nothing is bonded to the operator, the operon is "on". RNA polymerase binds to the promoter and transcription is initiated. The five structural genes are transcribed to one mRNA strand. The mRNA will then be translated into the enzymes that control the Trp synthesis pathway. The operon is turned "off" by a specific protein called the repressor. The repressor is a product of the regulator gene which is found some distance from the operon. Transcription of the regulator produces mRNA which is translated into the repressor. The repressor is inactive in this form and cannot bind properly to the operator with this conformation. To become active and bind properly to the operator, a co-repressor must associate with the repressor. The co--repressor for this system is Trp. This makes sense because *E. coli* does not want to synthesize Trp if it is available from the environment. The more Trp available, the more that can associate with repressor molecules. An active repressor binds to the operator blocking the attachment of RNA polymerase to the promoter. Without RNA polymerase, transcription and translation of

the structural genes can't occur and the enzymes needed for Trp synthesis are not made.

Repressible vs Inducible Systems

The Trp pathway is anabolic as Trp is being synthesized. The Trp and other regulated anabolic pathways are usually repressible because the system can be repressed by an overabundance of the end product. The end product, Trp, in this case, decreases or stops the transcription of the enzymes necessary for its production. Regulated catabolic pathways, on the other hand, are usually inducible because the pathway is stimulated rather than inhibited by a specific molecule. An example of an inducible system is lactose metabolism.

The Lac Operon

The genes that code for the enzymes needed for lactose catabolism are clustered on the same chromosome in what is called the Lac operon. The Lac operon has three components - Three Structural Genes:

These contain the genetic code for the three enzymes in the lac catabolic pathway

One Promoter:

DNA segment where RNA polymerase binds and starts transcription

One Operator:

DNA segment found between the promoter and structural genes. It determines if transcription will take place. If the operator in turned "on", transcription will occur. As in the Trp operon, the Lac operon is turned "off" by a specific protein called the repressor. The repressor is the product of the regulator gene which is found outside the operon. Transcription of the regulator produces mRNA which is translated into the repressor. But unlike the Trp operon, the repressor is active in this form and does not require a co-

repressor. The active repressor binds to the operator blocking the advancement of RNA polymerase to the structural genes. Without RNA polymerase, transcription and translation of the genes can't occur and the enzymes needed for Lac metabolism are not made. What turns the Lac operon "on"? Lactose does!

This makes sense because the cell only needs to make enzymes to catabolize lactose if lactose is present. When lactose enters the cell, allolactose, an isomer of lactose is formed. Allolactose binds to the repressor and alters its conformation so that it can't bind to the operator. RNA polymerase can now start transcription. The three structural genes are transcribed to one mRNA strand. The mRNA will then be translated into the enzymes that control lactose catabolism. In this sense, allolactose is an inducer.

Negative vs Positive Control

While the Trp operon is an example of repressible gene regulation and the Lac operon is an example of inducible gene regulation, both are examples of negative control of genes because both operons are shut "off" by an active repressor. Gene regulation would be positive, on the other hand, if an activator molecule turned the operon "on".

The Lac operon is also an example of a positive control system and is turned on by the cAMP-CAP complex, as the next section explains. *E. coli* can be described as a fussy eater. Its first choice at every meal is glucose because glucose supplies maximum energy for growth. Therefore, *E. coli* will only metabolize lactose if concentrations of glucose are low. For this to work, there must be a signal to tell the Lac operon that glucose is not available and to start transcribing the genes to metabolize lactose. This signal is a small molecule called cyclic AMP (cAMP). The amount of cAMP present in a cell is inversely proportional to the amount of glucose present. As a result, the absence of glucose results in an increase in cAMP in the cell. The following describes the situation where there is lactose but no glucose available to the cell: No glucose means high levels of cAMP. cAMP binds to a molecule known as CAP. CAP, when in association with cAMP, can bind to the promoter at the CAP binding site. Here, the cAMP-CAP complex stimulates transcription by helping RNA polymerase bind to the

promoter. RNA polymerase has a weak affinity for the Lac promoter and will not bind without this help. Remember with lactose present so is allolactose. Allolactose binds to the repressor and prevents it from binding to the operator. Therefore, transcription and translation of the genes can occur. The following depicts what happens when glucose and lactose are both present for *E. coli* to metabolize:

With glucose present, there is very little or no cAMP. It cannot bind to the CAP binding site. Without this complex, RNA polymerase cannot bind to the promoter and transcription cannot occur. Even though allolactose is present and blocks the action of the repressor, there is no transcription of the lac genes because glucose is present.

4.7. Mutations

Point Mutations

Point mutations are the most common type of mutation. A single point mutation, also called a base substitution, occurs when a single nucleotide is replaced with a different nucleotide. A point mutation results in a base pair substitution after replication and possibly a mutant protein after transcription and translation. There are three types of point mutations:

Silent Mutation:

A silent mutation causes no change in the activity of the protein. A silent mutation is usually the result of a substitution occurring in the third location of the mRNA codon Because the gentic code is degenerate (most amino acids are coded for by several alternative codons), the resulting new codon may still code for the same amino acid.

Missence Mutation:

A missence mutation is a nucleotide substitution that changes a codon so that it codes for a different amino acid in the protein. This usually results in a change of the activity of the protein. The change may be harmful or beneficial to the protein.

Nonsense Mutation:

A nonsense mutation is the same as a missense mutation except the resulting codon codes for a STOP signal. The result is a premature termination of translation. The protein is shorter than usual and does not contain all the amino acids that it should. Therefore, this protein is most likely nonfunctional.

Frameshift Mutations

Another type of mutation is a frameshift mutation which is caused by the insertion or a deletion of a base pair. An inserted or deleted nucleotide alters the triplet grouping of nucleotides into codons and shifts the reading frame so that all nucleotides downstream from the mutation will be improperly grouped. The result is a protein with extensive missense ending sooner or later in nonsense.

Mutagens

A mutation can be the result of different events. Errors made during replication, repair, or recombination can all lead to point or frameshift mutations. Mutations resulting from such errors are spontaneous mutations. A mutation can also result from the action of physical and chemical agents known as mutagens. We will now explore three mutagens: nitrous acid, base analogs, and UV light.

Nitrous Acid:

Nitrous Acid affects DNA complementation. The acid randomly modifies the base adenine so that it will pair with cytosine instead of thymine. This change is made evident during DNA replication when a new base pair appears in daughter cells in a later generation.

A Base Analog:

A base analog is a compound sufficiently similar to one of the four DNA bases but have different pairing properties. For example, 5-bromouracil is the analog of thymine but sometimes pairs with guanine and 2-aminopurine is the analog of adenine but sometimes pairs with cytosine. The incorporation of a base analog will to a base pair substitution in that appears in daughter cells in a later generation.

UV Light:

Exposure to direct UV light induces covalent linking between adjacent thymine nucleotides on a DNA strand forming a thymine dimer. These dimers cause the strand to buckle, disrupting normal base pairing. This prevents proper replication and transcription. Bacteria have enzymes to fix the damage created by UV light. An enzyme cuts the DNA at two points and removes the damaged portion. DNA polymerase synthesizes a new DNA segment using the healthy strand as a template. DNA ligase joins the new fragment to the old strand.

Mutation Rate

Mutations are random events and there is no way of knowing when a mutation will occur. Genes do, however mutate spontaneously at a characteristic rate, making it possible to assign probabilities to certain mutation events. The probability that a gene will mutate when a cell divides is called the mutation rate. Spontanoeus mutation rate for the average gene is 0.000000001. This means a mutation event is estimated to occur once in every million genes replicated. The presence of a mutagen increases the rate of mutation to 0.00001 to 0.001. This means that a mutation event is estimated to occur once in every hundred thousand to one hundred thousand genes in the presence of a mutagen.

Mutant Isolation

How can you tell if there are any mutant colonies in a culture? By either positive (direct) selection or by negative (indirect) selection.

Positive Selection:

Positive selection entails the growing the culture on a medium that will alow for the growth of only the mutant colonies. If, for example, we want to find mutants that resistant to penicillin, we grow the culture on a medium that contains pencillin. Only those colonies that are resistant to penicillin will grow and we ca identify them directly.

Negative Selection:

Negative selection is used to identify mutants that have lost the ability to perform a certain function that their parents had. Auxotrophic mutants, for example, are bacteria that have lost the ability to synthesize an essential nutrient.The replica-plating technique is used to identify mutants by negative selection. The replica-plating technique can be used, for example, to identify mutants that have lost the ability to synthesize the amino acid histidine. Therefore, mutants are His- and require histidine in order to survive.

Inoculate a histidine enriched medium with bateria. Incubate so that cells can form colonies. This is the master plate. Press a sterile velvet surface into the colonies of the master plate. Some cells from each of the colonies adhere to the velvet. Prepare two mediums, one with histidine, the other without histidine. Transfer cells from the velvet to each plate. Compare growth on the two plates after incubation. Colonies that grow on the histindine enriched medium but not on the medium lacking histidine are His- mutants.

4.8. Gene Transfer

Bacteria reproduce by the process of binary fission. In this process, the chromosome in the mother cell is replicated and a copy is allocated to each of the daughter cells. As a result, the two daughter cells are genetically identical. If the daughter cells are always identical to the mother, how are different strains of the same bacterial species created? The answer lies in certain events that change the bacterial chromosome and then these changes are passed on to future generations by binary fission. In this chapter, you will

explore some of the events that result in heritable changes in the genome: genetic transfer and recombination, plasmids and transposons.

4.9. Recombination

Genetic recombination refers to the exchange between two DNA molecules. It results in new combinations of genes on the chromosome. You are probably most familiar with the recombination event known as crossing over. In crossing over, two homologous chromosomes (chromosomes that contain the same sequence of genes but can have different alleles) break at corresponding points, switch fragments and rejoin. The result is two recombinant chromosomes. In bacteria, crossing over involves a chromosome segment entering the cell and aligning with its homologous segment on the bacterial chromosome. The two break at corresponding point, switch fragments and rejoin. The result, as before, is two recombinant chromosomes and the bacteria can be called a recombinant cell. The recombinant pieces left outside the chromosome will eventually be degraded or lost in cell division. But one question still remains...how did the chromosome segment get in to the cell? The answer is Genetic Transfer!

4.10. Genetic Transfer

Genetic transfer is the mechanism by which DNA is transferred from a donor to a recipient. Once donar DNA is inside the recipient, crossing over can occur. The result is a recombinant cell that has a genome different from either the donar or the recipient.

In bacteria genetic transfer can happen three ways:

1. Transformation
2. Transduction
3. Conjugation

Remember that a recombination event must occur after transfer in order that the change in the genome be heritable(passed on to the next generation).

4.11. Transformation

After death or cell lyses, some bacteria release their DNA into the environment. Other bacteria, generally of the same species, can come into contact with these fragments, take them up and incorporate them into their DNA by recombination. This method of transfer is the process of transformation. Any DNA that is not integrated into he chromosome will be degraded. The genetically transformed cell is called a recombinant cell because it has a different genetic makeup than the donar and the recipient. All of the descendants of the recombinant cell will be identical to it. In this way, recombination can give rise to genetic diversity in the population.

Griffith's Experiment

The transformation process was first demonstrated in 1928 by Frederick Griffith. Griffith experimented on *Streptococcus pneumoniae*, a bacteria that causes pneumonia in mammals. When he examined colonies of the bacteria on petri plates, he could tell that there were two different strains. The colonies of one strain appeared smooth. Later analysis revealed that this strain has a polysaccharide capsule and is virulent, that it, it causes pneumonia. The colonies of the other strain appeared rough. This strain has no capsules and is avirulent. When Griffith injected living encapsulated cells into a mouse, the mouse died of pneumonia and the colonies of encapsulated cells were isolated from the blood of the mouse. When living nonencapsulated cells were injected into a mouse, the mouse remained healthy and the colonies of nonencapsulated cells were isolated from the blood of the mouse. Griffith then heat killed the encapsulated cells and injected them into a mouse. The mouse remained healthy and no colonies were isolated. The encapsulated cells lost the ability to cause the disease. However, a combination of heat-killed encapsulated cells and living nonencapsulated cells did cause pneumonia and colonies of living encapsulated cells were isolated from the mouse. How can a combination of these two strains cause pneumonia when either strand alone does not cause the disease? If you guessed the process of transformation you are right! The living nonencapsulated cells came into contact with DNA fragments of the dead

capsulated cells. The genes that code for the capsule entered some of the living cells and a crossing over event occurred. The recombinant cell now has the ability to form a capsule and cause pneumonia. All of the recombinant's offspring have the same ability. That is why the mouse developed pneumonia and died.

4.12. Transduction

Another method of genetic transfer and recombination is transduction. This method involves the transfer of DNA from one bacterium to another with the use of a bacteriophage (phage). A phage is a virus that infects bacteria. The phage T4 and the phage lambda, for example, both infect *E. coli*. Because the phage reproductive system is important to understanding transduction, we will briefly review phage lifecycle. Phages are obligatory intracellular parasites and must invade a host cell in order to reproduce. T4 multiplies by the lytic cycle which kills the host and lamba multiplies by the lysogenic cycle which does not cause the death of the host cell. In lysogeny, the phage DNA remains latent in the host until it breaks out in a lytic cycle.

General Steps Of The Lytic Cycle:

1. Attachment of T4 to receptors on *E. coli* cell wall.
2. Penetration of the cell wall by tail core. Inject DNA into host.
3. *E. coli* DNA is hydrolyzed. Phage DNA directs biosynthesis of viral parts using the host cell's machinery.
4. The phages mature as the parts are assembled.
5. Lyses of *E. coli* and release of the new phages.

General Steps Of The Lysogenic Cycle:

Phage attaches to *E. coli* and injects DNA. Phage circularizes and can enter either the lytic or the lysogenic cycle. The lytic cycle would occur as previously described. In the lysogenic cycle the circular phage DNA recombines with *E. coli* DNA and the phage DNA is now called prophage. *E. coli* undergoes cell division, copying prophage and

passing to daughter. With more divisions there are more cells with the prophage. The prophage may exit the chromosome and start a lytic cycle at any time. Now that we have reviewed phage lifecycles, we can discuss transduction. Transduction can be generalized or specialized.

The Steps Of General Transduction:

1. A phage attaches to cell wall of bacterium and injects DNA.
2. The bacterial chromosome is broken down and biosynthesis of phage DNA and protein occurs.
3. Sometimes bacterial DNA can be packaged into the virus instead of phage DNA.
4. This phage is defective (can't destroy another host cell) because it does not carry its own genetic material.
5. The cell lyses, releasing viruses. The phage carrying bacterial DNA infects another cell. Crossing over between donor and recipient DNA can occur producing a recombinant cell.
6. In generalized transduction, any bacterial genes can be transferred because the host's chromosome is broken down into fragments.

Whatever piece of bacterial DNA happens to get packaged within the phage is the genetic material that will be transferred between cells. In specialized transduction, on the other hand, only certain bacterial genes can be transferred. These genes, as you will see, must exist on either side of the prophage. Specialized transduction requires a phage that uses the lysogenic cycle for reproduction.

The Steps In Specialized Transduction:

1. Remember that in the lysogenic cycle, phage DNA cn exist as a prophage integrated in the bacterial chromosome)
2. Occasionally when the prophage exits it can take adjacent bacterial genes with it.
3. The phage DNA directs synthesis of new phages.

4. The phage particles carry phage DNA and bacterial DNA.
5. The cell lyses, releasing the phages.
6. A phage carrying bacterial DNA infects another cell.
7. The joined phage and bacterial DNA circularize.
8. Along with the prophage, bacterial DNA integrates with the recipient chromosome by a cross over event. This forms a recombinant cell.

4.13. Conjugation

A third mechanism by which genetic transfer takes place is conjugation. This mechanism requires the presence of a special plasmid called the F plasmid. Therefore, we will briefly review plamid structure before continuing. Plasmids are small, circular pieces of DNA that are separate and replicate indepentently from the bacterial chromosome. Plasmids contain only a few genes that are usually not needed for growth and reproduction of the cell. However, in stressful situations, plasmids can be crucial for survial. The F plasmid, for example, facilites conjugation. This can give a bacterium new genes that may help it survive in a changing environment. Some plasmids can integrate reversibly into the bacterial chromosome. An integrated plasmid is called an episome.Bacteria that have a F plasmid are referred to as as F+ or male. Those that do not have an F plasmid are F- of female. The F plasmid consists of 25 genes that mostly code for production of sex pili. A conjugation event occurs when the male cell extends his sex pili and one attaches to the female. This attached pilus is a temporary cytoplasmic bridge through which a replicating F plasmid is transferred from the male to the female. When transfer is complete, the result is two male cells.The F plasmid can behave as an episome. When the F+ plasmid is integrated within the bacterial chromosome, the cell is called an Hfr cell (high frequency of recombination cell). The F plasmid always insetrs at the same spot for a bacterial species. The Hfr cell still behaves as a F+ cell, transferring F genes to a F-cell, but now it can take some of the bacterial chromosome with it. Replication of the Hfr chromosome begins at a fixed point within the F episome and the chromosome is transferred to the female as it replicates. Movement of the bacteria usually disrupts conjugation before the entire chromosome,

including the tail of the F episome can be transferred. Therefore, the recipient remains F- because the F plasmid is not entirely transferred. A cross over event can occur between homologous genes on the Hfr fragment and the F- DNA. Pieces of DNA not recombined will be degraded or lost in cell division. Now the recombinant genome can be passed on to future generations.

4,14, Plasmids

Plasmids are genetic elements that can also provides a mechanism for genetic change. Plasmids, as we discussed previously, are small, circular pieces of DNA that exist and replicate separately from the bacterial chromosome. We have already seen the importance of the F plasmid for conjugation, but other plasmids of equal importance can also be found in bacteria. One such plasmid is the R plasmid. Resistance or R plasmids carry genes that confer resistance to certain antibiotics. A R plasmid usually has two types of genes:

1. R-determinant: resistance genes that code for enzymes that inactivate certain drugs
2. RTF (Resistance Transfer Factor): genes for plasmid replication and conjugation.

Without resistance genes for a particular antibiotic, a bacterium is sensitive to that antibiotic and probably destroyed by it. But the presence of resistance genes, on the other hand, allows for their transcription and translation into enzymes that make the drug inactive. Resistance is a serious problem. The widespread use of antibiotics in medicine and agriculture has led to an increasing number of resistant strain pathogens.

These bacteria survive in the presence of the antibiotic and pass the resistance genes on to future generations. R plasmids can also be transferred by conjugation from one bacterial cell to another, further increasing numbers in the resistant population.

4.15. Transposons

Transposons (Transposable Genetic Elements) are pieces of DNA that can move from one location on the chromosome another, from plasmid to chromosome or vice versa or from one plasmid to another. The simplest transposon is an insertion sequence. An insertion sequence contains only one gene that codes frotransposase, the enzyme that catalyzes transposition. The transposase gene is flanked by two DNA sequences called inverted repeats because that two regions are upside-down and backward to each other. Transposase binds to these regions and cuts DNA to remove the gene. The transposon can enter a number of locations. When it invades a gene it usually inactivates the gene by interrupting the coding sequence and the protein that the gene codes for. Luckily, transposition occurs rarely and is comparable to spontaneous mutation rates in bacteria. Complex transposons consist of one or more genes between two insertion sequences. The gene, coding for antibiotic resistance, for example, is carried along with the transposon as it inserts elsewhere. It could insert in a plasmid and be passed on to other bacteria by conjugation.

4.16. Viral Genetics

Viruses have a simple structure consisting usually of nucleic acid packed into a protein head. They lack the metabolic machinery for isolated mutiplication and must invade a host cell in order to reproduce. This parasitic lifestyle gives rise to some interesting reproductive cycles which we will explore in this chapter.

Viral Structure

A virus is an obligate intracellular parasite. This means it must invade a host cell in order to reproduce. An isolated virus can not replicate itself or carry on metabolism because it lacks many of the enzymes and structures necessary for reproduction, protein synthesis and ATP generation. Therefore, it invades and takes control of a host’s metabolic machinery in order to multiply. Viruses are extremely small, ranging in size 20 to 14000 nm in length. They are structured to lead a parasitic life, composed of mostly nucleic acid surrounded by a protein coat.

Viral Nucleic Acid:

A virus contains only one type of nucleic acid which is either DNA or RNA. Yet, this genetic material can exist in many different forms. It can be double- stranded DNA, single- stranded DNA, double- stranded RNA or single- stranded RNA. The nucleic acid can be circular or linear.

The Protein Coat:

The protein coat or capsid as it is called, surrounds the nucleic acid and protects it from the environment. A capsid is composed of repeating units called capsomeres. The proteins making up the capsomere is determined by the viral nucleic acid. A virus may also be covered by an envelope composed of lipids,proteins and carbohydrates. The envelope may also have spikes projecting from the surface. Whether a virus has an envelope or spikes is again determined by the nucleic acid.

Host Range

A virus can only infect and reproduce within certain host cells. Viruses identify their hosts by specific receptor molecules on the outside of the host cell. Some viruses have broad host ranges and can infect many different types of cells. Other viruses, on the other hand, have extremely narrow ranges and can infect only very specific cells.

General Lifecycle

The general life cycle of a virus can be described by:

1)Recognition of Host

2)Genetic Material Enters Host

3)Replication Using Host Nucleotides

4)Protein Synthesis Using Host Enzymes, Ribosomes, tRNA, ATP

5)Self-assembly Of Capsids And Packaging Of Genome

6)Release From Host

There are many variations of this cycle depending on the type of virus and the host.

We will now examine bacteriophage lifecycle.

4.17. Bacteriophages

A bacteriophage (phage) is a virus that infects bacteria. The phage T4 and the phage lambda, for example, both infect E. coli. Phages, like all viruses, are obligatory intracellular parasites and must invade a host cell in order to reproduce. Viruses can multiply by two alternate mechanisms: the lytic or the lysogenic cycle. T4 multiplies by the lytic cycle which kills the host and lamba multiplies by the lysogenic cycle which does not cause death of the host cell. In lysogeny, the phage DNA remains latent in the host until it breaks out in a lytic cycle.

The Lytic Cycle

The lytic cycle can be described as follows:

1)Attachment: The T4 phage is a complex virus and has several tail fibers. T4 uses these fibers to attach to complementary receptor sites on E. coli's cell wall. Weak chemical bonds form between the attachment and receptor sites, adhering the virus to the host.

2)Penetration:

T4 injects its DNA into E. coli by releasing the enzyme, phage lysozyme, which breaks down a portion of the cell wall. T4 contracts is tail sheath driving the tail core into E. coli. The viral DNA passes through the core and into the cell. The capsid remains outside.

3)Biosynthesis:

Host protein synthesis is stopped by viral degradation of host DNA, and interference of host transcription and translation. T4 uses host nucleotides to replicate its DNA and host ribosomes, enzymes and amino acids to synthesize its enzymes and proteins. During this time no complete phages are found in the host and is called the eclipse period.

4)Maturation:

Spontaneous assembly of capsids and packaging of DNA inside the head. Tails fibers join the complex.

5)Release

Release of the viruses from the host.

The viral enzyme lysozyme, lyses (breaks open) E. coli's plasma membrane and cell wall. The E. coli cell dies. Burst time is the time from attachment of a virus to lyses and release of the new phage particles. Burst time ranges from 20 to 40 minutes.Burst size is the number of viruses released from the cell at the burst time. Burst size ranges from 50 to 200 viruses.

The Lysogenic Cycle

Some viruses, like phage lambda, can enter a lytic cycle or alternatively, may enter a lysogenic cycle. Phages that can multiply by either cycle are lysogenic or temperate phages. When lambda enters a lysogenic cycle, its DNA recombines with E. coli's chromosome. The inserted phage DNA is called a prophage. Repressor proteins, which are products of the prophage, prevent transcription of the other prophage genes. As a result, the synthesis and release of new phages is repressed. Every time the bacterial chromosome is replicated, prophage DNA is replicated, as well. The prophage can remain latent in the bacterial chromosome for many generations. A spontaneous event at any time may cause the virus to break out of its latent state and enter the lytic cycle.

5.0. BIOCHEMICAL AND BIOMEDICAL GENETICS

5.1. **Introduction**

It is obvious to us all that we all differ from one another, and that many of these differences ‘run in families’. Apart from identical twins people can readily be distinguished, from their facial features and many other attributes. This high level of individuality is reflected in our DNA. Both non-coding and coding DNA show a great deal of person to person variation. The existence of variation at a molecular and biochemical level has been known for nearly 100 years, that is long before the first human DNA sequences were read. Some of the early biochemical traits to be identified took more than 50 years to be elucidated at the level of the gene. In contrast more recently the gene defect in many genetic diseases has been elucidated by sequencing of DNA with no knowledge at all of the function of the protein product.

Genetic variation in functional regions of the genome can fall into several different categories with respect to its effect on the individual and the frequency of the allele in the population.

5.2. Variation in proteins

A single amino-acid substitution can have very severe effects but it may be unimportant if it is outside critical regions of functional importance.

Enzyme deficiencies: Inborn errors of metabolism

These are rare disorders in which an enzyme is deficient - which causes a block or 'error' in a metabolic pathway. They are usually recessive. Enzymes are catalysts - so that the half levels present in heterozygotes are sufficient, and these individuals are usually completely unaffected. The first inborn errors were described early in the 20th century by Sir Archibald Garrod. Garrod's concept of these disorders came mainly from his studies on the rare disorder Alkaptonuria. This is a relatively benign disorder but often diagnosed in infancy because of brown discoloration of the baby's nappy. The

disorder is characterised by the massive urinary excretion of the substance homogentisic acid which is not normally found in the urine. Although the affected individuals are usually quite healthy - in later life they are particularly prone to develop a form of arthritis known as ochronosis - because of deposition of a substance derived from homogentisic acid.

Garrod observed that frequently more than one sibling in a family was affected and that often the parents were related (the marriages were consanguinous) and learned that these observations could be readily explained if the defect was inherited as a recessive condition in terms of the recently rediscovered laws of Mendel. Garrod was able to predict the enzyme deficiency. As recently as 1997 the gene encoding homogentisic acid oxidase was cloned and the first mutations identified. From his study of alkaptonuria Garrod developed the concept that certain diseases of lifelong duration arise because an enzyme governing a single metabolic step is reduced in activity or missing. Many inborn errors of metabolism are now known. One of the most common is phenylketonuria which gives rise to a clinical disorder which includes severe mental retardation. Phenylketonuria is due to deficiency of the enzyme phenylalanine hydroxylase which converts phenylalanine to tryrosine (see the pathway above). An effective treatment of this condition is to withhold phenylalanine from the diet.

5.3. Defects of abundant and structural proteins

- Haemoglobin, the major protein in red cells, which gives blood its red colour was the earliest human protein to be studied in detail in relation to its genetic variability. The abundance of the protein meant that it could be sequenced directly and its colour meant that it could be readily detected. A very large number of variant haemoglobins are now known and many of these lead to disease. A well known example is sickle cell haemoglobin which in homogyzotes leads to sickle cell anaemia.

  The single substitution of valine for glutamic acid at position six of the beta-globin polypeptide chain gives rise to sickle cell disease in homozygotes because the

modified chain has a tendency to crystallise at low O_2 concentrations. The amino-acid substitution causes changes in electrophoretic mobility so that HbS and HbA can be separated by electrophoresis. The sickle cell trait in heterozygous carriers confers resistence to malaria.

- Collagen is a family of related structural proteins which are vital to the integrity of many tissues including skin and bones. The mature collagen molecule is comprised of three polypeptide chains wound in a triple helix. The chains first associate at their C termini and then twist together in a C to N direction. To be able to accomplish this collagen polypeptide chains have a special repeating structure of three amino acids, glycine - X - Y (X is usually proline and Y can be any of a wide range of aminoacids). A point mutation which by changing a single amino acid disrupts either the association of chains at their C termini or which prevents the triple helix formation may have severe consequences. One mutant chain can disrupt a triple helix with two wild-type chains, leading to a condition with **dominant inheritance**.

5.4. The first single gene polymorphisms

A polymorphism is defined as the occurrence in a population of two or more common alleles. A locus is regarded **polymorphic** if the frequency of the rarest allele is more than 0.01 - ie if the heterozyote frequency is 2% or more.

The existence of human blood groups has been known since the beginning of the 20th century. Landsteiner showed that when suspensions of red blood cells obtained from different people are mixed with blood serum obtained from other people that the red cells often agglutinated and a clear cut pattern of differences in reaction was observed. By working out the patterns of the agglutinations he first defined the ABO blood groups. Others (Epstein and Ottenberg) showed that the ABO blood group was inherited as a **Mendelian trait**.

This was the first example of a human **polymorphism** (other than eye colour) inherited as a single Mendelian trait.

ABO genotypes and phenotypes			
Genotype	Phenotype	red cell antigens	serum antibodies
AA	A	A	anti-B
AO	A	A	anti-B
BB	B	B	anti-A
BO	B	B	anti-A
AB	AB	A and B	neither
OO	O	neither	anti-A and anti-B

The ABO gene codes for a glycosyltransferase which adds a sugar residue to a carbohydrate structure known as the H antigen on the surface of red blood cells. The A allele codes for an enzyme which adds N-acetylgalactosamine, whereas the enzyme coded for by the B allele has two amino acid differences which alter is specificity and it adds D-galactose, forming the A and B antigens respectively. The O allele has a frameshift muation in the gene and thus produces a truncated and inactive product which cannot modify H. A phenotype people have natural antibodies to B antigen in their serum and vice versa. O phenotype individuals have antibodies directed against both A and B. AB individuals have no antibodies against either A or B antigens.

5.5. Common variations associated with disease susceptibility - pharmacogenetics

It has been known for a long time that common genetically determined variations which are not disease-causing in themselves can lead to disease in response to environmental exposure to substances to which other individuals do not respond adversely.

The first examples of this were discovered in the early 1950s. It seems very possible there will be a rush of new examples in the next few years as we enter the 'post-genome' era.

Inherited variations of serum cholinesterase came to light following the introduction and wide-spread use of the drug suxamethonium (succinyl dicholine) as a muscle relaxant in surgery. The effects of this drug are quite short lived in most people - but 1 in 2000 are unusually sensitive because the drug takes longer to degrade. It is usually metabolised by serum cholinesterase which hydrolyses the molecule as indicated. Some individuals carry a variant serum cholinesterase with reduced activity and altered kinetic properties.

Also in the 1950s it was noted that in some populations a significant number of individuals respond adversely to the anti-malarial drug primaquine or that some individuals are sensitive to fava (or broad) beans. These agents cause acute haemolysis in these people.

The sensitivities are due to deficiency of the enzyme glucose-6-phosphate dehydrogenase (G6PD). Males are more commonly affected than females. This is because the gene encoding G6PD is X-linked.

Measurement of the level of G6PD activity in different individuals shows that the distribution of activities is bimodal in males in these populations, and broad and almost unimodal in females. The levels of activity in females are not twice the level that are seen in males, because only one of the X chromosomes is active in each cell.

5.6. Common variations with no known (as yet) functional significance.

With the advent of the technique of gel electrophoresis, it was possible to assess the frequency of normal genetically determined polymorphism of proteins. This is because electrophoresis can readily separate proteins which differ in charge. Because the allelic forms of the protein can be separated, **codominant** inheritance can be observed.

Electrophoresis is the separation of charged particles in an electric field. The charge on protein depends on the ionisation of the R groups of the amino-acids.

'basic' proteins i.e. those rich in lysine, arginine, histidine are cathodal

'acid' proteins i.e. those rich in aspartic acid and glutamic acid are anodal.

5.7. Variations in non-coding, non-functional DNA

A wide variety of changes can result from mutations in DNA. These include single nucleotide changes as well as large deletions, insertions and rearrangements. Mutations will be covered in more detail in a later lecture.

Nucleotide changes in coding DNA:

Here we will consider just one example - mutations in the coding region of the gene which lead to an amino-acid substitution through alteration of a codon. For example: Aginine is positively charged at neutral pH because of the ionisation of the amino groups shown in bold.

Self Assessment Question

1. A protein is subjected to electrophoresis at pH7.0 but fails to migrate from the start line. What is a possible explanation? (Assuming all the electrical connections were made properly!)

6.0. HUMAN GENETICS

6.1. The human karyotype

There are 44 autosomes and 2 sex chromosomes in the human genome, for a total of 46. Karyotypes are pictures of homologous chromosomes lined up together during

Metaphase I of meiosis. The chromosome micrographs are then arranged by size and pasted onto a sheet.

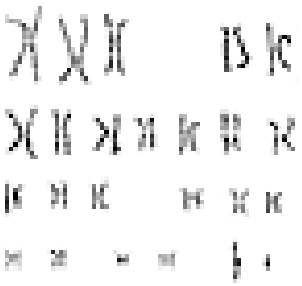


Fig. 2. This picture is from The Primate Cytogenetics Network at (http://www.selu.com/~bio/cyto/karyotypes/Hominidae/Hominidae.html).

6.2. Human chromosomal abnormalities.

A common abnormality is caused by nondisjunction, the failure of replicated chromosomes to segregate during Anaphase II. A gamete lacking a chromosome cannot produce a viable embryo. Occasionally a gamete with n+1 chromosomes can produce a viable embryo.

In humans, nondisjunction is most often associated with the 21st chromosome, producing a disease known as Down's syndrome (also referred to as trisomy 21). Sufferers of Down's syndrome suffer mild to severe mental retardation, short stocky body type, large tongue leading to speech difficulties, and (in those who survive into middle-age), a propensity to develop Alzheimer's Disease. Ninety-five percent of Down's cases result from nondisjunction of chromosome 21. Occasional cases result from a translocation in the chromosomes of one parent. Remember that a translocation occurs when one chromosome (or a fragment) is transferred to a non-homologous chromosome. The incidence of Down's Syndrome increases with age of the mother, although 25% of the cases result from an extra chromosome from the father.

Sex-chromosome abnormalities may also be caused by nondisjunction of one or more sex chromosomes. Any combination (up to XXXXY) produces maleness. Males with

more than one X are usually underdeveloped and sterile. XXX and XO women are known, although in most cases they are sterile. What meiotic difficulties might a person with Down's syndrome or extra sex-chromosomes face?

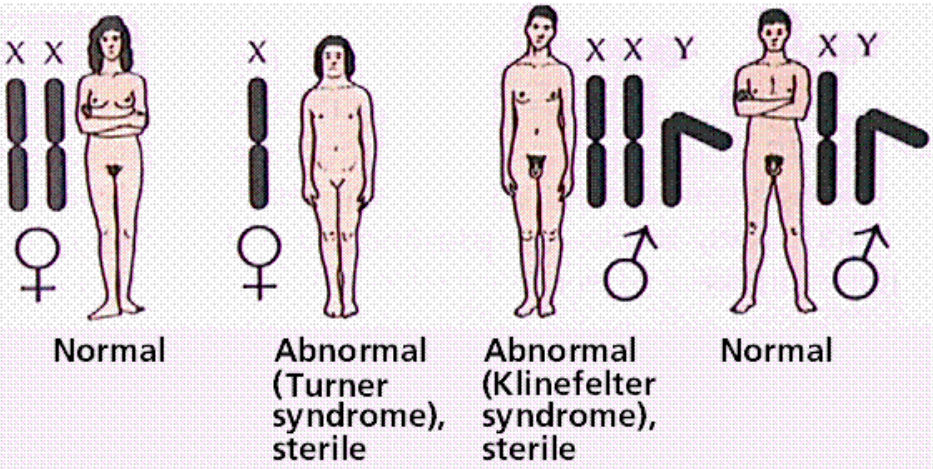


Fig. 3: Human sex chromosome abnormalities. Image from Purves *et al*., Life: The Science of Biology, 4th Edition, by Sinauer Associates and WH Freeman

Chromosome deletions may also be associated with other syndromes such as Wilm's tumor.

Prenatal detection of chromosomal abnormalities is accomplished chiefly by amniocentesis. A thin needle is inserted into the amniotic fluid surrounding the fetus (a term applied to an unborn baby after the first trimester). Cells are withdrawn have been sloughed off by the fetus, yet they are still fetal cells and can be used to determine the state of the fetal chromosomes, such as Down's Syndrome and the sex of the baby after a karyotype has been made.

6.3. Human Allelic Disorders (Recessive)

The first Mendelian trait in humans was described in 1905 (brachydactly) by Dr. Farabee. Now more than 3500 human genetic traits are known.

Albinism, the lack of pigmentation in skin, hair, and eyes, is also a Mendelian human trait. Homozygous recessive (aa) individuals make no pigments, and so have face, hair, and eyes that are white to yellow. For heterozygous parents with normal pigmentation (Aa), two different types of gametes may be produced: A or a. From such a cross 1/4 of the children could be albinos. The brown pigment melanin cannot be made by albinos. Several mutations may cause albinism: 1) the lack of one or another enzyme along the melanin-producing pathway; or 2) the inability of the enzyme to enter the pigment cells and convert the amino acid tyrosine into melanin.

Phenylketonuria (PKU) is recessively inherited disorder whose sufferers lack the ability to synthesize an enzyme to convert the amino acid phenylalanine into tyrosine Individuals homozygous recessive for this allele have a buildup of phenylalanine and abnormal breakdown products in the urine and blood. The breakdown products can be harmful to developing nervous systems and lead to mental retardation. 1 in 15,000 infants suffers from this problem. PKU homozygotes are now routinely tested for in most states. If you look closely at a product containing Nutra-sweet artificial sweetener, you will see a warning to PKU sufferers since phenylalanine is one of the amino acids in the sweetener. PKU sufferers are placed on a diet low in phenylalanine, enough for metabolic needs but not enough to cause the buildup of harmful intermediates.

Tay-Sachs Disease is an autosomal recessive resulting in degeneration of the nervous system. Symptoms manifest after birth. Children homozygous recessive for this allele rarely survive past five years of age. Sufferers lack the ability to make the enzyme N-acetyl-hexosaminidase, which breaks down the GM2 ganglioside lipid. This lipid accumulates in lysosomes in brain cells, eventually killing the brain cells. Although rare in the general population (1 in 300,000 births), it was (until recently) higher (1 in 3600 births) among Jews of eastern central European descent. One in 28 American Jews is thought to be a carrier, since 90% of the American Jewish population emigrated from those areas in Europe. Most Tay-Sachs babies born in the US are born to non-Jewish parents, who did not undergo testing programs that most US Jewish prospective parents had.

Sickle-cell anemia is an autosomal recessive. The recessive allele causes a single amino acid substitution in the beta chains of hemoglobin. When oxygen concentration is low, sickling of cells occurs. Heterozygotes make enough "good beta-chain hemoglobin" that they do not suffer as long as oxygen concentrations remain high, such as at sea-level.

6.4. Human Allelic Disorders (Dominant)

Autosomal dominants are rare, although they are (by definition) more commonly expressed. Achondroplastic dwarfism occurs, even though sufferers have reduced fertility.

Huntington's disease (also referred to as Woody Guthrie's disease, after the folk singer who died in the 1960s) is an autosomal dominant resulting in progressive destruction of brain cells. If a parent has the disease, 50% of the children will have it (unless that parent was homozygous dominant, in which case all children would have the disease). The disease usually does not manifest until after age 30, although some instances of early onset phenomenon are reported among individuals in their twenties.

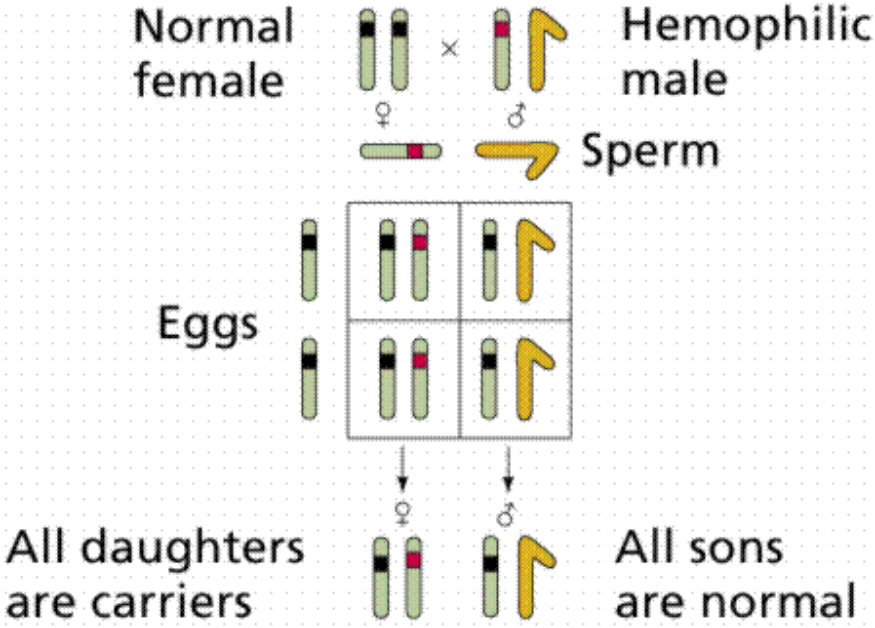
Polydactly is the presence of a sixth digit. In modern times the extra finger has been cut off at birth and individuals do not know they carry this trait. The extra digit is rarely functional.

6.5. Sex-linked Traits

Color blindness afflicts 8% of males and 0.04 % of human females. Color perception depends on three genes, each producing chemicals sensitive to different parts of the visible light spectrum. Red and green detecting genes are on the X-chromosome, while the blue detection is on an autosome.

Heamophilia is a group of diseases in which blood does not clot normally. Factors in blood are involved in clotting. Hemophiliacs lacking the normal Factor VIII are said to have Hemophilia A, the most common form. Normal Factor VIII can be supplied at a

high cost and health risk cost, although the development of biotechnologically engineered Factor VIII produced by bacteria lessens the health risk. England's Queen Victoria was a carrier for this disease. The allele was passed to two of her daughters and one son. Since royal families in Europe commonly intermarried, the allele spread, and may have contributed to the downfall of the Russian monarchy (Czar Nicholas' son Alexei suffered from hemophilia A inherited from his mother who carried Victoria's genetic secret).



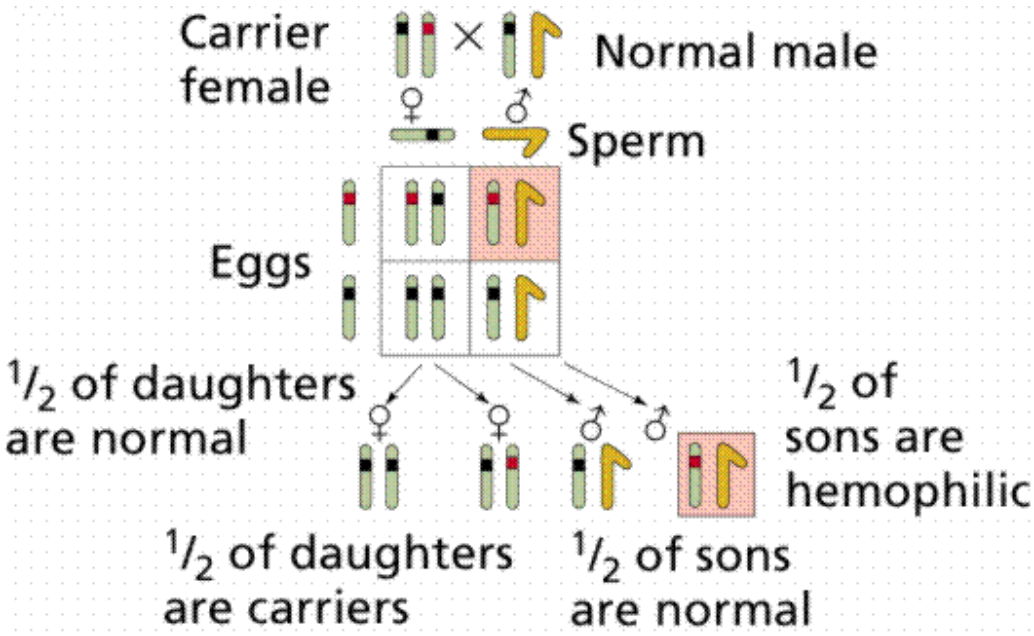


Fig. 4: Inheritance of a human sex-linked trait. Images from Purves *et al*., Life: The Science of Biology, 4th Edition, by Sinauer Associates and WH Freeman.

6.6. Muscular dystrophy is a term encompassing a variety of muscle wasting diseases. The most common type, Duchenne Muscular Dystrophy (DMD), affects cardiac and skeletal muscle, as well as some mental functions. DMD is an X-linked recessive occurring in 1 in 3500 newborns. Most sufferers die before their 20th birthday. In 1987, Louis Kunkel claimed to have isolated a protein, dystrophin, present in normal individuals (about 0.002 % of their muscle protein) but absent in two individuals with DMD. The lack of dystrophin is accompanied with a condition of muscle hardening known as fibrosis, which restricts blood supply to the muscle which then die. The gene technologies discussed in an earlier chapter have been employed to sequence and clone the dystrophin gene, which is the largest known human gene (some 2-3 million base pairs), with 60 exons and many large introns.

6.7. Diagnosis of Human Genetic Diseases

Restriction enzymes, such as *Hpa* I were used in a study on sickle-cell anemia. The probe hybridized in normal hemoglobin with two fragments 7000 or 7600 nucleotides

long. Sickle-cell hemoglobin had hybridization with a 13,000 nucleotide single sequence. A similar result has been obtained from amniocentesis studies, providing a tool to screen fetus and adult for sickle-cell. The markers where hybridization occurred are referred to as RFLPs (restriction-fragment-length polymorphisms). The longer fragment in sickle-cell individuals is interpreted as evidence of a mutation in the recognition sequence. Two nucleotide sequences close together on the same DNA molecule tend to stay together. In the sickle-cell DNA the beta-chain hemoglobin gene has become linked with another gene that somehow alters the recognition sequence at which *Hpa* I hybridizes. Heterozygotes will have both long and short fragments, while a single type (short or long) will occur in homozygous dominant and recessive, respectively.

Huntington's disease was studied by James F. Gusella and his research team, who used RFLPs to identify a marker. Testing a large library of human DNA fragments, Gusella et al. found the needle in the haystack. The enzyme used was *Hind* III. Four fragments have been identified in an American family that has members suffering from the disease. The presence of fragment A has been identified in individuals who suffer from (or will suffer from) Huntington's. Pattern A occurs in 60 percent of the population, as well as the Huntington's sufferers. A Venezuelan family of 3000 members is descended from a German sailor who had Huntington's. This family had a strong correlation between Fragment C and the disease. Pattern C is much less common among the general population in this country. Many individuals do not wish to know if they will develop this disease; Woody Guthrie's children have chosen not to be tested.

Cystic fibrosis (CF) has also been studied with RFLP technology. CF is the most common genetic disease in Caucasians.

6.8. Radioactive probes

Hemophiliacs suffer from defective Factor VIII, which can be detected in fetuses 20 weeks old. A more accurate test, which can also be administered earlier during pregnancy, involves the use of a radioactive probe (36 nucleotide RNA fragment) which

hybridizes restriction fragments. The gene for hemophilia is 186,000 base pairs, and has 26 exons separated by 25 introns. Mutations in the gene can be detected by RFLPs. This technology has also been used to detect the single base-pair difference between normal and mutated beta-chains, a screen for sickle-cell anemia. A DNA probe has been developed that hybridizes with the gene for dystrophin. The previous screening for Duchenne Muscular Dystrophy was a sex screen, with option to abort a male. The new technique allows differentiation between the healthy and diseased male fetus, so parents have more information with which to make an informed choice (if they chose). The hybridization only occurs if the normal dystrophin gene is present, no hybridization occurs in the DMD sufferer.

7.0. PEDIGREE ANALYSIS

7.1. Introduction

Humans are unique among organisms in many ways, but one way which is near and dear to a geneticist's heart is that humans are not susceptible to genetic experimentation. In practice, we humans actually share this characteristic with many long lived organisms who delay first births. In short, it is not terribly convenient to perform experimental crosses if one has to wait 15 years between generations. However, for humans, one also has to add that our system of morality uniquely does not allow such experimentation on humans. This is an unfortunate state of affairs since there is no other organism for which practical knowledge of their genetics would be more useful, especially in the case of the genetics of heritable diseases.

In a perfect example of necessity being the mother of invention, however, it has been found that human genetics may readily be inferred so long as good records have been kept within large families. This formal mechanism of inference is termed *pedigree analysis*. In this lecture we will discuss many aspects of human genetics, in particular considering strategies of *pedigree analysis* whereby we will attempt to infer the genetics of human conditions based on knowledge of marriage (mating) and affliction in large extended families.

7.2. Autosome

An *autosome* is a eucaryote chromosomes other than a sex chromosomes.

Sex chromosome [X chromosome, Y chromosome]

a. Eucaryote chromosomes of many species (of, particularly, animals) may be distinguished into two general categories: Those that are involved in the determination of gender (*sex chromosomes*) and those that are not (autosomes).
b. In humans, the female gender is determined by the absence of a *Y chromosome* and the presence of an *X chromosome* while the male is determined by the presence of a *Y chromosome*. Since females have only *X chromosomes*, the only combination of chromosomes that normally occurs is *XX* (i.e., female) and *XY* (i.e., male).
c. Note that a human lacking an *X chromosome* is also inviable since the *X chromosome* carries a number of valuable genes (i.e., just YY and just Y are not viable). The same cannot be said for the *Y chromosome* which is of value only in the determination of gender.
d. Note that not all species employ chrosomal (i.e., genetic) differences to determine gender. In fact, there are a whole range of strategies by which gender is determined in the natural world, and it is not unheard of for organisms to switch gender during normal, adult maturation.

7.2. Chromosome number

Humans possess 23 chromosomes per haploid. Of these 23, 22 are autosomes while the 23rd is either one of two sex chromosomes, X or Y.

The chromosome number found in other species usually differs from that in humans. Chromosome number, however, tends to be invariant within species (typically by definition).

7.3. Karyotype

Metaphase mitosis chromosomes may be arranged so that they are visible through a light microscope.

In such an arrangement, called a *karyotype*, chromosomes display a variety of differentiating characteristics including length from centromere (arm length), overall length, and patterns of banding (the latter upon staining with particular chromosome staining dyes).

From this *karyotype* chromosomes are named numerically, for the most part with increasing number with decreasing size (i.e., length). Thus, autosome number 1 is the largest of human chromosomes while autosome number 21 is the smallest (note that autosome number 22 is not actually the smallest of human chromosomes---a violation of convention that apparently represents an early cytogenetical error that was caught only after everyone was used to this erroneous designation, therefore doomed to be forever an exception to the rule).

7.4. Homologous [nonhomologous] chromosome

A chromosome of the same *type* found in an individual. That is, the *homologue* of one chromosome 21 in a human individual's cell is the other chromosome 21. A *nonhomologous* chromosome (to chromosome 21) would be chromosome numbers 1-20, 22, X, or Y.

7.5. Nondisjunction

A failure of chromosomes to separate during meiosis or mitosis. In meiosis this leads to gametes which contain a surplus or deficit of one or more chromosome.

Trisomies and monosomies

A consequence of a meiotic nondisjunction.A *trisomy* is the presence of three or more homologous chromosomes upon fertilization (and thus in the resulting individual). A *monosomy* has only a single example of a particularly type of chromosome rather than the expected two (i.e., in a diploid individual). Note that the vast majority of human autosomal *trisomies* and *monosomies* are inviable, most prior to birth.

Trisomy 21 [Down syndrome]

The only long term viable human trisomy. *Down syndrome* exhibiting individuals have, in the simplest of cases, three autosome number 21s rather than the normal two.

Translocation

The breakage and joining of a section of one chromosome to a second, nonhomologous chromosome. *Translocations* are a means by which partial trisomies may be attained (i.e., an individual may hold three copies of a particular section of a given chromosome with the third copy attached to a nonhomologous chromosome).

7.6. Dosage compensation

By definition, most males contain a deficit of one X chromosome. This represents a sex genetic material monosomy. However, this situation is, obviously, not lethal. Why not? The answer is that for the X chromosome, it is the presence of two active copies rather than the presence of only one which represents the pathological situation. Thus, males are for the most part not detrimentally affected by their monosomy X. But wait a minute, females have two X chromosomes. Isn't that an obvious contradiction of the above statements? The answer, of course, is, yes it is. But one would suppose that such a situation would represent strong selective pressure to do something about the problem, and that something is the inactivation of one of the two X chromosomes located in each female cell. Thus, females have effectively one X chromosome per cell. This strategy is termed *dosage compensation* (that is, the dose of genes and gene products per X

chromosome is the same for both males and females because female are rendered with effectively the same number of X chromosome as males).

7.7, Barr body

Inactivated X chromosomes found in the cells of females are called *Barr bodies*. They were named this before it was understood that *Barr bodies* are inactivated X chromosomes.

7.8. Mosaicism [mosaic]

X chromosomes are inactivated to form Barr bodies both at random and, effectively, irreversibly. Thus, at some point to during the development of the soma, one of the two X chromosomes in each cell then present is inactivated. Since the information found on each of the two X chromosomes is not necessary identical, Barr body formation leads to individual cells in a females body effectively being genetically of only one of two types, depending on which X chromosome was inactivated. This situation is known as *mosaicism*. That is, a female's body consists of a mixture of cells each having had different X chromosomes inactivated. Note that a *mosaic* can arise by many other mechanisms involving the change in genetic information (e.g., nondisjunction during mitosis) in the course of soma development.

7.9. Sex chromosome trisomies and monosomies

A great range of variation in sex chromosome number exists among viable individuals, a situation which contrasts starkly with the general inviability of autosomal trisomies and monosomies. Such individuals as XXX females (or higher such as XXXX) tend toward sterility but are otherwise relatively normal. XO females (only one sex chromosome), on the other hand, are seriously affected. XXY males (and higher such as XXXY) also tend to develop normally but nevertheless suffer from sterility. XYY males (and higher such as XYYY), on the other hand, tend to be both relatively normal and fertile.

7.10. Pedigree

Pedigrees are a convention for keeping track of human genetic traits used to infer genotype. *Pedigrees* are the human equivalent of test crosses. In a visualization of a *pedigree*:

i. males are designated with square symbols.
ii. females with round symbols
iii. lines are drawn to indicated matings, parent-offspring relationships, and relationships between siblings.

Traits associated with dominant, recessive, sex linked, etc. alleles and loci display characteristic patterns in *pedigrees* just as they do when following traits in any organism by any means (i.e., in addition to historical).

Autosomal dominant allele [e.g., Huntington's Disease, brown eyes]

A phenotype associated with an autosomal dominant allele will, ideally, be present in every individual carrying that allele. In principle, at least, there are no silent carriers of dominant alleles. (This is because for many alleles, their interaction with genes found at different loci and with the environment is sufficient to make the actual expression of an allele variable---genetics can be very complicated.). For particularly serious early onset diseases caused by dominant alleles, it is likely that neither parent possesses a copy of the offending allele thus implying that allele formed through mutation during or after gamete formation. Such diseases are consequently very rare.

Autosomal recessive alleles [silent carriers] include: albinism, cystic fibrosis, certain types of hemophilia, Tay-Sachs disease, PKU, blue eyes.

A pedigree following a trait associated with an autosomal recessive allele is often marked by a skipping of generations. That is, children may express a trait which their parents do not. In such a situation, both parents are heterozygotes, also known as *silent*

carriers. Thus, as in any typical cross between two heterozygotes, one-fourth of all offspring are expected to be homozygous for one of the alleles (i.e.Mendelian ratio).

Note that alleles that cause serious, early onset disease are far more likely to be caused by *autosomal recessives* than autosomal dominants, sex-linked dominants, sex-linked recessives, etc. since the latter each have much higher probabilities of being expressed (and thus selected against) than the former.

For less serious conditions, there is nothing which precludes one or both parents from being homozygotes: In the case of one parent being a homozygote recessive and the other a homozygote dominant, no offspring will be affected. In the case of one parent being a homozygote recessive and the other a heterozygote, half of the offspring, on average, will be affected. In the case of one both parents being homozyote recessives, all offspring will be affected.

7.11. Sex-linked dominant alleles

Males are haploid for X chromosome:

Since males carry only a single X chromosome, they are effectively haploid for that chromosome. Thus, there is no such thing as recessive or dominant allele when it comes to the genes found on a male's X chromosome.

Haploid phenotype reflect the presence of whatever allele is present. However, since females are diploid for the X chromosome, a given allele found on the X chromosome may be considered to be recessive or dominant when found in females.

A *sex linked* dominant allele has a variation on the pattern displayed by autosomal dominant alleles. That is: one-half of the offspring of an afflicted heterozygote female will be similarly afflicted (gender independent).

Only the female progeny of males will be afflicted (because the male donates an X chromosome to his female progeny).

As with any *sex-linked allele*, males can pass the allele only on to their daughters, not their sons.

7.12. Alleles found on Y chromosome

Though rare, a trait associated with a loci found on the Y chromosome would be passed from father to son, only, and not skip generations. That is, it would always be expressed in the haploid state and never found in females.

7.13. Sex-linked recessive alleles

Sex-linked recessives show interesting patterns that result in part from genetics and in part from allele frequencies. That is, like all recessive traits, the likelihood of finding a homozygous recessive at any given loci (given random mating) is a function of the frequency of that allele in the general population (likelihood equals frequency squared).

For a rare allele, the likelihood of finding the allele in a heterozygous state is much greater (equal simply to the frequency of the allele) than in a homozygous individual.

Thus, for relatively rare recessive alleles found on X chromosomes, the likelihood that a women will carry the allele is much higher than the likelihood that the women will display the phenotype associated with that allele. Half of their sons, however, will receive an X chromosome carrying that recessive allele.

Since the sons are haploid for the X chromosome, they display that associated trait so long as they carry one copy of the allele.

Thus, for *sex-linked recessive alleles* the likelihood that a male will display the associated trait is equal to the frequency of the allele in the general population while the probability that any given female will carry it is equal to frequency of the allele squared.

As with any *sex-linked allele*, males can pass the allele only on to their daughters, not their sons. Examples include red-green color blindness, certain types of hemophilia.

7.14. Codominance

Codominance is neither recessive nor dominant. Certain allelic combinations exert neither dominance nor recessiveness. Instead, both alleles exert an influence on phenotype. Such allelic combinations are said to display *codominance*. *Codominance* both helps and hinders pedigree analysis.

It's a hindrance because it implies that at least three phenotypes are possible for three genotypes (AA, Aa, and aa). However, this also can make pedigree analysis terribly easy because phenotype maps 1:1 onto genotype.

7.15. ABO blood group

A slightly complicated example of a system in which there exists both multiple alleles, recessive alleles, dominant alleles, and codominance. Particularly, red blood cell surface markers (a.k.a., antigens) come in three varieties (actually more, but this is sufficient for both our discussion here and to describe most individuals), variety *A*, variety *B*, and variety *O*. Permissible genotypes include *AA*, *AB*, *AO*, *BB*, *BO*, and *OO*. Both *A* and *B* display dominance toward *O*, and codominance toward each other.

Thus, *AA* and *AO* individuals display phenotype A, *AB* individuals display phenotype *AB*, *BB* and *BO* individuals display phenotype *B*, and *OO* individuals display phenotype O.

7.16. Antibodies to blood antigens:

Individuals which do not have the *B* allele see the B antigen as foreign and consequently have antibodies which are reactive with the B antigen. Similarly, individuals who do not have an *A* allele have antibodies which are reactive to A antigen. The existence of these antibodies makes it inadvisable to receive blood (i.e., via a transfusions) from individuals whose red blood cell display the wrong (A or B) maker."The blood plasma of many people contains genetically determined antibodies referred to as agglutinins or isoantibodies. These are antibody a (anti-A), which attacks

agglutinogen A (i.e., the A protein), and antibody b (anti-B) which attacks B. The antibodies formed by each of the four blood types are shown in Figure 19-4. You do not have antibodies that attack the agglutinogens of your own erythrocytes. A person with blood type A does not have antibody a. But you do have an antibody against any agglutinogen you yourself do not synthesize. Suppose type A blood is accidentally given to a person who does not have A agglutinogens. The person's body recognizes that the A protein is foreign and therefore treats it as an antigen. Antibody a's rush to the foreign erythrocytes, attack them, and cause them to *agglutinate* (clump)--hence the names agglutinogen and agglutinins. This reaction is another example of an antigen-antibody response. In practice, only matching blood types are used for transfusions." Consequently, the *universal donor* is individuals whose red blood cell display no antigen (*OO*). On the other hand, the *universal recipient* is individuals whose red blood cell display both antigens (*AB*).

7.17. Additional donor-recipient permutations

AO and *AA* individuals can receive blood from *AA*, *AO*, and *OO* individuals (but not *BB* or *BO* individuals). *BO* and *BB* individuals can receive blood from *BB*, *BO*, and *OO* individuals (but not *AA* or *AO* individuals). Note, however, that "in practice, only matching blood types are used for transfusions.

7.18. Rh blood group

A simple one locus, two allele, dominant-recessive red blood cell marker. Rh^+ gives the marker-present phenotype while Rh^- is the marker-absent phenotype.The *rH* is an abbreviation of rhesus (as in rhesus macaque), the animal in which *rH* factors were first described.

"The *Rh system* is so named because it was first worked out in the blood of the Rhesus monkey. Like the ABO grouping, the Rh system is based on agglutinogens that lie on the surface of erythrocytes. Individuals whose erythrocytes have the Rh agglutinogens are designated Rh^+. Those who lack Rh agglutinogens are designated Rh^-. . . Under

normal circumstances, human plasma does not contain anti-Rh antibodies. However, if an Rh^- person receives Rh^+ blood, the body starts to make anti-Rh antibodies that will remain in the blood. If a second transfusion of Rh^+ blood is given later, the previously formed anti-Rh antibodies will react against the donated blood and a severe reaction may occur. One of the most common problems with Rh incompatibility arises from preganancy. During pregnancy, some of the fetus's blood may leak from the placenta (afterbirth) into the mother's blood stream. If the fetus is Rh^+ and the woman is Rh^-, she, upon exposure tot he Rh^+ fetal cells, will make anti-Rh antibodies. If the woman becomes pregnant again, her anti-Rh antibodies will cross the placenta and make their way into the blood stream of the baby. If the fetus is Rh^-, no problem will occur, since Rh^- blood does not have the Rh antigen. If the fetus is Rh^+, and antigen-antibody response called *hemolysis* may occur in the fetal blood. Hemolysis means a breakage of erythrocytes resulting in the liberation of hemoglobin. The hemolysis brought on by fetal-maternal incompatibility is called erythroblastosis fetalis or hemolytic disease of newborn. When a baby is born with erythroblastosis, all the blood is slowly removed and replaced with antibody-free blood. It is even possible to transfuse blood into the unborn child if erythroblastosis is suspected. More important, though, is the fact that erythroblastosis can be prevented with an injection of an anti-Rh $gamma_2$-globulin antibody preparation, administered to Rh^- mothers right after delivery or abortion. These antibodies tie up the fetal agglutinogens by producing antibodies. Thus the fetus of the next pregnancy is protected. In the case of an Rh^+ mother and an Rh^- child, there are no complications, since the fetus cannot make antibodies.

7.19. Sickle cell anemia

On the surface *sickle cell anemia* appears to be caused by a simple autosomal recessive allele. Thus, those homozygous for the *sickle cell* allele are stricken with *sickle cell anemia* while heterozygotes and homozygous, dominant individuals are not. However, the number of those stricken with *sickle cell anemia* is actually greater than would be expected given the severity of this disease. Consequently, there exists the possibility that something more is going on than first meets the eye and that that

something might be of interest. As it turns out, what is going on is that the *sickle cell* heterozygote, though otherwise pretty much normal, is simultaneously more resistant to malaria than homozygous, dominant individuals. In regions in which malaria causes significant disease and mortality, an advantage is accrued by the heterozygote even after factoring in the likelihood that some fraction of that individual's progeny will be afflicted by *sickle cell anemia*. *Sickle cell anemia* consequently is an example of a case of overdominance maintaining a polymorphism in populations living in malaria-prone areas (that is, the greater evolutionary fitness of the heterozygote maintains the frequency of the deleterious allele at artificially high levels).

7.20. Inviable

Dead or incapable of surviving due, usually, to a genetic defect. For example, most trisomies create sufficient problems with development that the fetus is rendered *inviable*, i.e., incapable of surviving outside of the womb and, in many cases, inside the womb as well.

Self Assessment Assignments

1. A "disease" which usually is a consequence of nondisjunction is?
2. A carrier is an individual who "carries" a certain allele. A carrier may express the corresponding trait or may not. The latter is called a "silent carrier." The following pedigree follows a trait that is very rare in most human populations. Assuming no marriage between related individuals, circle all of the persons you are certain are carriers.
3. Describe three characteristics of a normal human karyotype.
4. Name at least one way in which a species might do without sex chromosomes? (give one answer)
5. Give me an example of a phenotypically normal human who does not have 46 chromosomes (ignore fertility issues; use whatever method you consider appropriate to effectively describe/identify such an individual).

6. A couple is unable to conceive a child. Among other tests, both individuals are karyotyped to rule out sex chromosome number abnormalities. After superficial examination, both individuals are deemed karyotypically normal (i.e., they both have 46 chromosomes, the female with two X chromosomes, the male with one X and one Y). Can we rule out Klinefelter's syndrome? Why or why not?
7. Turner syndrome individuals are often viable more because they are mosaics rather than because of an inherent robustness associated with the underlying karyotypic abnormality. Propose a mechanism whereby such a mosaic might form (hint: start with a karyotypically normal zygote).
8. How are the following concepts related: (i) dosage compensation and (ii) the lack of a severe phenotypic consequence associated seen with some trisomies?
9. Given the ABO blood system, what is the genotype of a universal donor and a universal recipient?

Tutor Marked Assignment

Can you figure out your own *ABO* genotype?

8.0. FURTHER CONSIDERATION OF VARIOUS DEVIATIONS FROM BASIC PRINCIPLES

8.1. Polymorphism

Polymorphism in biology occurs when two or more clearly different phenotypes exist in the same population of a species, in other words, the occurrence of more than one *form* or *morph*. In order to be classified as such, morphs must occupy the same habitat at the same time and belong to a panmictic population or a population that is randomly mating.

Polymorphism is common in nature; it is related to biodiversity, genetic variation and adaptation; it usually functions to retain variety of form in a population living in a varied environment. The most common example is sexual dimorphism, which occurs in many organisms. Other examples are mimetic forms of butterflies and human haemoglobin and blood types.

Polymorphism results from evolutionary processes, as does any aspect of a species. It is heritable, and is modified by natural selection. In polyphenism, an individual's genetic make-up allows for different morphs, and the switch mechanism that determines which morph is shown is environmental. In *genetic polymorphism* the genetic make-up determines the morph. Ants exhibit both types in a single population.

Polymorphism as described here involves morphs of the phenotype. The term is also used somewhat differently by molecular biologists to describe certain point mutations in the genotype, such as SNPs (single nucleotide proteins).

In general use, polymorphism is quite a broad term, in biology it has been given a specific meaning.

The term omits characters showing *continuous variation* (such as weight), even though this has a heritable component. Polymorphism deals with forms in which the variation is discrete (discontinuous) or strongly bimodal or polymodal.

Morphs must occupy the same habitat at the same time: this excludes geographical races and seasonal forms. The use of the words *morph* or *polymorphism* for what is a visibly different *geographical race or variant* is common, but incorrect. The significance of geographical variation is in that it may lead to allopatric speciation, whereas true polymorphism takes place in panmictic populations.

The term was first used to describe *visible forms*, but nowadays it has been extended to include *cryptic morphs*, for instance blood types, which can be revealed by a test.

Rare variations are not classified as polymorphisms; and mutations by themselves do not constitute polymorphisms. To qualify as a polymorphism there has to be some kind of balance between morphs underpinned by inheritance. The criterion is that the frequency of the *least* common morph is too high simply to be the result of new mutations or, as a rough guide, that it is greater than 1 percent (though that is far higher than any normal mutation rate for a single allele).

Various synonymous terms exist for the various polymorphic forms of an organism. The most common are *morph* and *morpha*, while a more formal term is *morphotype*. *Form* and *phase* are sometimes also used, but are easily confused in zoology with, respectively, "form" in a population of animals, and "phase" as a color or other change in an organism due to environmental conditions (temperature, humidity, etc.). Phenotypic traits and characteristics are also possible descriptions, though that would imply just a limited aspect of the body.

In the taxonomic nomenclature of zoology, the word "morpha" plus a Latin name for the morph can be added to a binomial or trinomial name. However, this invites confusion with geographically-variant ring species or subspecies, especially if polytypic. Morphs have no formal standing in the ICZN. In botanical taxonomy, the concept of morphs is represented with the terms "variety", "subvariety" and "form", which are formally regulated by the ICBN. Horticulturalists sometimes confuse this usage of "variety" both with cultivar ("variety" in viticultural usage, rice agriculture jargon, and informal gardening lingo) and with the legal concept "plant variety" (protection of a cultivar as a form of intellectual property).

Selection, whether natural or artificial, changes the frequency of morphs within a population; this occurs when morphs reproduce with different degrees of success. A genetic (or *balanced*) polymorphism usually persists over many generations, maintained by two or more opposed and powerful selection pressures. Apes have similar blood groups to humans; this suggests rather strongly that this kind of polymorphism is quite ancient, at least as far back as the last common ancestor of the apes and man, and possibly even further.

The relative proportions of the morphs may vary; the actual values are determined by the effective fitness of the morphs at a particular time and place. The mechanism of heterozygote advantage assures the population of some alternative alleles at the locus or loci involved. Only if competing selection disappears will an allele disappear. However, heterozygote advantage is not the only way a polymorphism can be

maintained. Apostatic selection, whereby a predator consumes a common morph whilst overlooking rarer morphs is possible and does occur. This would tend to preserve rarer morphs from extinction.

A polymorphic population does not initiate speciation; nor does it prevent speciation. It has little or nothing to do with species splitting. However, *it has a lot to do with the adaptation of a species to its environment*, which may vary in colour, food supply, predation and in many other ways. Polymorphism is one good way the opportunities get to be used; it has survival value, and the selection of modifier genes may reinforce the polymorphism.

Investigation of polymorphism requires a coming together of field and laboratory technique. In the field:

- detailed survey of occurrence, habits and predation
- selection of an ecological area or areas, with well-defined boundaries
- capture, mark, release, recapture data
- relative numbers and distribution of morphs
- estimation of population sizes

And in the laboratory:

- genetic data from crosses
- population cages
- chromosome cytology if possible
- use of chromatography or similar techniques if morphs are cryptic (for example, biochemical)

Both types of work are equally important. Without proper field-work the significance of the polymorphism to the species is uncertain; without laboratory breeding the genetic basis is obscure. Even with insects the work may take many years; examples of Batesian mimicry noted in the nineteenth century are still being researched.

8.2. Genetic polymorphism

Since all polymorphism has a genetic basis, *genetic polymorphism* has a particular meaning. Genetic polymorphism is the simultaneous occurrence in the same locality of two or more discontinuous forms in such proportions that the rarest of them cannot be maintained just by recurrent mutation. The definition has three parts: a) sympatry: one interbreeding population; b) discrete forms; and c) not maintained just by mutation.

Genetic polymorphism is actively and steadily maintained in populations by natural selection, in contrast to *transient polymorphisms* where a form is progressively replaced by another. By definition, genetic polymorphism relates to a balance or equilibrium between morphs. The mechanisms that conserve it are types of balancing selection.

8.3. Mechanisms of balancing selection

1. Heterosis (or heterozygote advantage) where the heterozygote at a locus is fitter than either homozygote".
2. Frequency dependent selection: The fitness of a particular phenotype is dependent on its frequency relative to other phenotypes in a given population. Example: prey switching, where rare morphs of prey are actually fitter due to predators concentrating on the more frequent morphs.
3. Fitness varies in time and space. Fitness of a genotype may vary greatly between larval and adult stages, or between parts of a habitat range.
4. Selection acts differently at different levels. The fitness of a genotype may depend on the fitness of other genotypes in the population: this covers many natural situations where the best thing to do (from the point of view of survival and reproduction) depends on what other members of the population are doing at the time.

8.4. Pleiotropism

Most genes have more than one effect on the phenotype of an organism (pleiotropism). Some of these effects may be visible, and others cryptic, so it is often important to look beyond the most obvious effects of a gene to identify other effects. Cases occur where a gene affects an unimportant visible character, yet a change in fitness is recorded. In such cases the gene's other (cryptic or 'physiological') effects may be responsible for the change in fitness.

If a neutral trait is pleiotropically linked to an advantageous one, it may emerge because of a process of natural selection. It was selected but this doesn't mean it is an adaptation. The reason is that, although it was selected, there was no selection for that trait."

8.5. Epistasis

Epistasis occurs when the expression of one gene is modified by another gene. For example, gene A only shows its effect when allele B1 (at another locus) is present, but not if it is absent. This is one of the ways in which two or more genes may combine to produce a coordinated change in more than one characteristic (for instance, in mimicry). Unlike the supergene, epistatic genes do not need to be closely linked or even on the same chromosome.

Both pleiotropism and epistasis show that a gene need not relate to a character in the simple manner that was once supposed.

8.6. The origin of supergenes

Although a polymorphism can be controlled by alleles at a single locus (e.g. human ABO blood groups), the more complex forms are controlled by supergenes consisting of several tightly linked genes on a single chromosome. Batesian mimicry in butterflies and heterostyly in angiosperms are good examples. There is a long-standing debate as to how this situation could have arisen, and the question is not yet resolved.

Whereas a gene family (several tighly linked genes performing similar or identical functions) arises by duplication of a single original gene, this is usually not the case with supergenes. In a supergene some of the constituent genes have quite distinct functions, so they must have come together under selection. This process might involve suppression of crossing-over, translocation of chromosome fragments and possibly occasional cistron duplication. That crossing-over can be suppressed by selection has been known for many years.

Debate has centred round the question of whether the component genes in a super-gene could have started off on separate chromosomes, with subsequent reorganization, or if it is necessary for them to start on the same chromosome. Originally, it was held that chromosome rearrangement would play an important role. This explanation was accepted by E. B. Ford and incorporated into his accounts of ecological genetics.

However, today many believe it more likely that the genes start on the same chromosome. They argue that supergenes arose *in situ*. This is known as Turner's sieve hypothesis.

8.7. Examples of polymorphism

8.7.1. Sexual dimorphism

We meet genetic polymorphism daily, since our species (like most other eukaryotes) uses sexual reproduction, and of course, the sexes are differentiated. However, even if the sexes were identical in superficial appearance, the division into two sexes is a dimorphism, albeit cryptic. This is because the phenotype of an organism includes its sexual organs and its chromosomes, and all the behaviour associated with reproduction. So research into sexual dimorphism has addressed two issues: first, the advantage of sex in evolutionary terms; second, the role of visible sexual differentiation.

The system is relatively stable (with about half of the population of each sex) and heritable, usually by means of sex chromosomes. Every aspect of this everyday

phenomenon bristles with questions for the theoretical biologist. Why is the ratio ~50/50? How could the evolution of sex occur from an original situation of asexual reproduction, which has the advantage that every member of a species could reproduce? Why the visible differences between the sexes? These questions have engaged the attentions of biologists such as Charles Darwin, August Weismann, Ronald Fisher, George C. Williams, John Maynard Smith and W. D. Hamilton, with varied success.

Of the many issues involved, there is widespread agreement on the following: the advantage of sexual and hermaphroditic reproduction over asexual reproduction lies in the way recombination increases the genetic diversity of the ensuing population. The advantage of sexual reproduction over hermaphroditic is not so clear. In forms that have two separate sexes, same sex combinations are excluded from mating which decreases the amount of diversity compared with hermaphrodites by at least twice. So, why are almost all progressive species bi-sexual, considering the asexual process is more efficient and simple, whilst hermaphrodites produce a more diversified progeny? It has been suggested that differentiation into two sexes has evolutionary advantages allowing changes to concentrate in the male part of the population and at the same time preserving the existing genotype distribution in the females. This enables the population to better meet the challenges of infection, parasitism, predation and other hazards of the varied environment.

8.7.2. Human polymorphisms

Apart from sexual dimorphism, there are many other examples of human genetic polymorphisms. Infectious disease has been a major factor in human mortality, and so has affected the evolution of human populations. Evidence is now strong that many polymorphisms are maintained in human populations by balancing selection.

8.7.3. Human blood groups

All the common blood types, such as the ABO system, are genetic polymorphisms. Here we see a system where there are more than two morphs: the phenotypes are A, B, AB and O are present in all human populations, but vary in proportion in different parts of the world. The phenotypes are controlled by multiple alleles at one locus. These polymorphisms are seemingly never eliminated by natural selection; the reason came from a study of disease statistics.

Statistical research has shown that the various phenotypes are more, or less, likely to suffer a variety of diseases. For example, an individual's susceptibility to cholera (and other diarrheal infections) is correlated with their blood type: those with type O blood are the most susceptible, while those with type AB are the most resistant. Between these two extremes are the A and B blood types, with type A being more resistant than type B. This suggests that the pleiotropic effects of the genes set up opposing selective forces, thus maintaining a balance. Geographical distribution of blood groups (the differences in gene frequency between populations) is broadly consistent with the classification of "races" developed by early anthropologists on the basis of visible features.

8.7.4. Sickle-cell anaemia

Such a balance is seen more simply in sickle-cell anaemia, which is found mostly in tropical populations in Africa and India. An individual homozygous for the recessive sickle haemoglobin, HgbS, has a short expectancy of life, whereas the life expectancy of the standard haemoglobin (HgbA) homozygote and also the heterozygote is normal (though heterozygote individuals will suffer periodic problems). The sickle-cell variant survives in the population because the heterozygote is resistant to malaria and the malarial parasite kills a huge number of people each year. This is balancing selection or genetic polymorphism, balanced between fierce selection against homozygous sickle-cell sufferers, and selection against the standard HgbA homozygotes by malaria. The heterozygote has a permanent advantage (a higher fitness) so long as malaria exists; and it has existed as a human parasite for a long time. Because the heterozygote

survives, so does the HgbS allele survive at a rate much higher than the mutation rate in Sickle-cell disease.

8.7.5. Duffy system

The Duffy antigen is a protein located on the surface of red blood cells, encoded by the *FY (DARC)* gene. The protein encoded by this gene is a non-specific receptor for several chemokines, and is the known entry-point for the human malarial parasites *Plasmodium vivax* and *Plasmodium knowlesi*. Polymorphisms in this gene are the basis of the Duffy blood group system.

In humans, a mutant variant at a single site in the FY cis-regulatory region abolishes all expression of the gene in erythrocyte precursors. As a result, homozygous mutants are strongly protected from infection by *P. vivax*, and a lower level of protection is conferred on heterozygotes. The variant has apparently arisen twice in geographically distinct human populations, in Africa and Papua New Guinea. It has been driven to high frequencies on at least two haplotypic backgrounds within Africa. Recent work indicates a similar, but not identical, pattern exists in baboons (*Papio cynocephalus*), which suffer a mosquito-carried malaria-like pathogen, *Hepatocystis kochi*. Researchers interpret this as a case of convergent evolution.

8.7.6. G6PD

G6PD (Glucose-6-phosphate dehydrogenase) human polymorphism is also implicated in malarial resistance. G6PD alleles with reduced activity are maintained at a high level in endemic malarial regions, despite reduced general viability. Variant A (with 85% activity) reaches 40% in sub-Saharan Africa, but is generally <1% outside Africa and the Middle East.

8.7.7. Cystic fibrosis

Cystic fibrosis, a congenital defect which affects about one in 2000 children, is caused by a mutant form of the CF transmembrane regulator gene, CFTR. The transmission is

Mendelian: the normal gene is dominant, so all heterozygotes are healthy, but those who inherit two mutated genes have the condition. The mutated allele is present in about 1:25 of the population (mostly heterozygotes), which is much higher than expected from the rate of mutation alone. Sufferers from this disease have shortened life expectancy (and males are usually sterile if they survive), and the disease was effectively lethal in pre-modern societies. The incidence of the disease varies greatly between ethnic groups, but is highest in Caucasian populations.

Although over 1500 mutations are known in the CFTR gene, by far the most common mutant is DF508. This mutant is being kept at a high level in the population despite the lethal or near-lethal effects of the mutant homozygote. It seems that some kind of heterozygote advantage is operating. Early theories that the heterozygotes might enjoy increased fertility have not been borne out. Present indications are that the bacterium which causes typhoid fever enters cells using CFTR, and experiments with mice suggest that heterozygotes are resistant to the disease. If the same were true in humans, then heterozygotes would have had an advantage during typhoid epidemics. Cystic fibrosis is a prime target for gene therapy research.

8.7.8. Human taste morphisms

A famous puzzle in human genetics is the genetic ability to taste phenylthiourea (phenylthiocarbamide or PTC), a morphism which was discovered in 1931. This substance, which to some of us is bitter, and to others tasteless, is of no great significance in itself, yet it is a genetic dimorphism. Because of its high frequency (which varies in different ethnic groups) it must be connected to some function of selective value. The ability to taste PTC itself is correlated with the ability to taste other bitter substances, many of which are toxic. Indeed, PTC itself is toxic, though not at the level of tasting it on litmus paper. Variation in PTC perception may reflect variation in dietary preferences throughout human evolution, and might correlate with susceptibility to diet-related diseases in modern populations. There is a statistical correlation between PTC tasting and liability to thyroid disease.

Fisher, Ford and Huxley tested orangutans and chimpanzees for PTC perception with positive results, thus demonstrating the long-standing existence of this dimorphism. The recently identified PTC gene, which accounts for 85% of the tasting variance, has now been analysed for sequence variation with results which suggest selection is maintaining the morphism.

8.7.9. Lactose tolerance/intolerance

The ability to metabolize lactose, a sugar found in milk and other dairy products, is a prominent dimorphism that has been linked to recent human evolution.

8.7.10. MHC molecules

The genes of the major histocompatibility complex (MHC) are highly polymorphic, and this diversity plays a very important role in resistance to pathogens. This is true for other species as well.

8.7.11. The cuckoo

Over fifty species in this family of birds practise brood parasitism; the details are best seen in the British or European cuckoo (*Cuculus canorus*). The female lays 15–20 eggs in a season, but only one in each nest of another bird. She removes some or all of the host's clutch of eggs, and lays an egg which closely matches the host eggs. Although, in Britain, the hosts are always smaller than the cuckoo itself, the eggs she lays are small, and coloured to match the host clutch but thick-shelled. This latter is a defence which protects the egg if the host detects the fraud.

The intruded egg develops exceptionally quickly; when the newly-hatched cuckoo is only ten hours old, and still blind, it exhibits an urge to eject the other eggs or nestlings. It rolls them into a special depression on its back and heaves them out of the nest. The cuckoo nestling is apparently able to pressure the host adults for feeding by mimicking the cries of the host nestlings. The diversity of the cuckoo's eggs is extraordinary, the forms resembling those of its most usual hosts. In Britain these are:

1. Meadow pipit (*Anthus pratensis*): brown eggs speckled with darker brown.
2. European robin (*Erithacus rubecula*): whitish-grey eggs speckled with bright red.
3. Reed warbler (*Acrocephalus scirpensis*): light dull green eggs blotched with olive.
4. Redstart (*Phoenicurus phoenicurus*): clear blue eggs.
5. Hedge sparrow (*Prunella modularis*): clear blue eggs, unmarked, not mimicked. This bird is an uncritical fosterer; it tolerates in its nest eggs that do not resemble its own.

Each female cuckoo lays one type only; the same type laid by her mother. In this way female cuckoos are divided into groups (known as *gentes*, singular *gens*), each parasitises the host to which it is adapted. The male cuckoo has its own territory, and mates with females from any gens; thus the population (all gentes) is interbreeding.

The standard explanation of how the inheritance of gens works is as follows. The egg colour is inherited by sex chromosome. In birds sex determination is ZZ/ZW, and unlike mammals, the heterogametic sex is the female. The determining gene (or super-gene) for the inheritance of egg colour is believed to be carried on the W chromosome, which is directly transmitted in the female line. The female behaviour in choosing the host species is set by imprinting after birth, a common mechanism in bird behaviour.

Ecologically, the system of multiple hosts protects host species from a critical reduction in numbers, and maximises the egg-laying capacity of the population of cuckoos. It also extends the range of habitats where the cuckoo eggs may be raised successfully. Detailed work on the Cuckoo started with E. Chance in 1922, and continues to the present day; in particular, the inheritance of gens is still a live issue.

8.7.12. Grove snail

The grove snail, *Cepaea nemoralis*, is famous for the rich polymorphism of its shell. The system is controlled by a series of multiple alleles. The shell colour series is brown (genetically the top dominant trait), dark pink, light pink, very pale pink, dark yellow and light yellow (the bottom or universal recessive trait). Bands may be present or absent;

and if present from one to five in number. Unbanded is the top dominant trait, and the forms of banding are controlled by modifier genes. In many ecologies, the snail is regularly predated by the song thrush *Turdus philomelos*, which breaks them open on *thrush anvils* (large stones). Here fragments accumulate, permitting researchers to analyse the snails taken. The thrushes hunt by sight, and capture selectively those forms which match the habitat *least well*. Snail colonies are found in woodland, hedgerows and grassland, and the predation determines the proportion of phenotypes (morphs) found in each colony.

A second kind of selection also operates on the snail, whereby certain heterozygotes have a physiological advantage over the homozygotes. In addition, apostatic selection is likely, with the birds preferentially taking the most common morph. This is the 'search pattern' effect, where a predominantly visual predator persists in targeting the morph which gave a good result, even though other morphs are available.

Despite the predation, the polymorphism survives in almost all habitats, though the proportions of morphs varies considerably. The alleles controlling the polymorphism form a super-gene with linkage so close as to be nearly absolute. This control saves the population from a high proportion of undesirable recombinants, and it is hypothesised that selection has brought the loci concerned together.

To sum up, in this species predation by birds appears to be the main (but not the only) selective force driving the polymorphism. The snails live on heterogeneous backgrounds, and thrush are adept at detecting poor matches. The inheritance of physiological and cryptic diversity is preserved also by heterozygous advantage in the super-gene.

A similar system of genetic polymorphism occurs in the White-lipped Snail *Cepaea hortensis*, a close relative of the grove snail. In Iceland, where there are no song thrushes, a correlation has been established between temperature and colour forms. Banded and brown morphs reach higher temperatures than unbanded and yellow

snails. This may be the basis of the physiological selection found in both species of snail.

8.7.13. Two-spotted ladybird beetle

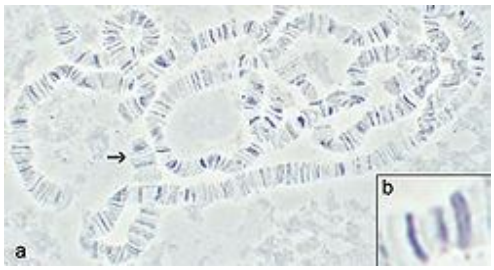
Adalia bipunctata, the two-spotted ladybird, is highly polymorphic. Its basic form is red with two black spots, but it has many other forms, the most important being melanic, with black elytra and red spots. The curious fact about this morphism is that, although the melanic forms are more common in industrial areas, its maintenance has nothing to do with cryptic camouflage and predation. The Coccinellidae as a whole are highly noxious, and experiments with birds and other predators have found this species quite exceptionally distasteful. Therefore, their colour is warning (aposematic) colouration, and all the morphs are quite conspicuous against green vegetation. The field studies identify differing proportions of morphs at different times of year and in different places, which indicates a high level of selection. However, the basis of that selection is still not known for sure, though many theories have been proposed. Since all the morphs are aposematically coloured, it seems unlikely that the difference between the colour of morphs is directly under selection. Perhaps pleiotropic effects of the genes acting on colour also affect the beetle's physiology, and hence its relative fitness. A similar polymorphic system is found in many other species in this family: *Harmonia axyridis* is a good example.

8.7.14. Ants

Ants exhibit a range of polymorphisms. First, there is their characteristic haplodiploid sex determination system, whereby all males are haploid, and all females diploid. Second, there is differentiation between both the females and males based mostly on feeding of larvae, which determines, for example, whether the imago is capable of reproduction. Lastly, there is differentiation of size and 'duties' (particularly of females), which are usually controlled by feeding and/or age, but which may sometimes be genetically controlled. Thus the order exhibits both genetic polymorphism and extensive polyphenism.

8.7.15. Chromosome polymorphism in *Drosophila*

In the 1930s Dobzhansky and his co-workers collected *Drosophila pseudoobscura* and *D. persimilis* from wild populations in California and neighbouring states. Using Painter's technique they studied the polytene chromosomes and discovered that the wild populations were polymorphic for chromosomal inversions. All the flies look alike whatever inversions they carry: this is an example of a cryptic polymorphism. Accordingly, Dobzhansky favoured the idea that the morphs became fixed in the population by means of Sewall Wright's drift. However, evidence rapidly accumulated to show that natural selection was responsible:



Drosophila polytene chromosome

1. Values for heterozygote inversions of the third chromosome were often much higher than they should be under the null assumption: if no advantage for any form the number of heterozygotes should conform to N_s (number in sample) = $p^2+2pq+q^2$ where 2pq is the number of heterozygotes

2. Using a method invented by l'Heretier and Teissier, Dobzhansky bred populations in *population cages*, which enabled feeding, breeding and sampling whilst preventing escape. This had the benefit of eliminating migration as a possible explanation of the results. Stocks containing inversions at a known initial frequency can be maintained in controlled conditions. It was found that the various chromosome types do not fluctuate at random, as they would if selectively neutral, but adjust to certain frequencies at which they become stabilised. With *D. persimilis* he found that the caged population followed the values expected on the Hardy-Weinberg equilibrium when conditions were optimal

(which disproved any idea of non-random mating), but with a restricted food supply heterozygotes had a distinct advantage.

3. Different proportions of chromosome morphs were found in different areas. There is, for example, a polymorph-ratio cline in *D. robusta* along an 18-mile (29 km) transect near Gatlinburg, TN passing from 1,000 feet (300 m) to 4,000 feet. Also, the same areas sampled at different times of year yielded significant differences in the proportions of forms. This indicates a regular cycle of changes which adjust the population to the seasonal conditions. For these results selection is by far the most likely explanation.

4. Lastly, morphs cannot be maintained at the high levels found simply by mutation, nor is drift a possible explanation when population numbers are high.

By the time Dobzhansky published the third edition of his book in 1951, he was persuaded that the chromosome morphs were being maintained in the population by the selective advantage of the heterozygotes, as with most polymorphisms. Later he made yet another interesting discovery. One of the inversions, known as PP, was quite rare up to 1946, but by 1958 its proportion had risen to 8%. Not only that, but the proportion was similar over an area of some 200,000 square miles (520,000 km^2) in California. This cannot have happened by migration of PP morphs from, say, Mexico (where the inversion is common) because the rate of dispersal (at less than 2 km/year) is of the wrong order. The change therefore reflected a change in prevailing selection whose basis was not yet known.

8.7.16.Darwin's finches

Whereas Darwin spent just five weeks in the Galápagos, and David Lack spent three months, Peter and Rosemary Grant and their colleagues have made research trips to the Galápagos for about thirty years, particularly studying Darwin's finches. Here we look briefly at the case of the large cactus finch *Geospiza conirostris* on Isla Genovesa (formerly Tower Island) which is formed from a shield volcano, and is home to a variety of birds. These birds, like all well-studied groups, show various kinds of morphism.

Males are dimorphic in song type: songs A and B are quite distinct. Also, males with song A have shorter bills than B males. This is also a clear difference. With these beaks males are able to feed differently on their favourite cactus, the prickly pear *Opuntia*. Those with long beaks are able to punch holes in the cactus fruit and eat the fleshy aril pulp which surrounds the seeds, whereas those with shorter beaks tear apart the cactus base and eat the pulp and any insect larvae and pupae (both groups eat flowers and buds). This dimorphism clearly maximises their feeding opportunities during the non-breeding season when food is scarce.

Territories of type A and type B males are random if not mated but alternate if mated: no two breeding males of the same song type shared a common boundary. This initially suggested the possibility of assortative mating by female choice. However, further work showed that "the choice of a male by a female is independent of any conditioning influence of her father's song type and there is no evidence of assortative mating by bill type... Hence there is no direct evidence of reproductive subdivision in the population". In 1999 Peter Grant agreed that "sympatric speciation is unlikely to occur".

If the population is panmixic, then *Geospiza conirostris* exhibits a balanced genetic polymorphism and not, as originally supposed, a case of nascent sympatric speciation. The selection maintaining the polymorphism maximises the species' niche by expanding its feeding opportunity. The genetics of this situation cannot be clarified in the absence of a detailed breeding program, but two loci with linkage disequilibrium is a possibility.

Another interesting dimorphism is for the bills of young finches, which are either "pink" or "yellow". All species of Darwin's finches exhibit this morphism, which lasts for two months. No interpretation of this phenomenon is known.

8.8. **Genetic variation**

We are all unique because each of us has a unique combination of genetic variants, including SNPs in our DNA. A SNP or “snip” can affect our inherited risk of disease. A SNP, pronounced “snip”, is a single variation in the nucleotide sequence of DNA and stands for Single Nucleotide Polymorphism that can affect our inherited risk and a

multitude of other characteristics. Most of our features, internal and external are determined or influenced in some way by such variations in our DNA, i.e. by SNPs. Each of us is different because we carry a unique combination of genetic variants, including SNPs.

SNPs are the result of alterations to DNA, usually called mutations. These mutations accumulate very gradually as DNA is passed on from parent to child, generation after generation. The slow rate of change to our DNA explains why children are so like their parents. However, the fact that our DNA can change, given enough time, explains why we are all different in size, shape, color and many other characteristics. Such differences are the result of the many SNPs that have arisen in the DNA of our species and its predecessors.

Even though the differences between people around us are often easy to see, it is nonetheless important to bear in mind that humans are on average 99.9% genetically identical. This means that if you were to compare your chromosomes with those of a random person, we would expect to find, on average, one SNP that differs every thousand DNA nucleotides. Our DNA is also surprisingly similar to that of other animals and organisms. Our blood group and eye color are determined by which alleles we inherit. Alleles are different versions of the same gene. An allele is the version of a particular SNP or chromosome segment that you inherited from either your mother or father. Your cells carry 23 pairs of chromosomes, where one was inherited from your mother and the other from your father. This means that for any nucleotide or SNP located on an autosomal chromosome you have inherited two versions (one maternal and the other paternal). These are usually referred to as your two alleles for that particular location in the genome. For the vast majority of chromosome pairs, you will have inherited the same allele from both parents (for example, two copies of the cytosine nucleotide, represented by the letter C). A systematic examination of the nucleotide sequence of your 23 chromosome pairs would reveal millions of locations where the nucleotides you inherited from your parents were different (for example, a C allele from your mother and a T allele from your father).

8.9. Genetic Reshuffling: Microscopic bodies that carry our genes

We inherit 46 chromosomes from our parents, which are divided into 23 pairs. Chromosomes are compact packages of DNA contained within single cells. Unraveled, the ultra-thin strands of DNA from a single human cell are about 3 meters long. The only way for the entire body to carry such vast amounts of DNA is by winding it into complex bundles, known as chromosomes, that take up less space inside the cell nuclei. We inherit 46 chromosomes from our parents, which are divided into 23 pairs. One chromosome from each pair comes from our father and the other from our mother. It is this mix of chromosomes from our parents that determine our characteristics, including our propensity to develop various diseases.

Recombination takes place when large segments of DNA are exchanged between each pair of chromosomes. The chromosomes you inherited from your parents were actually a mosaic of chromosomes they inherited from their parents, that is, your grandparents. This is due to a process called recombination. Recombination takes place when germ cells (egg or sperm) are produced, when large segments of DNA are exchanged between each pair of chromosomes. This kind of genetic shuffling means that any chromosome you inherited from your mother is in fact a mosaic of chromosomes she inherited from her parents, and so on. This reshuffling increases the possible number of combinations of genetic variants, which in turn ensures greater variability of characteristics among individuals.

X + Y = male

X + X = female

8.10. How genes determine sex

Of the 23 chromosome pairs, 22 are known as autosomal, where the paired chromosomes are almost identical in size and content. The remaining pair consists of sex chromosomes, known as such because they carry the genes responsible for sex determination. For this pair, if you inherited an X chromosome from both mother and father, you are female. Alternatively, if you inherited an X chromosome from your

mother and a Y chromosome from your father, then you are male. The X and Y chromosomes are very different in size and content.

8.11. Polyploidy

Cells (and their owners) are polyploid if they contain more than two haploid (n) sets of chromosomes; that is, their chromosome number is some multiple of n greater than the 2n content of diploid cells. For example, triploid (3n) and tetraploid cell (4n) cells are polyploid.

Polyploidy in plants

Polyploidy is very common in plants, especially in angiosperms. From 30% to 70% of today's angiosperms are thought to be polyploid. Species of coffee plant with 22, 44, 66, and 88 chromosomes are known. This suggests that the ancestral condition was a plant with a haploid (n) number of 11 and a diploid (2n) number of 22, from which evolved the different polyploid descendants.

In fact, the chromosome content of most plant groups suggests that the basic angiosperm genome consists of the genes on 7–11 chromosomes. Domestic wheat, with its 42 chromosomes, is probably hexaploid (6n), where n (the ancestral haploid number) was 7.

Some other examples:

Plant	Probable ancestral haploid number	Chromosome number	Ploidy level
domestic oat	7	42	6n

Peanut	10	40	4n
sugar cane	10	80	8n
banana	11	22, 33	2n, 3n
white potato	12	48	4n
tobacco	12	48	4n
Cotton	13	52	4n
Apple	17	34, 51	2n, 3n

Polyploid plants not only have larger cells but the plants themselves are often larger. This has led to the deliberate creation of polyploid varieties of such plants as watermelons, marigolds, and snapdragons.

Origin of Polyploidy

Polyploidy has occurred often in the evolution of plants.

The process can begin if diploid (2n) gametes are formed. These can arise in at least two ways.

- The gametes may be formed by mitosis instead of meiosis.
- Plants, in contrast to animals, form germ cells (sperm and eggs) from somatic tissues. If the chromosome content of a precursor somatic cell has accidentally doubled (e.g., as a result of passing through S phase of the cell cycle without following up with mitosis and cytokinesis), then gametes containing 2n chromosomes are formed.

Polyploidy also occurs naturally in certain plant tissues.

- As the endosperm (3n) develops in corn (maize) kernels (Zea mays), its cells undergo successive rounds (as many as 5) of endoreplication producing nuclei that range as high as 96n.
- When rhizobia infect the roots of their legume host, they induce the infected cells to undergo endoreplication producing cells that can become 128n (from 6 rounds of endoreplication).

Polyploidy can also be induced in the plant-breeding laboratory by treating dividing cells with colchicine. This drug disrupts microtubules and thus prevents the formation of a spindle. Consequently, the duplicated chromosomes fail to separate in mitosis. Onion cells exposed to colchicine for several days may have over 1000 chromosomes inside.

Polyploidy and Speciation

When a newly-arisen tetraploid (4n) plant tries to breed with its ancestral species (a backcross), triploid offspring are formed. These are sterile because they cannot form gametes with a balanced assortment of chromosomes.

However, the tetraploid plants can breed with each other. So in one generation, a new species has been formed. Polyploidy even allows the formation of new species derived from different ancestors.

In 1928, the Russian plant geneticist Karpechenko produced a new species by crossing a cabbage with a radish. Although belonging to different genera (*Brassica* and *Raphanus* respectively), both parents have a diploid number of 18. Fusion of their respective gametes (n=9) produced mostly infertile hybrids.

However, a few fertile plants were formed, probably by the spontaneous doubling of the chromosome number in somatic cells that went on to form gametes (by meiosis). Thus these contained 18 chromosomes — a complete set of both cabbage (n=9) and radish (n=9) chromosomes.

Fusion of these gametes produced vigorous, fully-fertile, polyploid plants with 36 chromosomes. (They had the roots of the cabbage and the leaves of the radish.)

These plants could breed with each other but not with either the cabbage or radish ancestors, so Karpechenko had produced a new species.

The process also occurs in nature. Three species in the mustard family (Brassicaceae) appear to have arisen by hybridization and polyploidy from three other ancestral species:

- *B. oleracea* (cabbage, broccoli, etc.) hybridized with B. nigra (black mustard) → B. carinata (Abyssinian mustard).
- *B. oleracea* x *B. campestris* (turnips) → *B. napus* (rutabaga)
- *B. nigra* x *B. campestris* → *B. juncea* (leaf mustard)

Modern wheat and perhaps some of the other plants listed in the table above have probably evolved in a similar way.

Polyploidy in animals

Polyploidy is much rarer in animals. It is found in some insects, fishes, amphibians, and reptiles. Until recently, no polyploid **mammal** was known. However, the 23 September 1999 issue of **Nature** reported that a polyploid (tetraploid; **4n** = 102) rat has been found in Argentina.

Polyploid cells are larger than diploid ones; not surprising in view of the increased amount of DNA in their nucleus. The liver cells of the Argentinian rat are larger than those of its diploid relatives, and its sperm are huge in comparison. Normal mammalian sperm heads contain some 3.3 picograms (10^{-12} g) of DNA; the sperm of the rat contains 9.2 pg.

Although only one mammal is known to have all its cells polyploid, many mammals have polyploid cells in certain of their organs, e.g, the liver.

Polyploidy is a term used to describe cells and organisms containing more than two paired (homologous) sets of chromosomes. Most eukaryotic species are diploid, meaning they have two sets of chromosomes — one set inherited from each parent. However **polyploidy** is found in some organisms and is especially common in plants. In addition, polyploidy also occurs in some tissues of animals who are otherwise diploid, such as human muscle tissues. This is known as **endopolyploidy**. (Monoploid organisms also occur; a monoploid has only one set of chromosomes. These include the vast majority of prokaryotes.)

Polyploidy refers to a numerical change in a whole set of chromosomes. Organisms in which a particular chromosome, or chromosome segment, is under- or overrepresented are said to be **aneuploid** (from the Greek words meaning "not," "good," and "fold"). Therefore the distinction between aneuploidy and polyploidy is that aneuploidy refers to a numerical change in part of the chromosome set, whereas polyploidy refers to a numerical change in the whole set of chromosomes.

Polyploidy may occur due to abnormal cell division, either during mitosis, or commonly during metaphase I in meiosis.

Polyploidy occurs in some animals, such as goldfish, salmon, and salamanders, but is especially common among ferns and flowering plants including both wild and cultivated species. Wheat, for example, after millennia of hybridization and modification by humans, has strains that are **diploid** (two sets of chromosomes), **tetraploid** (four sets of chromosomes) with the common name of durum or macaroni wheat, and **hexaploid** (six sets of chromosomes) with the common name of bread wheat. Many agriculturally important plants of the genus *Brassica* are also tetraploids. Polyploidization is a mechanism of sympatric speciation because polyploids are usually unable to interbreed with their diploid ancestors.

Polyploidy can be induced in plants and cell cultures by some chemicals: the best known is colchicine, which can result in chromosome doubling, though its use may have

other less obvious consequences as well. Oryzalin also will double the existing chromosome content.

Polyploid types are labeled according to the number of chromosome sets in the nucleus:

- **triploid** (three sets; 3x), for example seedless watermelons, common in the phylum Tardigrad
- **tetraploid** (four sets; 4x), for example Salmonidae fish
- **pentaploid** (five sets; 5x), for example Kenai Birch (*Betula papyrifera* var. *kenaica*)
- **hexaploid** (six sets; 6x), for example wheat, kiwifruit
- **octaploid** (eight sets; 8x), for example *Acipenser* (genus of sturgeon fish), dahlias
- **decaploid** (ten sets; 10x), for example certain strawberries
- **dodecaploid** (twelve sets; 12x), for example the plant *Celosia argentea* and the amphibian *Xenopus ruwenzoriensis*

Polyploidy in animals (non-human)

Examples in animals are more common in the 'lower' forms such as flatworms, leeches, and brine shrimp. Polyploid animals are often sterile, so they often reproduce by parthenogenesis. Polyploid lizards are also quite common and parthenogenetic. Polyploid mole salamanders (mostly triploids) are all female and reproduce by kleptogenesis, "stealing" spermatophores from diploid males of related species to trigger egg development but not incorporating the males' DNA into the offspring. While mammalian liver cells are polyploid, rare instances of polyploid mammals are known, but most often result in prenatal death.

One of the few known exceptions to this 'rule' is an octodontid rodent of Argentina's harsh desert regions, known as the Plains Viscacha-Rat (*Tympanoctomys barrerae*). This rodent is not a rat, but kin to guinea pigs and chinchillas. Its "new" diploid [2n]

number is 102 and so its cells are roughly twice normal size. Its closest living relation is *Octomys mimax*, the Andean Viscacha-Rat of the same family, whose 2n = 56. It is surmised that an *Octomys*-like ancestor produced tetraploid (i.e., 4n = 112) offspring that were, by virtue of their doubled chromosomes, reproductively isolated from their parents; but that these likely survived the ordinarily catastrophic effects of polyploidy in mammals by shedding (via translocation or some similar mechanism) the "extra" set of sex chromosomes gained at this doubling. (The closely related Golden Vizcacha Rat, 2n = 96, is thought to have arisen via roughly the same process).

Polyploidy in humans

True polyploidy rarely occurs in humans, although it occurs in some tissues (especially in the liver). Aneuploidy is more common.

Polyploidy occurs in humans in the form of triploidy, with 69 chromosomes (sometimes called 69,XXX), and tetraploidy with 92 chromosomes (sometimes called 92,XXXX). Triploidy, usually due to polyspermy, occurs in about 2–3% of all human pregnancies and ~15% of miscarriages. The vast majority of triploid conceptions end as miscarriage and those that do survive to term typically die shortly after birth. In some cases survival past birth may occur longer if there is mixoploidy with both a diploid and a triploid cell population present.

Triploidy may be the result of either digyny (the extra haploid set is from the mother) or diandry (the extra haploid set is from the father). Diandry is mostly caused by reduplication of the paternal haploid set from a single sperm, but may also be the consequence of dispermic (two sperm) fertilization of the egg. Digyny is most commonly caused by either failure of one meiotic division during oogenesis leading to a diploid oocyte or failure to extrude one polar body from the oocyte. Diandry appears to predominate among early miscarriages while digyny predominates among triploidy that survives into the fetal period. However, among early miscarriages, digyny is also more common in those cases <8.5 weeks gestational age or those in which an embryo is present. There are also two distinct phenotypes in triploid placentas and fetuses that are

dependent on the origin of the extra haploid set. In digyny there is typically an asymmetric poorly grown fetus, with marked adrenal hypoplasia and a very small placenta. In diandry, a partial hydatidiform mole develops. These parent-of-origin effects reflect the effects of genomic imprinting. Complete tetraploidy is more rarely diagnosed than triploidy, but is observed in 1–2% of early miscarriages. However, some tetraploid cells are commonly found in chromosome analysis at prenatal diagnosis and these are generally considered 'harmless'. It is not clear whether these tetraploid cells simply tend to arise during *in vitro* cell culture or whether they are also present in placental cells *in vivo*. There are, at any rate, very few clinical reports of fetuses/infants diagnosed with tetraploidy mosaicism.

Mixoploidy is quite commonly observed in human preimplantation embryos and includes haploid/diploid as well as diploid/tetraploid mixed cell populations. It is unknown whether these embryos fail to implant and are therefore rarely detected in ongoing pregnancies or if there is simply a selective process favoring the diploid cells.

Polyploidy in plants

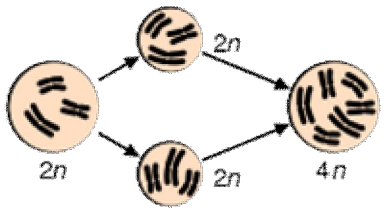


Fig. 5: Speciation via polyploidy: A diploid cell undergoes failed meiosis, producing diploid gametes, which self-fertilize to produce a tetraploid zygote.

Polyploidy is pervasive in plants and some estimates suggest that 30–80% of living plant species are polyploid, and many lineages show evidence of ancient polyploidy (paleopolyploidy) in their genomes. Huge explosions in angiosperm species diversity appear to have coincided with the timing of ancient genome duplications shared by

many species. It has been established that 15% of angiosperm and 31% of fern speciation events are accompanied by ploidy increase. Polyploid plants can arise spontaneously in nature by several mechanisms, including meiotic or mitotic failures, and fusion of unreduced (2n) gametes. Both autopolyploids (e.g. potato) and allopolyploids (e.g. canola, wheat, cotton) can be found among both wild and domesticated plant species. Most polyploids display heterosis relative to their parental species, and may display novel variation or morphologies that may contribute to the processes of speciation and eco-niche exploitation. The mechanisms leading to novel variation in newly formed allopolyploids may include gene dosage effects (resulting from more numerous copies of genome content), the reunion of divergent gene regulatory hierarchies, chromosomal rearrangements, and epigenetic remodeling, all of which affect gene content and/or expression levels. Many of these rapid changes may contribute to reproductive isolation and speciation.

Lomatia tasmanica is an extremely rare Tasmanian shrub which is triploid and sterile, and reproduction is entirely vegetative with all plants having the same genetic structure.

There are few naturally occurring polyploid conifers. One example is the giant tree *Sequoia sempervirens* or Coast Redwood which is a hexaploid (6x) with 66 chromosomes (2n = 6x = 66), although the origin is unclear.

Polyploid crops

Polyploid plants tend to be larger and better at flourishing in early succession habitats such as farm fields.] In the breeding of crops, the tallest and best thriving plants are selected for. Thus, many crops (and agricultural weeds) may have unintentionally been bred to a higher level of ploidy.

The induction of polyploidy is a common technique to overcome the sterility of a hybrid species during plant breeding. For example, Triticale is the hybrid of wheat (*Triticum turgidum*) and rye (*Secale cereale*). It combines sought-after characteristics of the

parents, but the initial hybrids are sterile. After polyploidization, the hybrid becomes fertile and can thus be further propagated to become triticale.

In some situations polyploid crops are preferred because they are sterile. For example many seedless fruit varieties are seedless as a result of polyploidy. Such crops are propagated using asexual techniques such as grafting.

Polyploidy in crop plants is most commonly induced by treating seeds with the chemical colchicine.

Examples of polyploid crops

- Triploid crops: apple, banana, citrus, ginger, watermelon
- Tetraploid crops: apple, durum or macaroni wheat, cotton, potato, cabbage, leek, tobacco, peanut, kinnow, Pelargonium
- Hexaploid crops: chrysanthemum, bread wheat, triticale, oat, kiwifruit.
- Octaploid crops: strawberry, dahlia, pansies, sugar cane

Some crops are found in a variety of ploidies: tulips and lilies are commonly found as both diploid and as triploid; daylilies (Hemerocallis cultivars) are available as either diploid or tetraploid; apples and kinnows can be diploid, triploid, or tetraploid.

Autopolyploidy

Autopolyploids are polyploids with multiple chromosome sets derived from a single species. Autopolyploids can arise from a spontaneous, naturally occurring genome doubling, like the potato. Others might form following fusion of 2n gametes (unreduced gametes). Bananas and apples can be found as autotriploids. Autopolyploid plants typically display polysomic inheritance, and are therefore often infertile and propagated clonally perfect.

Allopolyploidy

Allopolyploids are polyploids with chromosomes derived from different species. Precisely it is the result of doubling of chromosome number in an F1 hybrid. *Triticale* is an example of an allopolyploid, having six chromosome sets, allohexaploid, four from wheat (*Triticum turgidum*) and two from rye (*Secale cereale*). *Amphidiploid* is another word for an allopolyploid. Some of the best examples of allopolyploids come from the Brassicas, and the Triangle of U describes the relationships among the three common diploid Brassicas (*B. oleracea, B. rapa,* and *B. nigra*) and three allotetraploids (*B. napus, B. juncea,* and *B. carinata*) derived from hybridization among the diploids.

Paleopolyploidy

Ancient genome duplications probably occurred in the evolutionary history of all life. Duplication events that occurred long ago in the history of various evolutionary lineages can be difficult to detect because of subsequent diploidization (such that a polyploid starts to behave cytogenetically as a diploid over time) as mutations and gene translations gradually make one copy of each chromosome unlike its other copy.

In many cases, these events can be inferred only through comparing sequenced genomes. Examples of unexpected but recently confirmed ancient genome duplications include baker's yeast (*Saccharomyces cerevisiae*), mustard weed/thale cress (*Arabidopsis thaliana*), rice (*Oryza sativa*), and an early evolutionary ancestor of the vertebrates (which includes the human lineage) and another near the origin of the teleost fishes. Angiosperms (flowering plants) have paleopolyploidy in their ancestry. All eukaryotes probably have experienced a polyploidy event at some point in their evolutionary history.

Karyotype

A karyotype is the characteristic chromosome complement of a eukaryote species. The preparation and study of karyotypes is part of cytology and, more specifically, cytogenetics.

Although the replication and transcription of DNA is highly standardized in eukaryotes, the same cannot be said for their karotypes, which are highly variable between species in chromosome number and in detailed organization despite being constructed out of the same macromolecules. In some cases there is even significant variation within species. This variation provides the basis for a range of studies in what might be called evolutionary cytology.

Paralogous

The term is used to describe the relationship among duplicated genes or portions of chromosomes that derived from a common ancestral DNA. Paralogous segments of DNA may arise spontaneously by errors during DNA replication, copy and paste transposons, or whole genome duplications.

Homologous

The term is used to describe the relationship of similar chromosomes that pair at mitosis and meiosis. In a diploid, one homolog is derived from the male parent (sperm) and one is derived from the female parent (egg). During meiosis and gametogenesis, homologous chromosomes pair and exchange genetic material by recombination, leading to the production of sperm or eggs with chromosome haplotypes containing novel genetic variation.

Homoeologous

The term *homoeologous*, also spelled *homeologous*, is used to describe the relationship of similar chromosomes or parts of chromosomes brought together following inter-species hybridization and allopolyploidization, and whose relationship was completely homologous in an ancestral species. In allopolyploids, the homologous chromosomes within each parental sub-genome should pair faithfully during meiosis, leading to disomic inheritance; however in some allopolyploids, the homoeologous chromosomes of the parental genomes may be nearly as similar to one another as the homologous

chromosomes, leading to tetrasomic inheritance (four chromosomes pairing at meiosis), intergenomic recombination, and reduced fertility.

Example of homoeologous chromosomes

Durum wheat is the result of the inter-species hybridization of two diploid grass species *Triticum urartu* and *Aegilops speltoides*. Both the diploid ancestors had two sets of 7 chromosomes, which were similar in terms of size and genes contained on them. Durum wheat contains two sets of chromosomes derived from *Triticum urartu* and two sets of chromosomes derived from *Aegilops speltoides*. Each chromosome pair derived from the *Triticum urartu* parent is **homoeologous** to the opposite chromosome pair derived from the *Aegilops speltoides* parent, though each chromosome pair unto itself is **homologous**.

Reproductive isolation mechanisms

The mechanisms of reproductive isolation or hybridization barriers are a collection of mechanisms, behaviors and physiological processes that prevent the members of two different species that cross or mate from producing offspring, or which ensure that any offspring that may be produced is not fertile. These barriers maintain the integrity of a species over time, reducing or directly impeding gene flow between individuals of different species, allowing the conservation of each species' characteristics.

The mechanisms of reproductive isolation have been classified in a number of ways. Zoologist Ernst Mayr classified the mechanisms of reproductive isolation in two broad categories: those that act before fertilization (or before mating in the case of animals, which are called pre-copulatory) and those that act after. These have also been termed pre-zygotic and post-zygotic mechanisms. The different mechanisms of reproductive isolation are genetically controlled and it has been demonstrated experimentally that they can evolve in species whose geographic distribution overlaps (sympatric speciation) or as the result of adaptive divergence that accompanies allopatric speciation.

Isolation mechanisms that occur before breeding or copulation (pre-zygotic isolation)

Pre-zygotic isolation mechanisms are the most economic in terms of the biological efficiency of a population, as resources are not wasted on the production of a descendent that is weak, non-viable or sterile.

Temporal or habitat isolation

Any of the factors that prevent potentially fertile individuals from meeting will reproductively isolate the members of distinct species. The types of barriers that can cause this isolation include: different habitats, physical barriers, and a difference in the time of sexual maturity or flowering. When factors change, especially physical barriers, often, species will branch off.

An example of the ecological or habitat differences that impede the meeting of potential pairs occurs in two fish species of the family *Gasterosteidae* (sticklebacks). One species lives all year round in fresh water, mainly in small streams. The other species lives in the sea during winter, but in spring and summer individuals migrate to river estuaries to reproduce. The members of the two populations are reproductively isolated due to their adaptations to distinct salt concentrations. An example of reproductive isolation due to differences in the mating season is found in the toad species *Bufo americanus* and *Bufo fowleri* . The members of these species can be successfully crossed in the laboratory producing healthy, fertile hybrids. However, mating does not occur in the wild even though the geographical distribution of the two species overlaps. The reason for the absence of inter-species mating is that *B. americanus* mates in early summer and *B. fowleri* in late summer. Certain plant species, such as *Tradescantia canaliculata* and *T. subaspera*, are sympatric throughout their geographic distribution yet they are

reproductively isolated as they flower at different times of the year. In addition, one species grows in sunny areas and the other in deeply shaded areas.

Sexual isolation by behavior or conduct

The different mating rituals of animal species creates extremely powerful reproductive barriers, termed sexual or behavior isolation, that isolate apparently similar species in the majority of the groups of the animal kingdom. In dioecious species, males and females have to search for a partner, be in proximity to each other, carry out the complex mating rituals and finally copulate or release their gametes into the environment in order to breed.

The songs of birds, insects and many other animals are part of a ritual to attract potential partners of their own species. The song presents specific patterns recognizable only by members of the same species, they therefore represent a mechanism of reproductive isolation. The recording is the song of a species of cicada, recorded in Lower Hutt, New Zealand on 15th February 2006.

Mating dances, the songs of males to attract females or the mutual grooming of pairs, are all examples of typical courtship behavior that allows both recognition and reproductive isolation. This is because each of the stages of courtship depend on the behavior of the partner. The male will only move onto the second stage of the exhibition if the female shows certain responses in her behavior. He will only pass onto the third stage when she displays a second key behavior. The behaviors of both interlink, are synchronized in time and lead finally to copulation or the liberation of gametes into the environment. No animal that is not physiologically suitable for fertilization can complete this demanding chain of behavior. In fact, the smallest difference in the courting patterns of two species is enough to prevent mating (for example, a specific song pattern acts as an isolation mechanism in distinct species of grasshopper of the genus *Chorthippus*. Even where there are minimal morphological differences between species, differences in behavior can be enough to prevent mating. For example, *Drosophila melanogaster* and *D. simulans* which are considered twin species due to their morphological similarity,

do not mate even if they are kept together in a laboratory. *Drosophila ananassae* and *D. pallidosa* are twin species from Melanesia. In the wild they rarely produce hybrids, although in the laboratory it is possible to produce fertile offspring. Studies of their sexual behavior show that the males court the females of both species but the females show a marked preference for the males of their own species which they readily mate with. A different regulator region has been found on Chromosome II of both species that affects the selection behavior of the females.

Pheromones are an important causal factor in the sexual isolation of insect species. The function of these compounds is basically to allow the identification of individuals of the same species and of the same or different sex. When they are volatile compounds they function as a wide reaching chemical signal. In other cases their dispersion is reduced and are only sensed at a short distance or by contact. In species of the *melanogaster* group of *Drosophila*, the pheromones of the females are mixtures of different compounds, there is a clear dimorphism in the type and/or quantity of compounds present for each sex. In addition, there are differences in the quantity and quality of constituent compounds between related species, it is assumed that the pheromones serve to distinguish between individuals of each species. An example of the role of pheromones in sexual isolation is found in 'corn borers' in the genus *Ostrinia*. There are two twin species in Europe that occasionally cross. The females of both species produce pheromones that contain a volatile compound which has two isomers, E and Z; 99% of the compound produced by the females of one species is in the E isomer form, while the females of the other produce 99% isomer Z. The production of the compound is controlled by just one locus and the interspecific hybrid produces an equal mix of the two isomers. The males, for their part, almost exclusively detect the isomer emitted by the females of their species, such that the hybridization although possible is scarce. The perception of the males is controlled by one gene, distinct from the one for the production of isomers, the heterozygous males show a moderate response to the odour of either type. In this case, just 2 'loci' produce the effect of ethological isolation between species that are genetically very similar.

Sexual isolation between two species can be asymmetrical. This can happen when the mating that produces descendants only allows one of the two species to function as the female progenitor and the other as the male, while the reciprocal cross does not occur. For instance, half of the wolves tested in the Great Lakes area of America show mitochondrial DNA sequences of coyotes. While mitochondrial DNA from wolves is never found in coyote populations. This probably reflects an asymmetry in inter-species mating due to the difference in size of the two species as male wolves take advantage of their greater size in order to mate with female coyotes, while female wolves and male coyotes do not mate.

Mechanical isolation

The flowers of many species of *Angiosperm* have evolved to attract and reward a single or a few pollinator species (insects, birds, mammals). Their wide diversity of form, colour, fragrance and presence of nectar is, in many cases, the result of coevolution with the pollinator species. This dependency on its pollinator species also acts as a reproductive isolation barrier.

Mating pairs may not be able to couple successfully if their genitals are not compatible. The relationship between the reproductive isolation of species and the form of their genital organs was signaled for the first time in 1844 by the French entomologist Léon Dufour. Insects' rigid carapaces act in a manner analogous to a lock and key, as they will only allow mating between individuals with complementary structures, that is, males and females of the same species (termed *co-specifics*).

Evolution has led to the development of genital organs with increasingly complex and divergent characteristics, which will cause mechanical isolation between species. Certain characteristics of the genital organs will often have converted them into mechanisms of isolation. However, numerous studies show that organs that are anatomically very different can be functionally compatible, indicating that other factors also determine the form of these complicated structures.

Mechanical isolation also occurs in plants and this is related to the adaptation and coevolution of each species in the attraction of a certain type of pollinator (where pollination is zoophilic) through a collection of morphophysiological characteristics of the flowers (called floral syndrome), in such a way that the transport of pollen to other species does not occur.

Gametic Isolation

The synchronous spawning of many species of coral in marine reefs means that inter-species hybridization can take place as the gametes of hundreds of individuals of tens of species are liberated into the same water at the same time. Approximately a third of all the possible crosses between species are compatible, in the sense that the gametes will fuse and lead to individual hybrids. This hybridization apparently plays a fundamental role in the evolution of coral species. However, the other two-thirds of possible crosses are incompatible. It has been observed that in sea urchins of the genus *Strongylocentrotus* the concentration of spermatocytes that allow 100% fertilization of the ovules of the same species is only able to fertilize 1.5% of the ovules of other species. This inability to produce hybrid offspring, despite the fact that the gametes are found at the same time and in the same place, is due to a phenomenon known as *gamete incompatibility*, which is often found between marine invertebrates, and whose physiological causes are not fully understood.

In some *Drosophila* crosses, the swelling of the female's vagina has been noted following insemination. This has the effect of consequently, preventing the fertilization of the ovule by sperm of a different species.

In plants the pollen grains of a species can germinate in the stigma and grow in the style of other species. However, the growth of the pollen tubes may be detained at some point between the stigma and the ovules, in such a way that fertilization does not take place. This mechanism of reproductive isolation is common in the Angiosperms and is called *cross-incompatibility* or *incongruence*. A relationship exists between self-incompatibility and the phenomenon of cross-incompatibility. In general crosses

between individuals of a self-compatible species (SC) with individuals of a self-incompatible (SI) species give hybrid offspring. On the other hand, a reciprocal cross (SI x SC) will not produce offspring, because the pollen tubes will not reach the ovules. This is known as *unilateral incompatibility*, which also occurs when two SC or two SI species are crossed.

Isolation mechanisms that occur after breeding or copulation (post-zygotic isolation)

A number of mechanisms which act after fertilisation preventing successful inter-population crossing are discussed below.

Zygote mortality and non-viability of hybrids

A type of incompatibility that is found as often in plants as in animals occurs when the ovule is fertilized but the zygote does not develop, or it develops and the resulting individual has a reduced viability. This is the case for crosses between species of the frog genus, where widely differing results are observed depending of the species involved. In some crosses there is no segmentation of the zygote (or it may be that the hybrid is extremely non-viable and changes occur from the first mitosis). In others, normal segmentation occurs in the blastula but gastrulation fails. Finally, in other crosses, the initial stages are normal but errors occur in the final phases of embryo development. This indicates differentiation of the embryo development genes (or gene complexes) in these species and these differences determine the non-viability of the hybrids.

Similar results are observed in mosquitos of the Culex genus, but the differences are seen between reciprocal crosses, from which it is concluded that the same effect occurs in the interaction between the genes of the cell nucleus (inherited from both parents) as occurs in the genes of the cytoplasmic organelles which are inherited solely from the female progenitor through the cytoplasm of the ovule.

In Angiosperms, the successful development of the embryo depends on the normal functioning of its endosperm.

The failure of endosperm development and its subsequent abortion has been observed in many interploidal crosses (that is, those between populations with a particular degree of intra or interspecific ploidy, and in certain crosses in species with the same level of ploidy. The collapse of the endosperm, and the subsequent abortion of the hybrid embryo is one of the most common post-fertilization reproductive isolation mechanism found in angiosperms.

Hybrid sterility

Mules are hybrids with interspecific sterility. A hybrid has normal viability but is deficient in terms of reproduction or is sterile. This is demonstrated by the mule and in many other well known hybrids. In all of these cases sterility is due to the interaction between the genes of the two species involved; to chromosomal imbalances due to the different number of chromosomes in the parent species; or to nucleus-cytoplasmic interactions such as in the case of *Culex* described above.

Hinnies and mules are hybrids resulting from a cross between a horse and an ass or between a mare and a donkey, respectively. These animals are nearly always sterile due to the difference in the number of chromosomes between the two parent species. Both horses and donkeys belong to the genus *Equus*, but *Equus caballus* has 64 chromosomes, while *Equus asinus* only has 62. A cross will produce offspring (mule or hinny) with 63 chromosomes that will not form pairs which means that they do not divide in a balanced manner during meiosis. It is curious that they can cross with each other but the mule and the hinny are actually animals created by humans, as in the wild the species ignore each other and do not cross. In order to obtain mules or hinnies it is necessary to train the progenitors to accept copulation between the species or create them through artificial insemination.

The sterility of many of the interspecific hybrids among the angiosperms is a widely recognised and studied phenomenon. There are a variety of causes that can determine the interspecific sterility of hybrids in plants, these may be genetic, related to the genomes or the interaction between nuclear and cytoplasmic factors, as will be discussed in the corresponding section. Nevertheless, it should be pointed out that - on the contrary to the situation in animals - hybridization in plants is a stimulus for the creation of new species. Indeed, although the hybrid may be sterile it can continue to multiply in the wild through the mechanisms of asexual reproduction, be they vegetative propagation or apomixis or the production of seeds. Indeed, interspecific hybridization can be associated with polyploidia and, in this way, the origin of new species that are called allopolyploids. *Rosa canina*, for example, is the result of multiple hybridizations, or there is a type of wheat that is a allohexaploid that contains the genomes of three different species.

Multiple mechanisms

In general, the barriers that separate species do not consist of just one mechanism. The twin species of *Drosophila*, *D. pseuoobscura* and *D. persimilis*, are isolated from each other by habitat (*persimilis* generally lives in colder regions at higher altitudes), by the timing of the mating season (*persimilis* is generally more active in the morning and *pseuoobscura* at night) and by behavior during mating (the females of both species prefer the males of their respective species). In this way, although the distribution of these species overlaps in wide areas of the west of the United States of America, these isolation mechanisms are sufficient to keep the species separated. Such that, only a few fertile females have been found amongst the other species among the thousands that have been analyzed. However, when hybrids are produced between both species, the gene flow between the two will continue to be impeded as the hybrid males are sterile. Also, and in contrast with the great vigor shown by the sterile males, the descendants of the backcrosses of the hybrid females with the parent species are weak and notoriously non-viable. This last mechanism restricts even more the genetic interchange between the two species of fly in the wild.

Hybrid gender: Haldane's Rule

Haldane's Rule states that when one of the two sexes is absent in interspecific hybrids between two specific species, then the gender that is not produced, is rare or is sterile is the heterozygous (or heterogametic) sex. In mammals, at least, there is growing evidence to suggest that this is due to high rates of mutation of the genes determining masculinity in the Y chromosome.

It has been suggested that Haldane's Rule simply reflects the fact that the male gender is more sensitive than the female when the sex-determining genes are included in a hybrid genome. But there are also organisms in which the heterozygous sex is the female: birds and butterflies and the law is followed in these organisms. Therefore, it is not a problem related to sexual development, nor with the sex chromosomes. Haldane proposed that the stability of hybrid individual development requires the full gene complement of each parent species, so that the hybrid of the heterozygous sex is unbalanced (i.e. missing at least one chromosome from each of the parental species). For example, the hybrid male obtained by crossing *D. melanogaster* females with *D. simulans* males, which is non-viable, lacks the X chromosome of *D. Simulans*

The genetics of reproductive isolation barriers

The genetics of ethological isolation barriers will be discussed first. Pre-copulatory isolation occurs when the genes necessary for the sexual reproduction of one species differ from the equivalent genes of another species, such that if a male of species A and a female of species B are placed together they are unable to copulate. Study of the genetics involved in this reproductive barrier tries to identify the genes that govern distinct sexual behaviors in the two species. The males of *Drosophila melanogaster* and those of *D. simulans* conduct an elaborate courtship with their respective females, which are different for each species, but the differences between the species are more quantitative than qualitative. In fact the *simulans* males are able to hybridize with the *melanogaster* females. Although there are lines of the latter species that can easily cross there are others that are hardly able to. Using this difference, it is possible to

assess the minimum number of genes involved in pre-copulatory isolation between the *melanogaster* and *simulans* species and their chromosomal location.

In experiments, flies of the *D. melanogaster* line, which hybridizes readily with *simulans*, were crossed with another line that it does not hybridize with, or rarely. The females of the segregated populations obtained by this cross were placed next to *simulans* males and the percentage of hybridization was recorded, which is a measure of the degree of reproductive isolation. It was concluded from this experiment that 3 of the 8 chromosomes of the haploid complement of *D. melanogaster* carry at least one gene that affects isolation, such that substituting one chromosome from a line of low isolation with another of high isolation reduces the hybridization frequency. In addition, interactions between chromosomes are detected so that certain combinations of the chromosomes have a multiplying effect. Cross incompatibility or incongruence in plants is also determined by major genes that are not associated at the self-incompatibility *S locus*.

Post copulation or fertilization isolation mechanisms in animals

Reproductive isolation between species appears, in certain cases, a long time after fertilization and the formation of the zygote, as happens - for example - in the twin species *Drosophila pavani* and *D. gaucha*. The hybrids between both species are not sterile, in the sense that they produce viable gametes, ovules and spermatozoa. However, they cannot produce offspring as the sperm of the hybrid male do not survive in the semen receptors of the females, be they hybrids or from the parent lines. In the same way, the sperm of the males of the two parent species do not survive in the reproductive tract of the hybrid female. This type of post copulatory isolation appears as the most efficient system for maintaining reproductive isolation in many species.

In fact, the development of a zygote into an adult is a complex and delicate process of interactions between genes and the environment that must be carried out precisely, and if there is any alteration in the usual process, caused by the absence of a necessary gene or the presence of a different one, it can arrest the normal development causing

the non-viability of the hybrid or its sterility. It should be borne in mind that half of the chromosomes and genes of a hybrid are from one species and the other half come from the other. If the two species are genetically different, there is little possibility that the genes from both will act harmoniously in the hybrid. From this perspective, only a few genes would be required in order to bring about post copulatory isolation, as opposed to the situation described previously for pre-copulatory isolation.

In many species where pre-copulatory reproductive isolation does not exist, hybrids are produced but they are of only one sex. This is the case for the hybridization between females of *Drosophila simulans* and *Drosophila melanogaster* males: the hybridized females die early in their development so that only males are seen among the offspring. However, populations of *D. simulans* have been recorded with genes that permit the development of adult hybrid females, that is, the viability of the females is “rescued”. It is assumed that the normal activity of these speciation genes is to “inhibit” the expression of the genes that allow the growth of the hybrid. There will also be regulator genes.

A number of these genes have been found in the *melanogaster* species group. The first to be discovered was “Lhr” (Lethal hybrid rescue) located in Chromosome II of *D. simulans*. This dominant allele allows the development of hybrid females from the cross between *simulans* females and *melanogaster* males. A different gene, also located on Chromosome II of *D. simulans* is “Shfr” that also allows the development of female hybrids, its activity being dependent on the temperature at which development occurs. Other similar genes have been located in distinct populations of species of this group. In short, only a few genes are needed for an effective post copulatory isolation barrier mediated through the non-viability of the hybrids.

As important as identifying an isolation gene is knowing its function. The *Hmr* gene, linked to the X chromosome and implicated in the viability of male hybrids between *D. melanogaster* and *D. simulans*, is a gene from the proto-oncogene family *myb*, that codes for a transcriptional regulator. Two variants of this gene function perfectly well in each separate species, but in the hybrid they do not function correctly, possibly due to

the different genetic background of each species. Examination of the allele sequence of the two species shows that change of direction substitutions are more abundant than synonymous substitutions, suggesting that this gene has been subject to intense natural selection.

The Dobzhansky-Muller model proposes that reproductive incompatibilities between species are caused by the interaction of the genes of the respective species. It has been demonstrated recently that *Lhr* has functionally diverged in *D. simulans* and will interact with *Hmr* which, in turn, has functionally diverged in *D. melanogaster* to cause the lethality of the male hybrids. *Lhr* is located in a heterochromatic region of the genome and its sequence has diverged between these two species in a manner consistent with the mechanisms of positive selection. An important unanswered question is whether the genes detected correspond to old genes that initiated the speciation favoring hybrid non-viability, or are modern genes that have appear post-speciation by mutation, that are not shared by the different populations and that suppress the effect of the primitive non-viability genes. The *OdsH* (abbreviation of *Odysseus*) gene causes partial sterility in the hybrid between *Drosophila simulans* and a related species, *D. mauritiana*, which is only encountered on Mauritius, and is of recent origin. This gene shows monophyly in both species and also has been subject to natural selection. It is thought that it is a gene that intervenes in the initial stages of speciation, while other genes that differentiate the two species show polyphyly. *Odsh* originated by duplication in the genome of *Drosophila* and has evolved at very high rates in *D. mauritania*, while its paralogue, *unc-4*, is nearly identical between the species of the group *melanogaster*. Seemingly, all these cases illustrate the manner in which speciation mechanisms originated in nature, therefore they are collectively known as “speciation genes”, or possibly, gene sequences with a normal function within the populations of a species that diverge rapidly in response to positive selection thereby forming reproductive isolation barriers with other species. In general, all these genes have functions in the transcriptional regulation of other genes. The *Nup96* gene is another example of the evolution of the genes implicated in post-copulatory isolation. It regulates the production of one of the approximately 30 proteins required to form a

nuclear pore. In each of the *simulans* groups of *Drosophila* the protein from this gene interacts with the protein from another, as yet undiscovered, gene on the X chromosome in order to form a functioning pore. However, in a hybrid the pore that is formed is defective and causes sterility. The differences in the sequences of *Nup96* have been subject to adaptive selection, similar to the other examples of *speciation genes* described above.

Post-copulatory isolation can also arise between chromosomally differentiated populations due to chromosomal translocations and inversions. If, for example, a reciprocal translocation is fixed in a population, the hybrid produced between this population and one that does not carry the translocation will not have a complete meiosis. This will result in the production of unequal gametes containing unequal numbers of chromosomes with a reduced fertility. In certain cases, complete translocations exist that involve more than two chromosomes, so that the meiosis of the hybrids is irregular and their fertility is zero or nearly zero. Inversions can also give rise to abnormal gametes in heterozygous individuals but this effect has little importance compared to translocations. An example of chromosomal changes causing sterility in hybrids comes from the study of *Drosophila nasuta* and *D. albomicans* which are twin species from the Indo-Pacific region. There is no sexual isolation between them and the F1 hybrid is fertile. However, the F2 hybrids are relatively infertile and leave few descendants which have a skewed ratio of the sexes. The reason is that the X chromosome of *albomicans* is translocated and linked to an autosome which causes abnormal meiosis in hybrids. Robertsonian translocations are variations in the numbers of chromosomes that arise from either: the fusion of two acrocentric chromosomes into a single chromosome with two arms, causing a reduction in the haploid number, or conversely; or the fission of one chromosome into two acrocentric chromosomes, in this case increasing the haploid number. The hybrids of two populations with differing numbers of chromosomes can experience a certain loss of fertility, and therefore a poor adaptation, because of irregular meiosis.

Post copulation or fertilization isolation mechanisms in plants

In plants, hybrids often suffer from an autoimmune syndrome known as hybrid necrosis. In the hybirds, specific gene products contributed by one of the parents may be inappropriately recognized as foreign and pathogenic, and thus trigger pervasive cell death throughout the plant.In at least one case, a pathogen receptor, encoded by the most variable gene family in plants, was identified as being responsible for hybrid necrosis.

Incompatibility caused by microorganisms

In addition to the genetic causes of reproductive isolation between species there is another factor that can cause of post zygotic isolation: the presence of microorganisms in the cytoplasm of certain species. The presence of these organisms in a species and their absence in another causes the non-viability of the corresponding hybrid. For example, in the semi-species of the group *D. paulistorum* the hybrid females are fertile but the males are sterile, this is due to the presence of a mycoplasma in the cytoplasm which alters spermatogenesis leading to sterility. It is interesting that incompatibility or isolation can also arise at an intraspecific level. Populations of *D. simulans* have been studied that show hybrid sterility according to the direction of the cross. The factor determining sterility has been found to be the presence or absence of a microorganism *Wolbachia* and the populations tolerance or susceptibility to these organisms. This inter population incompatibility can be eliminated in the laboratory through the administration of a specific antibiotic to kill the microorganism. Similar situations are known in a number of insects, as around 15% of species show infections caused by this symbiont. It has been suggested that, in some cases, the speciation process has taken place because of the incompatibility caused by this bacteria. Two wasp species *Nasonia girault*i and *N. longicornis* carry two different strains of *Wolbachia*. Crosses between an infected population and one free from infection produces a nearly total reproductive isolation between the semi-species. However, if both species are free from the bacteria or both are treated with antibiotics there is no reproductive barrier. *Wolbachia* also induces incompatibility due to the weakness of the hybrids in populations of spider mites

(*Tetranychus urticae*, between *Drosophila recens* and *D. subquinaria* and between species of *Diabrotica* (beetle) and *Gryllus* (cricket). Selection for reproductive isolation

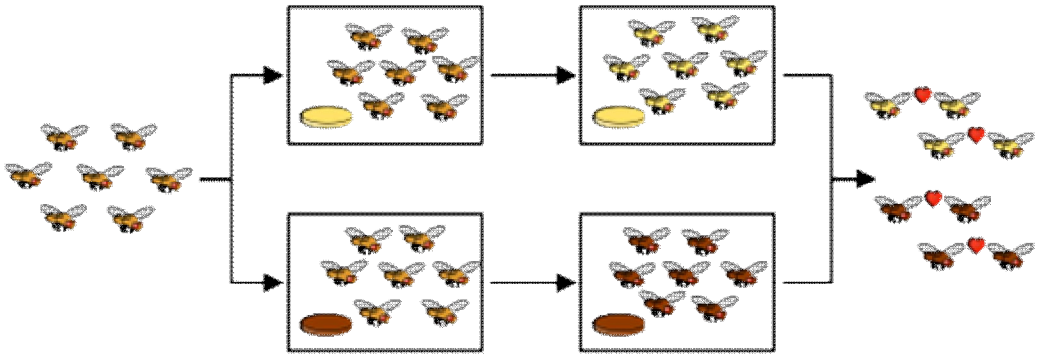
Selection for reproductive isolation between two *Drosophila* species.	
Generation	**Percentage of hybrids**
1	49
2	17,6
3	3,3
4	1,0
5	1,4
10	0,6

In 1950 K. F. Koopman reported results from experiments designed to examine the hypothesis that selection can increase reproductive isolation between populations. He used *D. pseudoobscura* and *D. persimilis* in these experiments. When the flies of these species are kept at 16°C approximately a third of the matings are interspecific. In the experiment equal numbers of males and females of both species were placed in containers suitable for their survival and reproduction. The progeny of each generation were examined in order to determine if there were any interspecific hybrids. These hybrids were then eliminated. An equal number of males and females of the resulting progeny were then chosen to act as progenitors of the next generation. As the hybrids were destroyed in each generation the flies that solely mated with members of their own species produced more surviving descendants than the flies that mated solely with individuals of the other species. In the table to the right it can be seen that for each

generation the number of hybrids continuously decreased up to the tenth generation when hardly any interspecific hybrids were produced. It is evident that selection against the hybrids was very effective in increasing reproductive isolation between these species. From the third generation, the proportions of the hybrids were less than 5%. This confirmed that selection acts to reinforce the reproductive isolation of two genetically divergent populations if the hybrids formed by these species are less well adapted than their parents.

These discoveries allowed certain assumptions to be made regarding the origin of reproductive isolation mechanisms in nature. Namely, if selection reinforces the degree of reproductive isolation that exists between two species due to the poor adaptive value of the hybrids, it is expected that the populations of two species located in the same area will show a greater reproductive isolation than populations that are geographically separated. This mechanism for “reinforcing” hybridization barriers in sympatric populations is called the "Wallace Effect", as it was first proposed by Alfred Russell Wallace at the end of the 19th century, and it has been experimentally demonstrated in both plants and animals.

The sexual isolation between *Drosophila miranda* and *D. pseudoobscura*, for example, is more or less pronounced according to the geographic origin of the flies being studied. Flies from regions where the distribution of the species is superimposed show a greater sexual isolation than exists between populations originating in distant regions.



Reproductive isolation mechanisms can be a consequence of allopatric speciation. A population of *Drosophila* was divided into sub populations that were selected to adapt to different food types. After a number of generations the two sub populations were mixed again. It was observed that the subsequent matings occurred between individuals belonging to the same adapted group.

On the other hand, interspecific hybridization barriers can also arise as a result of the adaptive divergence that accompanies allopatric speciation. This mechanism has been experimentally proved by an experiment carried out by Diane Dodd on *D. pseudoobscura*. A single population of flies was divided into two, with one of the populations fed with starch-based food and the other with maltose-based food. This meant that each sub population was adapted to each food type over a number of generations. After the populations had diverged over many generations, the groups were again mixed; it was observed that the flies would mate only with others from their adapted population. This indicates that the mechanisms of reproductive isolation can arise even though the interspecific hybrids are not selected against.

GENE INTERACTIONS

Between 1884 (the year Mendel died) and 1888 details of mitosis and meiosis were reported, the cell nucleus was identified as the location of the genetic material, and "qualities" were even proposed to be transmitted on chromosomes to daughter cells at

mitosis. In 1903 Walter Sutton and Theodore Boveri formally proposed that chromosomes contain the genes. The Chromosome Theory of Inheritance is one of the foundations of genetics and explains the physical reality of Mendel's principles of inheritance.

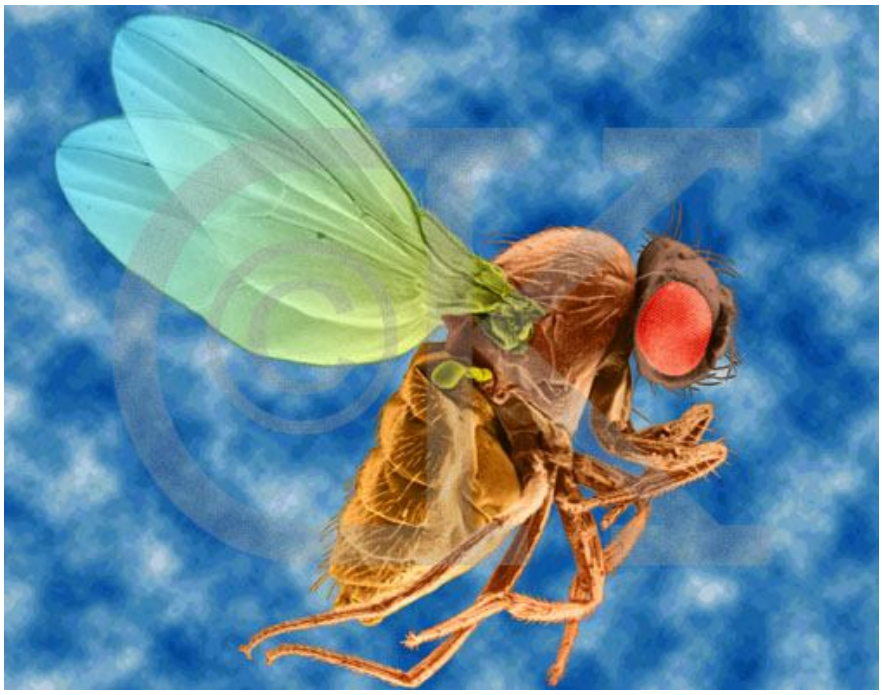


Fig. 7: Fruit Fly, *Drosophila melanogaster*, (SEM X60).

The location of many genes (Mendel's factors) was determined by Thomas Hunt Morgan and his coworkers in the early 1900's. Morgan's experimental organism was the fruit fly (*Drosophila melanogaster*). Fruit flies are ideal organisms for genetics, having a small size, ease of care, susceptibility to mutations, and short (7-9 day) generation time. The role of chromosomes in determination of sex was deduced by Morgan from work on fruit flies.

During Metaphase I, homologous chromosomes will line up. A karyotype can be made by cutting and arranging photomicrographs of the homologous chromosomes thus revealed at Metaphase I. Two types of chromosome pairs occur. Autosomes resemble each other in size and placement of the centromere, for example pairs of chromosome

21 are the same size, while pairs of chromosome 9 are of a different size from pair 21. Sex chromosomes may differ in their size, depending on the species of the organism they are from. In humans and *Drosophila*, males have a smaller sex chromosome, termed the Y, and a larger one, termed the X. Males are thus XY, and are termed heterogametic. Females are XX, and are termed homogametic. In grasshoppers, which Sutton studied in discovering chromosomes, there is no Y, only the X chromosome in males. Females are XX, while males are denoted as XO. Other organisms (notably birds, moths and butterflies) have males homogametic and females heterogametic. Males (if heterogametic) contribute either an X or Y to the offspring, while females contribute either X. The male thus determines the sex of the offspring. Remember that in meiosis, each chromosome is replicated and one copy sent to each gamete.

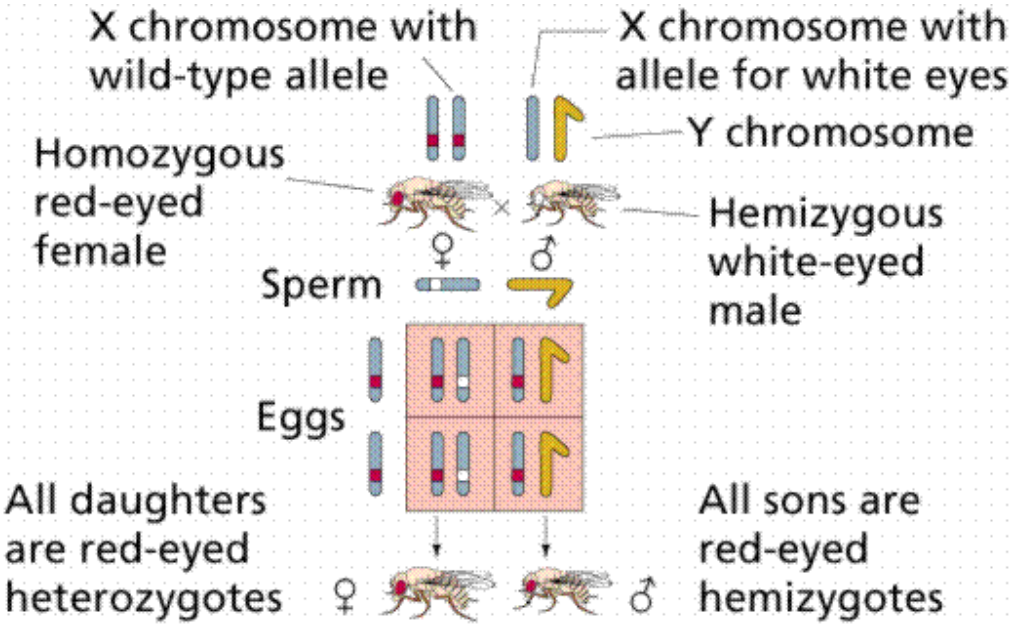
Morgan discovered a mutant eye color and attempted to use this mutant as a recessive to duplicate Mendel's results. He failed, instead of achieving a 3:1 F2 ratio the ratio was closer to 4:1 (red to white). Most mutations are usually recessive, thus the appearance of the white mutant presented Morgan a chance to test Mendel's ratios on animals. The F1 generation also had no white eyed females. Morgan hypothesized that the gene for eye color was only on the X chromosome, specifically in that region of the X that had no corresponding region on the Y. White eyed fruit flies were also more likely to die prior to adulthood, thus explaining the altered ratios. Normally eyes are red, but a variant (white) eyed was detected and used in genetic study. Cross a homozygous white eyed male with a homozygous red eyed female, and all the offspring have red eyes. Red is dominant over white. However, cross a homozygous white eyed female with a red eyed male, and the unexpected results show all the males have white eyes and all the females red eyes. This can be explained if the eye color gene is on the X chromosome.

Explanation

If the gene for eye color is on the X chromosome, the red eyed male in the second cross will pass his red eyed X to only his daughters, who in turn received only a recessive white-carrying X from their mother. Thus all females had red eyes like their

father. Since the male fruit fly passes only the Y to his sons, their eye color is determined entirely by the single X chromosome they receive from their mother (in this case white). Thus all the males in the second cross were white eyed.

These experiments introduced the concept of sex-linkage, the occurrence of genes on that part of the X that lack a corresponding location on the Y. Sex-linked recessives (such as white eyes in fruit flies, haemophilia, baldness, and colorblindness in humans) occur more commonly in males, since there is no chance of them being heterozygous. Such a condition is termed hemizygous.



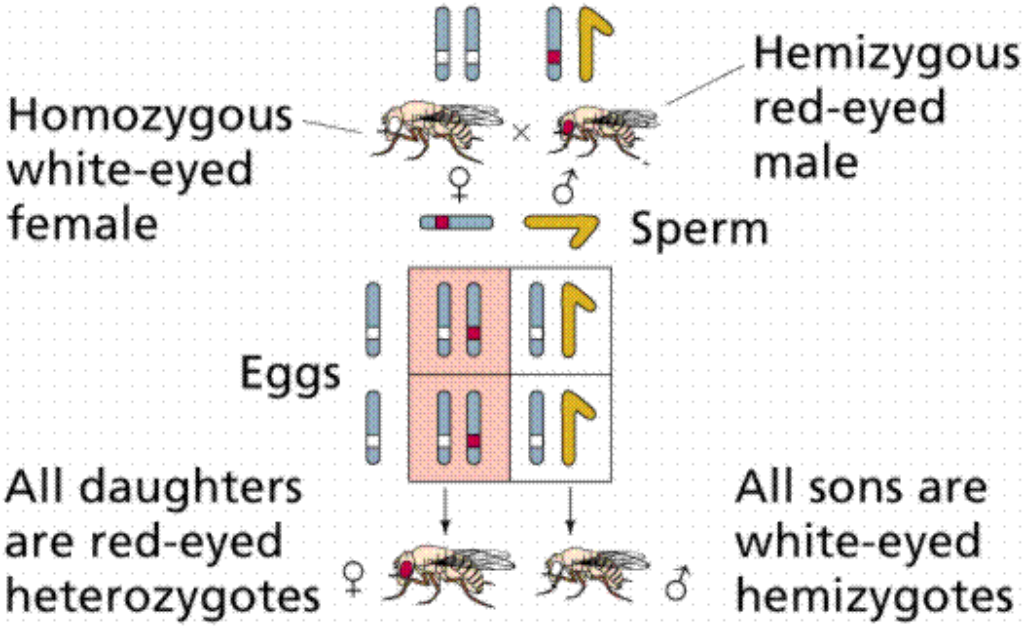


Fig. 8: Inheritance of eye color in fruit flies. Images from Purves *et al*., Life: The Science of Biology, 4th Edition, by Sinauer Associates and WH Freeman.

Characteristics of X-linked Traits

1. Phenotypic expression more common in males

2. Sons cannot inherit the trait from their fathers, but daughters can. Sons inherit their Y chromosome from their father. Only a few genes have been identified on the Y chromosome, among them the testis-determining factor (TDF) that promotes development of the male phenotype.

Barr bodies are interpreted as inactivated X chromosomes in mammalian females. Since females have two X chromosomes, the Lyon hypothesis suggests that one or the other X is inactivated in each somatic (non-reproductive) cell during embryonic

development. Cells mitotically produced from these embryonic cells likewise have the same inactivated X chromosome.

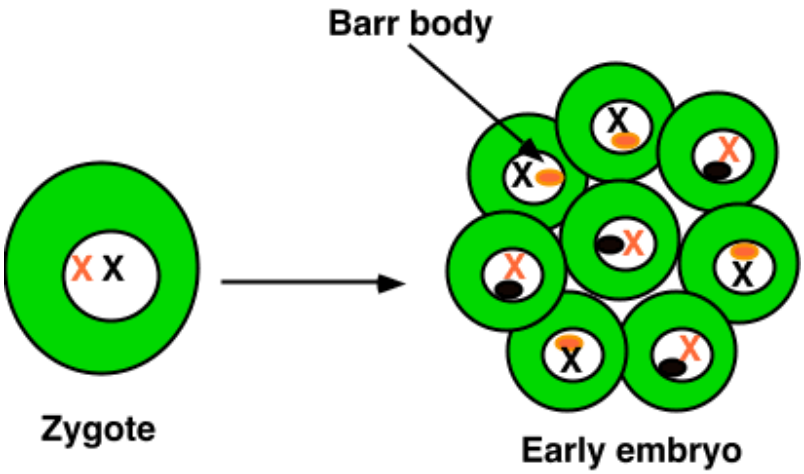


Fig. 7: The role of deactivated X chromosomes in mammalian female development.

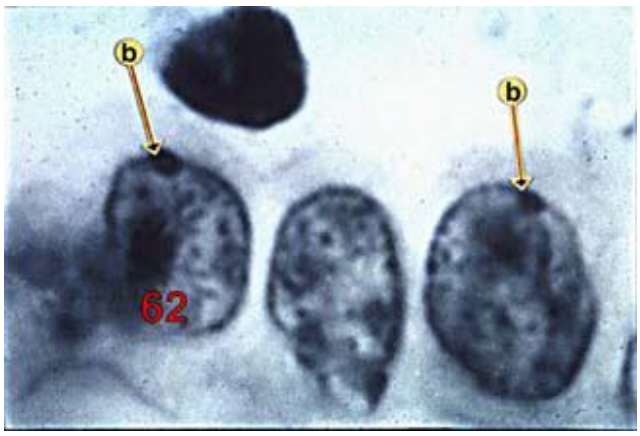


Fig. 8: Barr bodies (b) in mammalian cells.

Calico cats (sometimes called tortoiseshell) are almost always female since the calico trait is caused by some areas of the cat's fur expressing one allele and others expressing the other color. Can there be a male tortoiseshell cat? How would such a cat get its genes? Remember that fur color in cats is a sex-linked feature. Would the male calico be fertile or sterile?

The Modern View of the Gene

While Mendel discussed traits, we now know that genes are segments of the DNA that code for specific proteins. These proteins are responsible for the expression of the phenotype. The basic principles of segregation and independent assortment as worked out by Mendel are applicable even for sex-linked traits.

Codominant alleles

Codominant alleles occur when rather than expressing an intermediate phenotype, the heterozygotes express both homozygous phenotypes. An example is in human ABO blood types, the heterozygote AB type manufactures antibodies to both A and B types. Blood Type A people manufacture only anti-B antibodies, while type B people make only anti-A antibodies. Codominant alleles are both expressed. Heterozygotes for codominant alleles fully express both alleles. Blood type AB individuals produce both A and B antigens. Since neither A nor B is dominant over the other and they are both dominant over O they are said to be codominant.

Incomplete dominance

Incomplete dominance is a condition when neither allele is dominant over the other. The condition is recognized by the heterozygotes expressing an intermediate phenotype relative to the parental phenotypes. If a red flowered plant is crossed with a white flowered one, the progeny will all be pink. When pink is crossed with pink, the progeny are 1 red, 2 pink, and 1 white.

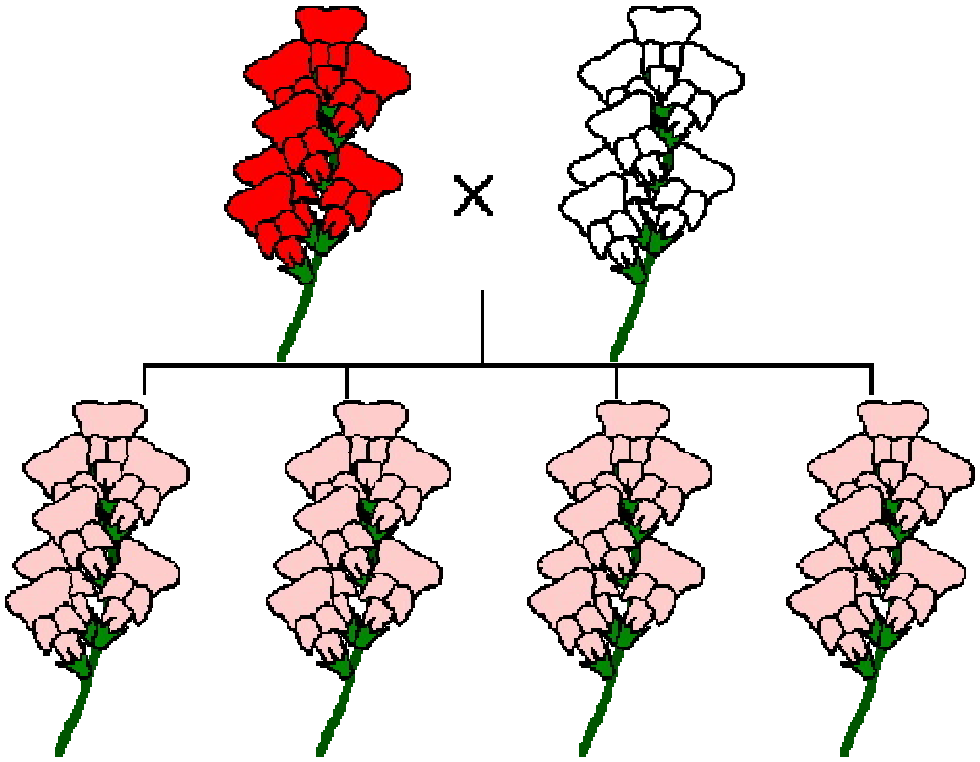


Fig. 8: Inheritance of flower color in snapdragons.

Flower color in snapdragons is an example of this pattern. Cross a true-breeding red strain with a true-breeding white strain and the F1 are all pink (heterozygotes). Self-fertilize the F1 and you get an F2 ratio of 1 red: 2 pink: 1 white. This would not happen if true blending had occurred (blending cannot explain traits such as red or white skipping a generation and pink flowers crossed with pink flowers should produce ONLY pink flowers).

Multiple alleles

Many genes have more than two alleles (even though any one diploid individual can only have at most two alleles for any gene), such as the ABO blood groups in humans, which are an example of multiple alleles. Multiple alleles result from different mutations of the same gene. Coat color in rabbits is determined by four alleles. Human ABO blood types are determined by alleles A, B, and O. A and B are codominants which are both

dominant over O. The only possible genotype for a type O person is OO. Type A people have either AA or AO genotypes. Type B people have either BB or BO genotypes. Type AB have only the AB (heterozygous) genotype. The A and B alleles of gene I produce slightly different glycoproteins (antigens) that are on the surface of each cell. Homozygous A individuals have only the A antigen, homozygous B individuals have only the B antigen, homozygous O individuals produce neither antigen, while a fourth phenotype (AB) produces both A and B antigens.

Interactions among genes

While one gene may make only one protein, the effects of those proteins usually interact (for example widow's peak may be masked by expression of the baldness gene). Novel phenotypes often result from the interactions of two genes, as in the case of the comb in chickens. The single comb is produced only by the rrpp genotype. Rose comb (b) results from R_pp. (_ can be either R or r). Pea comb (c) results from rrP_. Walnut comb, a novel phenotype, is produced when the genotype has at least one dominant of each gene (R_P_).

Epistasis

Epistasis is the term applied when one gene interferes with the expression of another (as in the baldness/widow's peak mentioned earlier). Bateson reported a different phenotypic ratio in sweet pea than could be explained by simple Mendelian inheritance. This ratio is 9:7 instead of the 9:3:3:1 one would expect of a dihybrid cross between heterozygotes. Of the two genes (C and P), when either is homozygous recessive (cc or pp) that gene is epistatic to (or hides) the other. To get purple flowers one must have both C and P alleles present.

Environment and Gene Expression

Phenotypes are always affected by their environment. In buttercup (*Ranunculus peltatus*), leaves below water-level are finely divided and those above water-level are

broad, floating, photosynthetic leaf-like leaves. Siamese cats are darker on their extremities, due to temperature effects on phenotypic expression. Expression of phenotype is a result of interaction between genes and environment. Siamese cats and Himalayan rabbits both animals have dark colored fur on their extremities. This is caused by an allele that controls pigment production being able only to function at the lower temperatures of those extremities. Environment determines the phenotypic pattern of expression.

Polygenic Inheritance

Polygenic inheritance is a pattern responsible for many features that seem simple on the surface. Many traits such as height, shape, weight, color, and metabolic rate are governed by the cumulative effects of many genes. Polygenic traits are not expressed as absolute or discrete characters, as was the case with Mendel's pea plant traits. Instead, polygenic traits are recognizable by their expression as a gradation of small differences (a continuous variation). The results form a bell shaped curve, with a mean value and extremes in either direction.

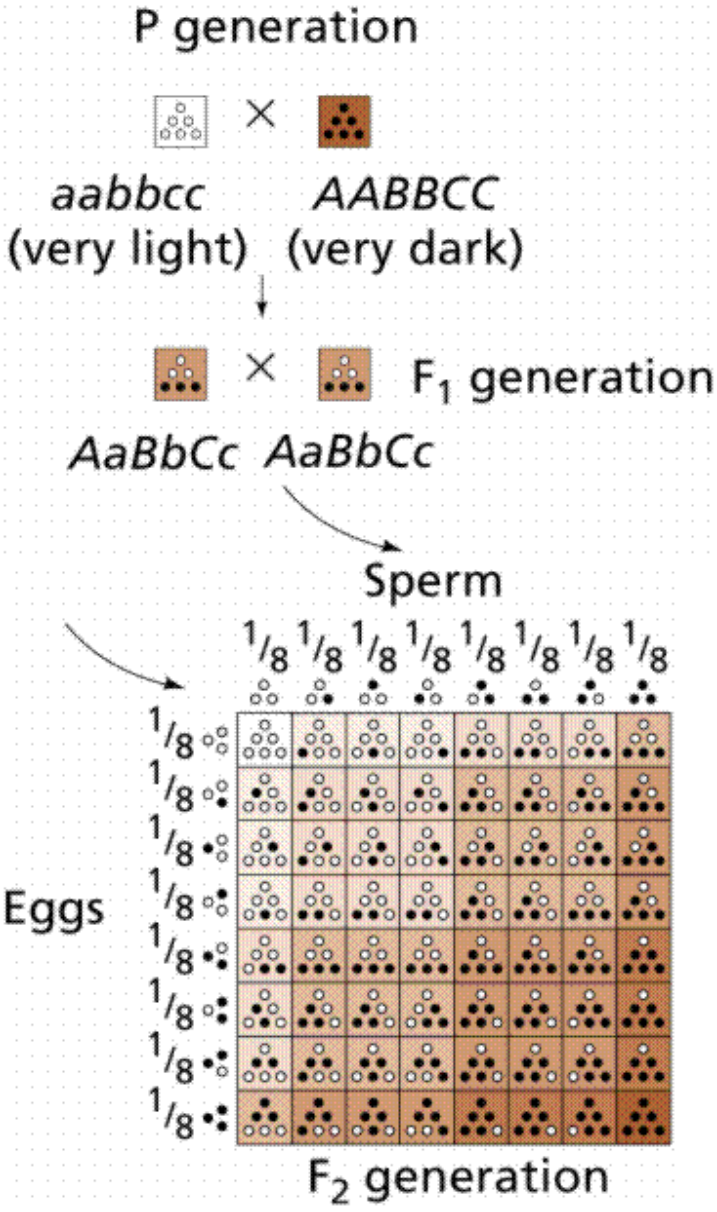
Height in humans is a polygenic trait, as is color in wheat kernels. Height in humans is not discontinuous. If you line up the entire class a continuum of variation is evident, with an average height and extremes in variation (very short [vertically challenged?] and very tall [vertically enhanced]). Traits showing continuous variation are usually controlled by the additive effects of two or more separate gene pairs. This is an example of polygenic inheritance. The inheritance of EACH gene follows Mendelian rules.

Usually polygenic traits are distinguished by:

2. Traits are usually quantified by measurement rather than counting.
3. Two or more gene pairs contribute to the phenotype.
4. Phenotypic expression of polygenic traits varies over a wide range.

Human polygenic traits include

1. Height
2. Weight
3. Eye Color
4. Intelligence
5. Skin Color
6. Many forms of behavior


P generation
aabbcc
(very light)
AABBCC
(very dark)
F1 generation
AaBbCc AaBbCc
Sperm
1/8 1/8 1/8 1/8 1/8 1/8 1/8 1/8
Eggs
1/8
1/8
1/8
1/8
1/8
1/8
1/8
1/8
F2 generation

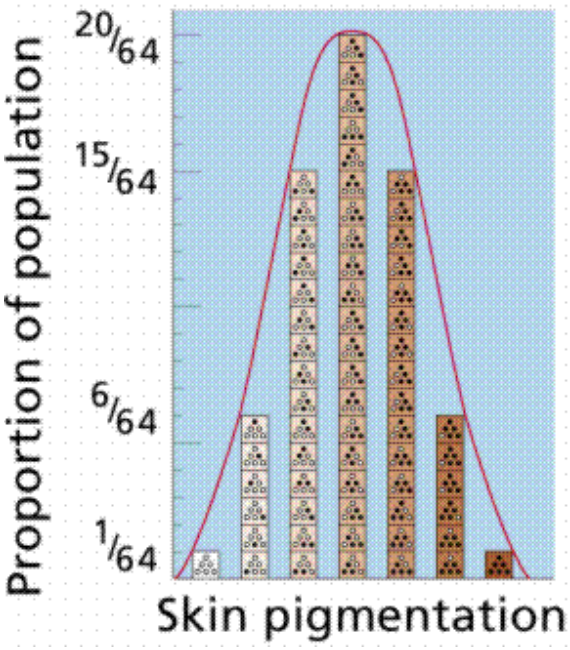


Fig. 9: The role of many genes (polygenic inheritance) in production of a continuum of phenotypes. From: *The Science of Biology*, 4th Edition

Pleiotropy

Pleiotropy is the effect of a single gene on more than one characteristic. An example is the "frizzle-trait" in chickens. The primary result of this gene is the production of defective feathers. Secondary results are both good and bad; good include increased adaptation to warm temperatures, bad include increased metabolic rate, decreased egg-laying, changes in heart, kidney and spleen. Cats that are white with blue eyes are often deaf, white cats with a blue and an yellow-orange eye are deaf on the side with the blue eye. Sickle-cell anaemia is a human disease originating in warm lowland tropical areas where malaria is common. Sickle-celled individuals suffer from a number of problems, all of which are pleiotropic effects of the sickle-cell allele.

Genes and Chromosomes

Linkage occurs when genes are on the same chromosome. Remember that sex-linked genes are on the X chromosome, one of the sex chromosomes. Linkage groups are

invariably the same number as the pairs of homologous chromosomes an organism possesses. Recombination occurs when crossing-over has broken linkage groups, as in the case of the genes for wing size and body color that Morgan studied. Chromosome mapping was originally based on the frequencies of recombination between alleles.

Since mutations can be induced (by radiation or chemicals), Morgan and his coworkers were able to cause new alleles to form by subjecting fruit flies to mutagens (agents of mutation, or mutation generators). Genes are located on specific regions of a certain chromosome, termed the gene locus (plural: loci). A gene therefore is a specific segment of the DNA molecule.

Alfred Sturtevant, while an undergraduate student in Morgan's lab, postulated that crossing-over would be less common between genes adjacent to each other on the same chromosome and that it should be possible to plot the sequence of genes along a fruit fly chromosome by using crossing-over frequencies. Distances on gene maps are expressed in map units (one map unit = 1 recombinant per 100 fertilized eggs; or a 1% chance of recombination).

The map for *Drosophila melanogaster* chromosomes is well known. Note that eye color and aristae length are far apart, as indicated by the occurrence of more recombinants (crossing-overs) between them, while wing length is closer to eye shape (as indicated by the low frequency of recombination between these two features).

REFERENCES

Baker H G (1959), "Reproductive methods as factors in speciation in flowering plants", *Cold Spring Harb Symp quant Biol* **24**: 177–191, PMID 13796002.

Grant V (1966), "The selective origin of incompatibility barriers in the plant genus

Fisher, Ronald. 1930. *The Genetical Theory of Natural Selection*

Postlethwalt, John (2009). *Modern Biology*. Holt, Rinehart and Winston. pp. 317.

Raven, P.H., Johnson, G.B. (1995). *Biology* (updated version). Third Edition. Wm. C. Brown publishers, Dubuque, Iowa. pp. 258-278.

Hartl, Daniel (2007). *Principles of Population Genetics*. Sinauer Associates. p. 95. ISBN 978-0-87893-308-2.

William B. Provine (1978). "The role of mathematical population geneticists in the evolutionary synthesis of the 1930s and 1940s". *Studies of the History of Biology* **1**: 167–192.

JBS Haldane (1927). "A Mathematical Theory of Natural and Artificial Selection, Part V: Selection and Mutation". *Mathematical Proceedings of the Cambridge Philosophical Society* **23**: 838–844

Richard C. Lewontin. The Genetic Basis of Evolutionary Change. Columbia University Press, 1974.

Stebbins G.L. 1950. *Variation and evolution in plants*. Chapter XII: The Karyotype. Columbia University Press N.Y.

White M.J.D. 1973. *The chromosomes*. 6th ed, Chapman & Hall, London. p28