KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University, Established Under Section 3 of UGC Act 1956)
Pollachi Main Road, Eachanari Post, Coimbatore - 641 021. INDIA
Phone : 0422-6471113-5, 6453777 Fax No : 0422 -2980022-3
Email : info@karpagam.com Web : www.kahedu.edu.in

DEPARTMENT OF BIOCHEMISTRY
SYLLABUS

SUBJECT NAME: CONCEPTS IN GENETICS
SUB.CODE: 17BCU304-B
SEMESTER: III **CLASS: II B.Sc., BIOCHEMISTRY**

SCOPE
This course covers genetics, the science of heredity, from its basic principles to the most recent advances in the field

OBJECTIVES
To demonstrate the basics of genetics and its recent advances enabling students to understand, analyze and solve problems in genetics.

Unit 1
Model organisms, Mendelism and chromosomal basis of heredity
Model organisms: *Escherichia coli*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Danio rerio* and *Arabidopsis thaliana*, Basic principles of heredity. Laws of probability & binomial expansion, formulating and testing genetic hypothesis, chromosomal basis of Mendelism -Sutton and Boveri hypothesis with experimental evidences.

Unit 2
Extensions of Mendelism, genetics of a gene, bacteria and viruses
Allelic variation and gene function- dominance relationships, multiple alleles, lethal alleles and null alleles. Pleiotropy gene interaction- epistatic and non-epistatic, interaction between gene(s) and environment. Penetrance and expressivity, norm of reaction and phenocopy. Complementation test, limitations of *cis-trans* test, intragenic complementation, rII locus of phage T4 and concept of cistron. Mechanism of genetic exchange - conjugation, transformation and transduction. Gene mapping in bacteria.

Unit 3
Genetics of eukaryotes and Human pedigree analysis
Linkage and crossing over, genetic mapping in eukaryotes, centromere mapping with ordered tetrads, cytogenetic mapping with deletions and duplications in *Drosophila*, detection of linked loci by pedigree analysis in humans and somatic cell hybridization for positioning genes on chromosomes. Pedigree conventions, characteristics of dominant and recessive inheritance. Applications of pedigree analysis.

Unit 4

Developmental genetics, epigenetics and chromosomal aberrations

Model organism for genetic analysis, *Drosophila* development, maternal effect genes, morphogens and zygotic gene activity in development, sex chromosomes and sex determination, dosage compensation of X-linked genes. Extra nuclear inheritance, tests for organelle heredity and maternal effect, epigenetic mechanisms of transcriptional regulation & genomic imprinting. Variations in chromosome number- monosomy and trisomy of sex and autosomes. Variations in chromosome structure- inversions, deletions, duplications and translocations.

Unit 5

Complex traits inheritance, population & evolutionary genetics

Inheritance of complex trait, analysis of quantitative traits, narrow and broad sense heritability, quantitative trait loci (QTL) and their identification. Hardy- Weinberg law, predicting allele and genotype frequencies and exceptions to Hardy-Weinberg principle. Molecular evolution - analysis of nucleotide and amino acid sequences, molecular phylogenies, homologous sequences, phenotypic evolution and speciation.

REFERENCES

1. Snustad, D.P., and Simmons, M.J., (2012). *Genetics.* (6th ed.), John Wiley & Sons. (Singapore), ISBN: 978-1-118-09242-2.
2. Pierce, B.A., (2012). *Genetics - A Conceptual Approach.* (4th ed.), W.H. Freeman & Co. (New York), ISBN:13:978-1-4292-7606-1 / ISBN:10:1-4292-7606-1.
3. Griffiths, A.J.F., Wessler, S. R., Carroll, S. B., and Doebley, J., (2010). *An Introduction to Genetic Analysis*. (10th ed.), W.H. Freeman & Company (New York), ISBN:10: 1-4292-2943-8.

TEXTBOOKS

1. Miglani, G.S., (2008), *Fundamentals of Genetics*. (2nd ed.), Narosa publications.
2. Manu, L.K., Mehra, L.A., Sadhana, S., and Choudary, R. (2009). *Essentials of Human Genetics*. (5th ed.). Universities Press.
3. Sambamurthy, A.V.S.S. (2009). *Genetics*. (2nd ed.), Narosa Publications.
4. Strickberger, M.W. (1985). *Genetics*. (3rd ed.) Mac Millian Publishing Co.,
5. Tamarin, R.H., (2004) *Principles of Genetics*. (7th ed.) Wm.C.Brown Publishers.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University Established Under Section 3 of UGC Act 1956)

Pollachi Main Road, Eachanari Post, Coimbatore - 641 021. INDIA

Phone : 0422-6471113-5, 6453777 Fax No : 0422 -2980022-3

Email : info@karpagam.com Web : www.kahedu.edu.in

DEPARTMENT OF BIOCHEMISTRY

LECTURE PLAN

SUBJECT NAME: CONCEPTS IN GENETICS

SUB.CODE: 16BCU304-B

SEMESTER: III **CLASS: II B.Sc., BIOCHEMISTRY**

S.No	Duration of Period	Topics to be Covered	Books referred with Page No.	Web page referred
UNIT I - MODEL ORGANISMS, MENDELISM AND CHROMOSOMAL BASIS OF HEREDITY				
1	1	Model organisms: *Escherichia coli, Saccharomyces cerevisiae, Caenorhabditis elegans, Danio rerio* and *Arabidopsis thaliana*	R1: 32-34	
2	1	Basic principles of heredity	R2: 45-49	
3	1	Laws of Probability, Binomial expansion	R1: 53-55	
4	1	Formulating and testing genetic hypothesis.	R1: 53-55	
5	1	Chromosomal basis of Mendelism - Sutton and Boveri hypothesis with experimental evidences	R2: 88-103	
6	1	Revision and Possible QP discussion		
Total No. of Hours planned for Unit I is 6 hours				
UNIT II – EXTENSIONS OF MENDELISM, GENETICS OF A GENE, BACTERIA AND VIRUSES				
1	1	Allelic variation and gene function - dominance relationships, multiple alleles, lethal alleles and null alleles	R2: 102-105	
2	1	Pleiotropy gene interaction - epistatic and non-epistatic, Interaction between gene(s) and environment.	T1: 8.4-8.16 R1: 33-36 R2: 106-116	
3	1	Penetrance and expressivity, norm of reaction and phenocopy.	R1: 154-158 T2: 36-37	
4	1	Complementation test, limitations of cis-trans test	R1: 322-327	
5	1	Intragenic complementation, rII locus of phage T4	R1: 205-	

		and concept of cistron	219, 326-336	
6	1	Mechanism of genetic exchange - conjugation, transformation and transduction	T3: 413-418 R1: 205-219	
7	1	Gene mapping in bacteria	T3: 413-418	
Total No. of Hours planned for Unit II is 7 hours				
UNIT III – GENETICS OF EUKARYOTES AND HUMAN PEDIGREE ANALYSIS				
1	1	Linkage and Crossing over, Genetic mapping in eukaryotes	R2: 159-168, 172-173 T4: 647-648, 762-763	
2	1	Centromere mapping with ordered tetrads, Cytogenetic mapping with deletions and duplications in Drosophila	R2: 172-187	
3	1	Detection of linked loci by pedigree analysis in humans	R1: 45-47, 571-572	
4	1	Somatic cell hybridization for positioning genes on chromosomes.	R1: 600-601 T2: 159-163	
5	1	Pedigree conventions, Characteristics of dominant and recessive inheritance	T4: 108-110, 207-208	
6	1	Applications of pedigree analysis	T4: 542-543, 712-714	
7	1	Revision and Possible QP discussion		
Total No. of Hours planned for Unit III is 7 hours				
UNIT IV – DEVELOPMENTAL GENETICS, EPIGENETICS AND CHROMOSOMAL ABERRATIONS				
1	1	Model organism for genetic analysis, Drosophila development, maternal effect genes	T4: 656-658	
2	1	Morphogens and zygotic gene activity in development	R1: 417-420	
3	1	Sex chromosomes and sex determination, dosage compensation of X-linked genes.	R1: 70-72, 78-82	
4	1	Extra nuclear inheritance	R1: 78-80	
5		Tests for organelle heredity and maternal effect	R1: 536-537	
6	1	Epigenetic mechanisms of transcriptional regulation & genomic imprinting	T4: 653-654	
7	1	Variations in chromosome number - monosomy and trisomy of sex and autosomes, Variations in	T1: 19.6-19.7, 19.2-	

		chromosome structure - inversions, deletions, duplications and translocations.	19.6	
Total No. of Hours planned for Unit IV is 7 hours				
UNIT V – COMPLEX TRAITS INHERITANCE, POPULATION & EVOLUTIONARY GENETICS				
1	1	Inheritance of complex trait, analysis of quantitative traits, Narrow and broad sense heritability	T5: 526-527 T4: 265-266	
2	1	Quantitative trait loci (QTL) and their identification	T5: 516-517, 661-662	
3	1	Hardy-Weinberg law, Predicting allele and genotype frequencies and exceptions to Hardy-Weinberg principle	R2: 673-676	
4	1	Molecular evolution - analysis of nucleotide and amino acid sequences	T5: 581-583	
5	1	Molecular phylogenesis	T5: 581-583	
6	1	Homologous sequences, phenotypic evolution and speciation	T5: 570-575	
7	1	Revision and Possible QP discussion		
Total No. of Hours planned for Unit V is 7 hours				

REFERENCES:

R1. Snustad, D.P., and Simmons, M.J., (2012). Genetics 6th ed., John Wiley & Sons. ISBN: 978-1-118-0942-2.

R2. Pierce, B.A., (2012). Genetics - A Conceptual Approach. 4th ed., W.H. Freeman & Co. ISBN: 13:978-1-4292-7606-1/ ISBN:10:1-4292-7606-1.

R3. Griffiths, A.F.J., Wessler, S.R., Carroll, S.B., and Doebley, J., (2010). An Introduction to Genetic Analysis 10th ed., W.H. Freeman & Company, ISBN:10: 1-4292-2943-8.

TEXTBOOKS:

T1. Miglani, G.S., (2008), *Fundamentals of Genetics*. (2nd ed.), Narosa publications.

T2. Manu, L.K., Mehra, L.A., Sadhana, S., and Choudary, R. (2009). *Essentials of Human Genetics*. (5th ed.). Universities Press.

T3. Sambamurthy, A.V.S.S. (2009). *Genetics*. (2nd ed.), Narosa Publications.

T4. Strickberger, M.W. (1985). *Genetics*. (3rd ed.) Mac Millian Publishing Co.,

T5. Tamarin, R.H., (2004) *Principles of Genetics*. (7th ed.) Wm.C.Brown Publishers.

CLASS: II B.Sc., BIOCHEMISTRY **COURSE NAME: CONCEPTS IN GENETICS**
COURSE CODE: 17BCU304B **UNIT:I-MODEL ORGANISMS, MENDELISM AND CHROMOSOMAL BASIS OF HEREDITY (BATCH-2017-2020)**

UNIT 1
SYLLABUS

Model organisms, Mendelism and chromosomal basis of heredity
Model organisms: *Escherichia coli*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Danio rerio* and *Arabidopsis thaliana*, Basic principles of heredity. Laws of probability & binomial expansion, formulating and testing genetic hypothesis, chromosomal basis of Mendelism -Sutton and Boveri hypothesis with experimental evidences.

Model organisms

are those that useful data sets have been already gathered to describe basic biological processes.
are more amenable to asking certain questions due to their simplicity of structure and features.

Characteristics of Model Organisms for Research and Instruction

A model system is a simpler, idealized system that can be accessible and easily manipulated. Therefore, when selecting living organisms as models to work with, certain criteria are used depending upon the experimental purposes. As a result, there is a wide range of characteristics common to model organisms, including: 1) rapid development with short life cycles, 2) small adult size, 3) ready availability, and 4) tractability. Being small, growing rapidly and being readily available are crucial in terms of housing them, given the budget and space limitations of research and teaching laboratories. Tractability relates to the ease with which they can be manipulated.

Escherichia coli

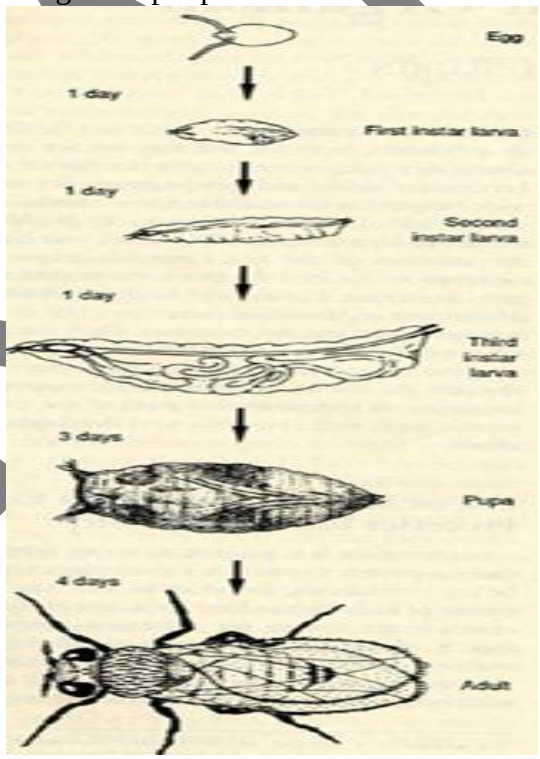
Escherichia coli, a prokaryotic organism without a nuclear membrane, is a representative living material often used in laboratories and classrooms. E. coli reproduces rapidly (under optimal situation 0.5 hr/generation) such that results for a number of experiments can be quickly obtained. Certain mutants of E. coli have been defined that cannot express certain proteins at saturation growth, and, therefore, die. E. coli was also the organism used to elucidate the regulation of the lac operon in genetics. E. coli's ability to take up exogenous genetic material under the procedure known as DNA-mediated cell transformation has also made it a popular model for studies using recombinant DNA. Using recombinant DNA techniques, E. coli can be manipulated in research laboratories and in the classroom to produce any DNA, RNA or protein of interest. Also, it is easy to manipulate both genetically and biochemically. Most importantly, it shares fundamental characteristics, such as DNA and messenger RNA, with all other organisms. The value of E. coli in recombinant DNA makes it a good model organism for students to study the genetic material.

Caenorhabditis elegans

In the last two decades, a nematode, Caenorhabditis elegans, has captured the hearts of developmental biologists and geneticists hoping to solve the enigma of cell development and related biological problems, such as aging. In addition, educators have begun to use it in

classrooms to illustrate central biological concepts, such as cell division. Its popularity as a model organism is because it is transparent, thus cells of interest can be observed using a dissecting microscope. It is small (about 1- 1.5 mm) and easy to cultivate, which makes it possible to house large numbers of C. elegans. It has a short life cycle (3 days), which makes the production of numerous generations possible. It can be crossed at will. Male and hermaphrodites are the two sexes. Hermaphrodites can self fertilize or mate with males to produce offspring. Thus, cross or self-fertilization can be manipulated as desired. There are numerous tools available to study C. elegans, including different types of antibodies and advanced microscopes. Its genome has been completely sequenced. This is quite attractive and useful in genetic studies allowing researchers to pick a gene of interest to study. In sum, as discussed, some organisms are easy and amenable for use in the laboratory and the classroom to enhance our understanding of human biology. They are called "model organisms". A model organism is one that possesses the virtues of tractability and accessibility that can be used in experimental manipulation both in school and research.

C. elegans is a popular research organism as it possesses all the characteristics mentioned, yet shares many essential biological properties with humans. For instance, researchers who study apoptosis (programmed cell death) use C. elegans as an experimental organism in the hope of finding treatments for certain types of human cancers, such as leukemia. By studying apoptosis in C. elegans, researchers hope to identify genes that switch-on cell death in cancer cells, thus, using the cell's own genetic machinery to rid the body of malignant cells. Because leukemia is the unregulated growth of white blood cells, identifying genes involved in apoptosis may provide researchers with a tool for treating the rapid proliferation of cancer cells.



In general, scientists have to work with organisms different from the ones they wish to apply their findings to for several reasons. Model organisms act as surrogates that enable experiments to be carried out under a more favorable environment than would be available in the original system.
The biological insights gained from using model organisms have helped to cure human diseases and improve people's understanding of life. Moreover, by studying organisms unrelated to humans, insight into scientific concepts can sometimes be more easily achieved. One example was the discovery of giant chromosomes in Drosophila's salivary glands that improved researchers' understanding of genetics.
Compared to general living organisms, model organisms are well-established experimental systems and are often ready to be used in classrooms. The available resources, such as experimental protocols, for these model organisms make the transition of their use to curricula relatively simple. Thus, many are amenable to be used as experimental organisms in teaching.
The fruit fly, Drosophila melanogaster, is the most obvious organism used in teaching. It is often used for students to learn Mendelian genetics. E. coli, yeast, and Rana pipiens are also popular model organisms in biology education. Within the past two decades, due to its importance and popular use in research, C. elegans has begun to take its place in the classroom as an important model organism. Below, I describe some of the model organisms, and the reasons that make them popular for classroom use.

Danio rerio - Zebrafish

Zebrafish have only been used as a model organism by a lot of labs in the last 15 years or so. In the 1970s, **George Streisinger**, a scientist at the University of Oregon, was interested in using a vertebrate model organism that was simpler than the mouse and easy to manipulate genetically.
As a tropical fish enthusiast, he chose the zebrafish, which is commonly found in pet shops and home aquaria. Streisinger's colleagues in Oregon, amongst them Chuck Kimmel, were impressed with the ease of using the fish and for Kimmel the embryo was particularly attractive for studying nervous system development. Since the mid 1990s, many more scientists have started using zebrafish as a model organism to study developmental biology. Currently there are at least 600 laboratories around the world that use zebrafish. A big step forward in the zebrafish field came when two large genetic searches for mutants were carried out in the early 1990s. One was led by Nobel prize winner Christiane Nüsslein-Volhard in Tübingen, Germany, and the other by Wolfgang Driever and Mark Fishman in Boston, USA. The identification of mutants is one of the most important strategies for the study of many areas of biology. A mutation often provides the first insight into the role of the gene in normal development or in a healthy individual. The screens looked at thousands of families of fish to find mutations that affected early development. In the end around 4000 mutants were identified, and the results were published in a single, mammoth issue of the journal Development in December 1996. Scientists have spent many years since then working with the mutants that came out of the screens. And many more screens have been carried out in individual labs since then. However, finding a mutant is just the first part of the story: you can see the effect of the mutation on the zebrafish embryo, but you do not know what gene is affected. Zebrafish researchers needed more genetic resources in order to understand and

characterise the mutants further. The community got together to initiate the sequencing of the whole zebrafish genome, which began in 2001 at the Wellcome Trust Sanger Institute in Cambridge. To date over 14,000 genes have been described in the zebrafish genome, and although sequencing is not fully complete, we have a working draft that has proved immensely useful to scientists around the world.

Arabidopsis thaliana

Arabidopsis is an angiosperm, a dicot from the mustard family (Brassicaceae). It is popularly known as thale cress or mouse-ear cress. While it has no commercial value — in fact is considered a weed — it has proved to be an ideal organism for studying plant development.
Some of its advantages as a model organism:

- It has one of the smallest genomes in the plant kingdom: 135 x 10^6 base pairs of DNA distributed in 5 chromosomes (2n = 10) and almost all of which encodes its 27,407 genes.
- Transgenic plants can be made easily using *Agrobacterium tumefaciens* as the vector to introduce foreign genes.
- The plant is small — a flat rosette of leaves from which grows a flower stalk 6–12 inches high.
- It can be easily grown in the lab in a relatively small space.
- Development is rapid. It only takes 5– 6 weeks from seed germination to the production of a new crop of seeds.
- It is a prolific producer of seeds (up to 10,000 per plant) making genetics studies easier.
- Mutations can be easily generated (e.g., by irradiating the seeds or treating them with mutagenic chemicals).
- It is normally self-pollinated so recessive mutations quickly become homozygous and thus expressed.
- Other members of its family cannot self-pollinate. They have an active system of **self-incompatibility**. Arabidopsis, however, has inactivating mutations in the genes — ***SRK*** and ***SCR*** — that prevent self-pollination in other members of the family.
- However, Arabidopsis can easily be cross-pollinated to
    - do genetic mapping and
    - produce strains with multiple mutations.

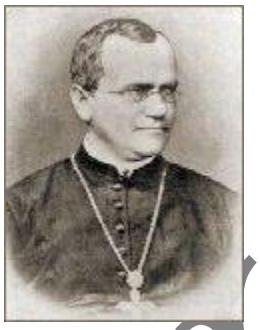
Drosophila melanogaster

The fruit fly, Drosophila melanogaster, has been the most popular eukaryotic organism used in classrooms. It has been used in heredity and biomedical research where the aims are to understand human genetics and developmental processes. It is also a popular model for teaching Mendelian genetics. Drosophila is very popular and successful as a model organism because it has short life cycle of two weeks, making it possible to study numerous generations in an academic year. It is easy to culture and inexpensive to house large numbers. Its size is amenable for cultivation in school laboratories. Also, it is large enough that many attributes can be seen with the naked eye or

under low-power magnification. Moreover, it has a very long history in biological research (since the early 1900s) and there are many useful tools to facilitate genetic study. For example, the use of antibodies makes the scoring of specific cells or cell types possible in Drosophila. Because of the above properties, Drosophila has been used in research and teaching in a great variety of disciplines, such as classical and molecular genetics. Some researchers use Drosophila to study how its body plan is controlled by a set of homeotic genes. The more research is done using Drosophila, the better we understand it, thus making it an attractive model organism for class use.

Basic Principles of Heredity

For thousands of years farmers and herders have been selectively breeding their plants and animals to produce more useful hybrids . It was somewhat of a hit or miss process since the actual mechanisms governing inheritance were unknown. Knowledge of these genetic mechanisms finally came as a result of careful laboratory breeding experiments carried out over the last century and a half.



Gregor Mendel
1822-1884

By the 1890's, the invention of better microscopes allowed biologists to discover the basic facts of cell division and sexual reproduction. The focus of genetics research then shifted to understanding what really happens in the transmission of hereditary traits from parents to children. A number of hypotheses were suggested to explain heredity, but Gregor Mendel , a little known Central European monk, was the only one who got it more or less right. His ideas had been published in 1866 but largely went unrecognized until 1900, which was long after his death. His early adult life was spent in relative obscurity doing basic genetics research and teaching high school mathematics, physics, and Greek in Brno (now in the Czech Republic). In his later years, he became the abbot of his monastery and put aside his scientific work.

While Mendel's research was with plants, the basic underlying principles of heredity that he discovered also apply to people and other animals because the mechanisms of heredity are essentially the same for all complex life forms.

Common edible peas

Through the selective cross-breeding of common pea plants (*Pisum sativum*) over many generations, Mendel discovered that certain traits show up in offspring without any blending of parent characteristics. For instance, the pea flowers are either purple or white--intermediate colors do not appear in the offspring of cross-pollinated pea plants. Mendel observed seven traits that are easily recognized and apparently only occur in one of two forms:

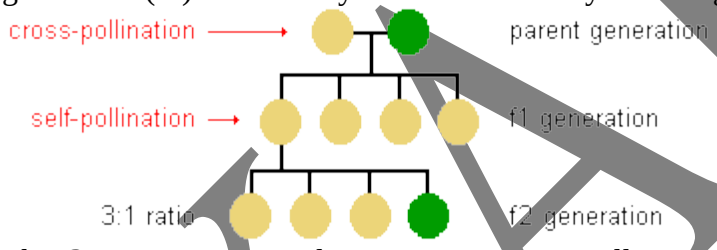
1. flower color is purple or white
2. flower position is axil or terminal
3. stem length is long or short
5. seed color is yellow or green
6. pod shape is inflated or constricted
7. pod color is yellow or green

4. seed shape is round or wrinkled

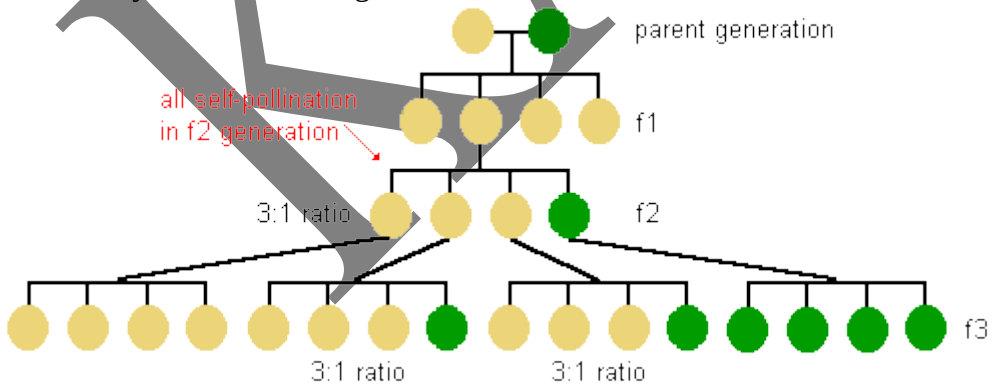
This observation that these traits do not show up in offspring plants with intermediate forms was critically important because the leading theory in biology at the time was that inherited traits blend from generation to generation. Most of the leading scientists in the 19th century accepted this "blending theory." Charles Darwin proposed another equally wrong theory known as "pangenesis". This held that hereditary "particles" in our bodies are affected by the things we do during our lifetime. These modified particles were thought to migrate via blood to the reproductive cells and subsequently could be inherited by the next generation. This was essentially a variation of Lamarck's incorrect idea of the "inheritance of acquired characteristics."

Mendel picked common garden pea plants for the focus of his research because they can be grown easily in large numbers and their reproduction can be manipulated. Pea plants have both male and female reproductive organs. As a result, they can either self-pollinate themselves or cross-pollinate with another plant. In his experiments, Mendel was able to selectively cross-pollinate purebred plants with particular traits and observe the outcome over many generations. This was the basis for his conclusions about the nature of genetic inheritance.

In cross-pollinating plants that either produce yellow or green pea seeds exclusively, Mendel found that the first offspring generation (f1) always has yellow seeds. However, the following generation (f2) consistently has a 3:1 ratio of yellow to green.



This 3:1 ratio occurs in later generations as well. Mendel realized that this underlying regularity was the key to understanding the basic mechanisms of inheritance.



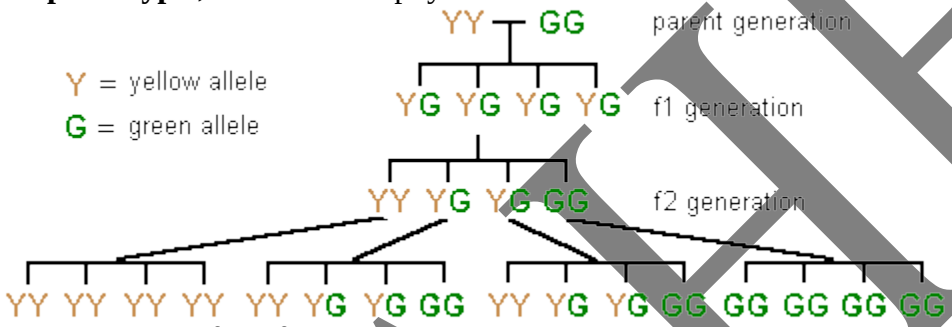
He came to three important conclusions from these experimental results:

1. that the inheritance of each trait is determined by "units" or "factors" that are passed on to descendents unchanged (these units are now

called genes)

2. that an individual inherits one such unit from each parent for each trait
3. that a trait may not show up in an individual but can still be passed on to the next generation.

It is important to realize that, in this experiment, the starting parent plants were **homozygous** for pea seed color. That is to say, they each had two identical forms (or **alleles**) of the gene for this trait--2 yellows or 2 greens. The plants in the f1 generation were all **heterozygous** . In other words, they each had inherited two different alleles--one from each parent plant. It becomes clearer when we look at the actual genetic makeup, or **genotype** , of the pea plants instead of only the **phenotype** , or observable physical characteristics.



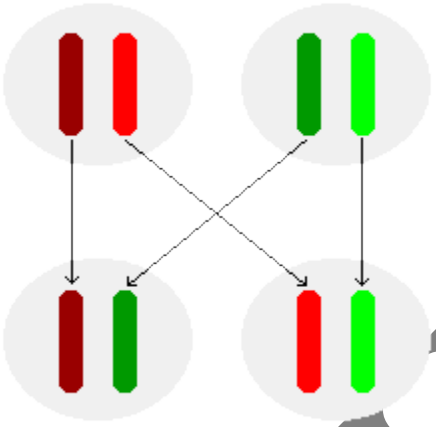
Note that each of the f1 generation plants (shown above) inherited a Y allele from one parent and a G allele from the other. When the f1 plants breed, each has an equal chance of passing on either Y or G alleles to each offspring.

With all of the seven pea plant traits that Mendel examined, one form appeared **dominant** over the other, which is to say it masked the presence of the other allele. For example, when the genotype for pea seed color is YG (heterozygous), the phenotype is yellow. However, the dominant yellow allele does not alter the **recessive** green one in any way. Both alleles can be passed on to the next generation unchanged.

Mendel's observations from these experiments can be summarized in two principles:

1. the principle of segregation
2. the principle of independent assortment

According to the **principle of segregation**, for any particular trait, the pair of alleles of each parent separate and only one allele passes from each parent on to an offspring. Which allele in a parent's pair of alleles is inherited is a matter of chance. We now know that this segregation of alleles occurs during the process of sex cell formation (i.e., meiosis).



Segregation of alleles in the production of sex cells

According to the **principle of independent assortment**, different pairs of alleles are passed to offspring independently of each other. The result is that new combinations of genes present in neither parent are possible. For example, a pea plant's inheritance of the ability to produce purple flowers instead of white ones does not make it more likely that it will also inherit the ability to produce yellow pea seeds in contrast to green ones. Likewise, the principle of independent assortment explains why the human inheritance of a particular eye color does not increase or decrease the likelihood of having 6 fingers on each hand. Today, we know this is due to the fact that the genes for independently assorted traits are located on different chromosomes.

These two principles of inheritance, along with the understanding of unit inheritance and dominance, were the beginnings of our modern science of genetics. However, Mendel did not realize that there are exceptions to these rules.

Laws of Probability & Binomial Expansion

Introduction to Probability:

The numbers of individuals in each ratio result from chance segregation of genes during gamete formation, and their chance combinations to form zygotes. Since these are chance events, accurate predictions about the results cannot be made.

This is especially true in cases where the progeny is limited to a small number such as in experiments in animal breeding and even more so in human pedigree studies. It is hardly possible to predict the appearance of a certain phenotype or genotype. But what we can say is that there is a certain probability of occurrence of a given genetic event.

In a general way we can say that the probability or chance that an event will occur can be defined as the proportion of times in which that event occurs in a very large number of trials.

If there are x trials and the event occurs on the average y times during the x trials, then the probability is expressed as y/x. The probability would be between zero and one. The closer the probability is to zero, the less chance is there for the event to occur. When probability is one, the event is certain to occur.

Rules of Probability:

1. The Multiplication Rule:

While considering probabilities it is important to note that the inheritance of certain genes such as for height of pea plants (tall or dwarf), flower colour (red or white), seed texture (round or wrinkled) are independent events. If we consider each gene separately, the probability of any one F2 plant being tall is 3/4, and of its being dwarf 1/4.

Similarly there is a probability of 3/4 that an F2 plant will bear red flowers and 1/4 that the flower will be white. Now what is the probability that an F2 plant would be both tall and coloured? Assuming that the inheritance of each gene is an independent event, the probability that a plant be coloured and tall is equal to the product of their individual probabilities, i.e., 3/4 x 3/4 = 9/16.

Thus there are 9 chances out of 16 that an F2 plant be coloured and tall. Stated in a general way, when the probability of an event is independent of that of another event, and the occurrence of one does not influence the occurrence of the other, probability that both events will occur together is the product of their individual probabilities.

2. The Addition Rule:

The probability that one of several mutually exclusive events will occur is the sum of their individual probabilities. This law is applicable when different types of events cannot occur together. If one occurs the other is excluded. When a coin is tossed there are two alternatives, either heads or tails will appear. If the probability for heads is 1/2, for tails 1/2, then the probability that either heads or tails will appear is 1/2 + 1/2 = 1.

Probability and Human Genetics:

In some recessive genetic traits in humans such as phenylketonuria (PKU), albinism and others, the birth of an affected child indicates that the parents are heterozygous carries for that trait. Due to the recessive nature of the gene the parents are healthy and normal. Such parents could be anxious to know what chances exist for any of their future children to show the genetic defect.

Since the trait is recessive they could expect affected children and normal children in the ratio 1:3. But since human families are limited in size, this ratio can be misleading. All we can say is that the chance or probability of having an affected child is 1/4 at each birth.

Moreover, even after the birth of an affected child, the same probability (1/4) exists for an affected child in all future pregnancies. The probability for a normal child at each birth is 3/4.

The rules of probability can be applied for predicting the ratio of boys and girls born in a family. Since the human male produces an equal number of X and Y sperm, the chance for a boy at any birth is 1/2, and for a girl also is 1/2. From the probability of each single conception it is possible to calculate the probability of successive births together.

For example, what is the probability that the first two children born in a family will both be males. To find this out we must determine the product of the separate probabilities at each conception, that is 1/2 x 1/2 = 1/4. Now consider a different question.

What is the probability that the third child in the family in which the first two are males, will also be a male? For the answer we must remember that the sex of any child is independent of the sex of the other children; therefore the probability for the third child to be male is 1/2.

Binomial Expansions of Probability:

There are many situations in genetics where we would like to know the probability that a combination of events will occur. For instance it may be required to determine the probability that two off-springs of a mating between Aa and aa parents will have a particular genetic constitution, namely 2 Aa, 2 aa, and 1 Aa and the other aa.

The occurrence of a particular genotype in a single offspring is an independent event, as it is not influenced by the genotype of any other offspring. The probability that 2 Aa offspring will be formed from this mating is therefore equal to the product of their separate probabilities.

Aa = 1/2 x 1/2 = 1/4 or 25%

Thus the probabilities for each sequence of two off-springs are as follows:

1st offspring	2nd offspring	Probabilities
Aa	*Aa*	1/2 × 1/2 = 1/4
AA	*aa*	1/2 × 1/2 = 1/4
aa	*Aa*	1/2 × 1/2 = 1/4
aa	*aa*	1/2 × 1/2 = 1/4

Thus the probability that both offspring are Aa is 1/4, that one is Aa the other aa is 2/4, that both are aa is 1/4. In other words the pattern for this distribution is 1: 2: 1.

This also represents the coefficients of raising two values of the binomial p and q to the power of 2:

$(p + q)2 = 1p2 + 2pq + 1q2$

or if we substitute Aa for p and aa for q

[(Aa) + (aa)]2 = 1(Aa) (Aa) + 2(Aa) (aa) + 1(aa) (aa)

Note that the probability for a complementary event such as the probability that both offspring will not be aa is 1 – the probability of the particular event, or 1 – 1/4 = 3/4. If the probabilities are calculated for the different combinations of genotypes possible among 3 children of the mating Aa x aa, the frequencies of each combination will be found to correspond with raising a binomial to the third power.

Probability that 3 offspring are Aa – 1/8

Probability that two are Aa and one aa = 3/8

Probability that two are aa and one Aa = 3/8

Probability that 3 offspring are aa = 1/8

or $(p + q)3 = 1p3 + 3p2q + 3pq2 + 1q3$

or [(Aa) + (aa)]3 = 1(Aa) (Aa) (Aa) + 3(Aa) (Aa) (aa) + 3(Aa) (aa) (aa) + 1(aa) (aa) (aa)

The probability for each particular combination of offspring can therefore be determined by the binomial coefficient for that combination relative to the total number of possible combinations.

In general we can say that when p is the probability that a particular event will occur and q or 1 – p is the probability of an alternative form of that event so that p + q = 1, then probability for each combination in which a succession of such events may occur is described by the binomial distribution.

Multinomial Distributions of Probability:

In the examples described above there are only two possible alternatives for each event. Sometimes a genetic cross or mating can produce three types of offspring. For instance Aa x Aa will produce AA, Aa and aa offspring in the ratio 1: 2: 1 so that their probabilities are 1/4 for AA, 1/2 for Aa and 1/4 for aa.

In such cases an additional term is added to the binomial to represent the third class of offspring. We now have a trinomial distribution (p + q + r)n, where p, q and r represent probabilities of AA, Aa and aa respectively.

For determining probabilities of trinomial combinations we use the formula:

$$\frac{n}{w \times y}\, p^w\, q^x\, r^y$$

where w, x and y are the numbers of offspring of each of the three different types, and p, q and r are their probabilities respectively. In a mating of Aa x Aa where only four offspring are produced, the probability of having exactly 1 AA homozygote, 2 Aa heterozygote and 1 aa homozygote would be

$$\frac{4}{1.2.1}(1/4)^1(1/2)^2(1/4)^1 = (24/2)\,(1/64) = 3/16$$

Other multinomial distributions can similarly be worked out. For example when a dihybrid cross is made in which four phenotypes can appear with frequencies p, q, r and s, the formula would be $\frac{n}{w \times yz} p^w q^x r^y s^z$.

As an illustration, consider the cross between smooth and yellow pea plants with wrinkled green plants.

When the F_1 heterozygote is self-fertilized, 19 offspring are obtained of which 8 are smooth yellow (probability 9/16), 5 are smooth green (probability 3/16), 4 wrinkled yellow (probability 3/16) and 2 wrinkled green (probability 1/16). The probability of having the 19 offspring in exactly this ratio (9:3:3:1) would be $\frac{8}{8.5.4.2}(9/16)^8\,(3/16)^5(3/16)^2 = 0.000574.$

Formulating and testing genetic hypothesis

Each cell in living organisms contains DNA, which is made of nucleotide subunits arranged in very long strands. By winding around structural proteins, the strands become condensed into compact units called chromosomes. Regions in the DNA, known as genes, carry specific instructions for making proteins. Genes represent unique combinations of nucleotides that provide information for the genetically-based traits that organisms display. While all humans have the same genes (that is, each human has a gene that codes for eye color, etc), humans are not genetically identical. Think of all the variation in the human population. We have different colors of hair, eyes, different heights, different tendencies towards disease, etc. One reason for the variation we see is that organisms, like humans, might have different forms of genes, which we term alleles. These alleles code for different proteins, and therefore result in the expression of different phenotypes. A dominant allele is

expressed phenotypically when present on either a single chromosome or on both homologous chromosomes (that is, when it's present on either the chromosome inherited from the mother, or the father, or both). A recessive allele is masked by a dominant allele and is expressed only when paired with another recessive allele on the homologous chromosome—that is, it is only expressed when the chromosomes inherited from both parents have the recessive allele. A pair of identical alleles (AA or aa) at a genes locus represents the homozygous condition. Two different alleles (Aa) at a genes locus represent the heterozygous condition.

For example, eye color in humans is determined by a single gene locus (although other loci can modify its effects). If the alleles at that locus are homozygous dominant (AA) or heterozygous (Aa), the eyes will be brown. If the alleles are homozygous recessive (aa), the eyes will be blue. Alleles for brown eyes are said to be dominant over alleles for blue eyes.

Morgan's forward cross was a red-eyed (wild-type) female with a white-eyed (mutant) male. For the reverse cross, Morgan used different flies with opposite alleles, in this case, a white-eyed (mutant) female and a red eyed (wild-type) male. The patterns of alleles in the offspring of these reciprocal crosses were markedly different and as a result, Morgan was able to determine that the white eye-color allele in fruit flies was sex-linked. In both Mendel's and Morgan's (who did experiments with fruit flies) experiments, it was important to begin with parents that are said to be true breeding. This means that the parents are homozygous for the alleles of interest for the study. Because the genetic makeup is known for the starting organisms, it enables scientists to track the genotypes (or the genetic makeup of the organisms) through generations All concepts about genetic information, the process of meiosis and experiments used to test for different patterns of inheritance are identical to what was presented in the Drosophila. It is important to remember that all parental virtual flies are true breeding and that there is no co-dominance or epistasis for any alleles being studied. There are several key differences between living Drosophila and what we will be doing with the computer simulation. The first difference is using virtual flies instead of live ones. These virtual flies are a new species we are naming, Drosophila spartaniensis. This species was created at MSU. The vast majority of the wild type individuals in the population are small ~3 mm long, with red eyes; two straight wings that extend beyond the abdomen; a tan body; and non-feathery antennae (just like the live wild type flies we have previously studied). Male flies are readily distinguished from females in D. spartaniensis by having sex combs on their forelegs and having short, blunt abdomens with three bands; whereas females have long, pointed abdomens with four bands. The second difference is the mutant traits being studied. Since we have created these flies then any traits we will be studying will be inherited in a variety of ways which may be different than a counterpart mutant trait in Drosophila melanogaster. The third difference is the instant life cycle. Whenever you mate the virtual flies, their offspring come up immediately on the screen; there is no waiting for 2-weeks to see them. Don't forget that you still need to keep track of which generation you are observing (e.g., parental or F1 or F2). The fourth difference is the simultaneous

viewing of both the forward and reverse crosses. This makes it a lot faster to use your logic trees to deductively select the mode of inheritance for each mutant trait being studied. The disadvantage is that you have a lot of flies on the screen at one time, and you have to make sure you are not mixing up the two crosses nor missing any offspring which may be further down on the problem page. The fifth difference is the method used to view the virtual flies (no need for the ether or tile or dissecting microscope). You still have to carefully sex and describe the fly phenotype and to record all the counts. What we've learned about meiosis and genetics allows us to examine various alternative hypotheses (modes) about how a particular allele might be inherited, and we can test their associated predictions by making crosses and examining their resulting offspring. If the crosses are designed correctly, then each mode of inheritance will lead to a distinctive phenotypic ratio in the offspring. We can choose among the alternative hypotheses by making statistical comparisons between the observed phenotypic ratios (the evidence) with the expected ratios that are predicted by each mode. Here are four common modes that explain the inheritance of a single allele:

1. Autosomal dominant
2. Autosomal recessive
3. Sex-linked dominant, and
4. Sex-linked recessive.

Below are some common modes that explain the inheritance of two alleles that are assorting independent of one another in a dihybrid cross analysis:

1. Both alleles are autosomal dominant
2. Both alleles are autosomal recessive
3. First allele is autosomal dominant while the second allele is autosomal recessive

Testing Hypotheses about Inheritance

There are also other modes that can be used to explain the inheritance of two alleles which are different than the ones listed above:

1. Both alleles are autosomal and located on the same chromosome (linked) [you still have to determine whether each allele is dominant or recessive]
2. One allele is autosomal while the other allele is sex-linked [you still have to determine whether each allele is dominant or recessive]

There are several ways to generate predictions from hypothesized modes of inheritance. We will use Punnett rectangles because they are simple method and easy to remember. For the alleles we will study in lab, it is necessary to predict the results both of parental crosses—producing the first (F1) generation—and of F1 crosses—producing the second (F2) generation.

The predicted phenotypic ratios are converted to expected frequencies and will be used to compare with the frequencies of real (or simulated) crosses to test your hypotheses.

Each mode of inheritance (i.e., alternative hypothesis) has a unique F2 generation prediction, and therefore the results of experiments can be compared to the predicted value to determine how alleles are passed on from parents to offspring. The best model will be the one whose predictions do not differ statistically from the actual cross values (do not result in the null hypothesis being rejected). We will use the Chi-Square Goodness-of-Fit test to determine whether or not our cross data actually matches our predictions.

There are four models that explain the inheritance of most traits: sex-linked dominant,sex-linked recessive, autosomal dominant, and autosomal recessive. When considering 1 or 2 traits, these possible models may be combined to create as many as 8 hypotheses about patterns of inheritance. Monohybrid Hypotheses

1) Mutant trait is autosomal recessive
2) Mutant trait is autosomal dominant
3) Mutant trait is sex-linked dominant
4) Mutant trait is sex-linked recessive Dihybrid Hypotheses
5) Mutant trait 1 is autosomal recessive and mutant trait 2 is autosomal recessive
6) Mutant trait 1 is autosomal dominant and mutant trait 2 is autosomal dominant
7) Mutant trait 1 is autosomal dominant and mutant trait 2 is autosomal recessive
8) Mutant trait 1 is autosomal recessive and mutant trait 2 is autosomal dominant

There are several ways to construct predictions, or expected data assuming a hypothesized mode of inheritance is true. We will use the Punnett square method because it is an easy - to - remember algorithm based on simple bookkeeping. For the types of traits we will study in lab,
it is necessary to predict the results of parental crosses (producing the first (F1) generation) and of F1 crosses (producing the second (F2) generation).
The predicted phenotypic ratios among the F2 generation are convertedto expected results and will be used to compare with the results of real (or simulated) crosses to test your hypotheses.

Using Punnett Squares to Predict Inheritance

The Punnett square technique involves six basic steps and is a careful way of predicting how the chromosomes bearing a trait of interest may sort during meiosis into gametes, then combine at fertilization.
These steps corresponding with an example are described below. You have a cross and are interest ed in determining how the trait for Brown eyes is inherited. You hypothesize it may be autosomal dominant, and will use a Punnett square to determine the expected phenotype ratio for this hypothesis.
Step 1: Set up the parental cross based on one hypothesis.
Determine if the cross is a monohybrid
or dihybrid cross. Assign a different letter to symbolize each trait:
the hypothesis you are using will determine which you assign to the wild type and which you assign to the mutant trait

Step 1: Establishing the Parents for the Cross
Both parents are true breeding (and express opposite traits)
Male = AA (Brown eyes) Female = aa (Blue eyes)
Step 2: Use the Punnett square to create the F1 generation of the parental cross.
The genotypes of the gametes of the parents are placed at the head of each column (for the male parent) and row (for the female parent). The remaining boxes in the Punnett square represent the genotypes of the offspring produced by combiningthe parental genotypes of one type of sperm and one type of egg. Fill in the progeny genotypes by writing in thegenotype of the mother's gametes (row) and the father's gametes (column) corresponding to each square in the body of the table. These entries are the expected genotypes of the F1 generation.
Step 2: Cross 1 (cross the true breeding parents to create the possible F1 offspring)
A A Aa Aa Aa Aa
Step 3: Analyze the F1 progeny to determine the genotypic and phenotypic ratios. To determine the genotypic ratios of the F1s, simply count the number of squares represented by each genotype Then, convert the number to a ratio of total genotypes. To determine the phenotypic ratios, refer back to the basic hypothesis being analyzed, and determine the ratio of offspring that express the p henotype
Step 3: Analyze the F1 generation
All offspring are heterozygous (Aa) and would have brown eyes.
(Genotype ratio and phenotype ratio are 100%)
Step 4: Construct a new Punnett square to predict the genotypes of the progeny of the F1 cross.
Follow the same procedures used in Step 2, using the genotypes of the F1 male and female parents. F1 male alleles will go in the boxes heading each column and F1 female alleles will go in the boxes heading each row. Continue to fill in the rest of the boxes as you did in Step 2.
Step 4: Cross 2 (cross two offspring from F1 to create the possible F2 offspring)
Aa AA Aa Aa aa
Step 5: Determine the genotypic and phenotypic ratios among the F2, the progeny of the F1 cross.
Use the same procedures and reasoning you used in Step 3. The F2 genotypic and phenotypic ratios you generate in this step are predictions about what the F2ratios should be if the basic hypothesis is true. This is the set of predictions you can test against a realcross to evaluate if the hypothesis might be true.
Step 5: Analyze the F2 generation
When the trait of interest (brown eyes) is hypothesized to be autosomal dominant, our prediction shows that the F2 offspring will display the brown and blue phenotypes in a 3:1 ratio (3 brown -eyed offspring for every 1 blue-eyed offspring).
All possible combinations of A and a alleles are represented in the genotypes of the F2 offspring. For every homozygous dominant (AA) offspring, we expect to see 2 heterozygote (Aa) offspring, a nd 1 homozygous recessive (aa). The final phenotypic and genotypic ratios are represented as:
1 AA : 2 Aa : 1 aa = genotype ratio 3 brown eye : 1 blue eye = phenotype ratio
Step 6: Convert the ratios to expected values, assuming that we are working with 100 offspring.

You will need to scale up the ratio to determine the actual number of offspring out of 100 that will bear that particular phenotype.

Step 6: Expected values

To determine expected values, we convert the ratio to a percentage. For example, ¼ of the offspring (25%) will be homozygous dominant, ½ of the offspring (50%) will be heterozygous, and the remaining ¼ (25%) will be homozygous recessive. Next, we multiple the percentages by the total number of offspring, in this case 100, to determine the actual number of offspring we might expect if this hypothesis is true. 0.25*100 = 25 offspring will be homozygous dominant, and express brown eyes. 0.50*100 = 50 offspring will be heterozygous and express brown eyes. 0.25*100 = 25 offspring will be homozygous recessive and express blue eyes.

Each model of inheritance (i.e. alternative hypothesis) has a unique F2 generation prediction, and therefore the results of experiments can be compared to the predicted value to determine how traits are passed on from parents to offspring. The best model will be the one whose predictions do not differ statistically from the experimental values. We will use the Chi-Squared Test Statistics to determine whether or not our experimental data actually matches our predictions.

Chi-square test

The topic of gene interaction includes a sometimes bewildering array of different phenotypic ratios. Although these ratios are easily demonstrated in established systems such as the ones illustrated in this chapter, in an experimental setting a researcher may observe an array of different progeny phenotypes and not initially know the meaning of this ratio. At this stage, a hypothesis is devised to explain the observed ratio. The next step is to determine whether the observed data are compatible with the expectations of the hypothesis.

In research generally, it is often necessary to compare experimentally observed numbers of items in several different categories with numbers that are predicted on the basis of some hypothesis. For example, you might want to determine whether the sex ratio in some specific population of insects is 1:1 as expected. If there is a close match, then the hypothesis is upheld, whereas, if there is a poor match, then the hypothesis is rejected. As part of this process, a judgment has to be made about whether the observed numbers are a close enough match to those expected. Very close matches and blatant mismatches generally present no problem in judgment, but inevitably there are gray areas in which the match is not obvious. Genetic analysis often requires the interpretation of numbers in various phenotypic classes. In such cases, a statistical procedure called the χ^2 (chi-square) test is used to help in making the decision to hold onto or reject the hypothesis.

The χ^2 test is simply a way of quantifying the various deviations expected by chance if a hypothesis is true. For example, consider a simple hypothesis that a certain plant is a heterozygote (monohybrid) of genotype *A*/*a*. To test this hypothesis, we would make a testcross to *a*/*a* and predict a 1:1 ratio of *A*/*a* and *a*/*a* in the progeny. Even if the hypothesis is true, we do not always expect an exact 1:1 ratio. We can model this experiment with a barrel full

of equal numbers of red and blue marbles. If we blindly removed samples of 100 marbles, on the basis of chance we would expect samples to show small deviations such as 52 red: 48 blue quite commonly and larger deviations such as 60 red:40 blue less commonly. The χ^2 test allows us to calculate the probability of such chance deviations from expectations if the hypothesis is true. But, if all levels of deviation are expected with different probabilities even if the hypothesis is true, how can we ever reject a hypothesis? It has become a general scientific convention that a probability value of less than 5 percent is to be taken as the criterion for rejecting the hypothesis. The hypothesis might still be true, but we have to make a decision somewhere, and the 5 percent level is the conventional decision line. The logic is that, although results this far from expectations are expected 5 percent of the time even when the hypothesis is true, we will mistakenly reject the hypothesis in only 5% of cases and we are willing to take this chance of error.

Let's consider an example taken from gene interaction. We cross two pure lines of plants, one with yellow petals and one with red. The F_1 are all orange. When the F_1 is salve to give an F_2, we find the following result:

orange	182
yellow	61
red	77
Total	320

What hypothesis can we invent to explain the results? There are at least two possibilities:

Hypothesis 1. Incomplete dominance

	(yellow) *G1/G1* × *G2/G2* (red)		
F_1	*G1/G2* (orange)		
			Expected numbers
F_2	$\frac{1}{4}$ *G1/G1*	(yellow)	80
	$\frac{1}{2}$ *G1/G2*	(orange)	160
	$\frac{1}{4}$ *G2/G2*	(red)	80

Hypothesis 2. Recessive epistasis of *r* (red) on *Y* (orange) and *y* (yellow)

	(yellow) *y/y* ; *R/R* × *Y/Y* ; *r/r* (red)		
F_1	*Y/y* ; *R/r* (orange)		
			Expected numbers
F_2	$\frac{9}{16}$ *Y/–* ; *R/–*	(orange)	180
	$\frac{3}{16}$ *y/y* ; *R/–*	(yellow)	60
	$\frac{3}{16}$ *Y/–* ; *r/r*	(red)	80
	$\frac{1}{16}$ *y/y* ; *r/r*	(red)	

The statistic χ^2 is always calculated from actual numbers, not from percentages, proportions, or fractions. Sample size is therefore very important in the χ^2 test, as it is in most considerations of chance phenomena. Samples to be tested generally consist of several classes. The letter *O* is used to represent the observed number in a class, and *E* represents the expected number for the same class based on the predictions of the hypothesis. The general formula for calculating χ^2 is as follows:

$$\text{Sum of } (O - E)^2/E \text{ for all classes}$$

For hypothesis 1, the calculation is as follows:

	O	E	$(O - E)^2$	$(O - E)^2/E$
orange	182	160	484	3.0
yellow	61	80	361	4.5
red	77	80	9	0.1
				$\chi^2 = 7.6$

To convert the χ^2 value into a probability, we use Table 4-1, which shows χ^2 values for different degrees of freedom (df). For any total number of progeny, if the number of individuals in two of the three phenotypic classes is known, then the size of the third class is automatically determined. Hence, there are only *2 degrees of freedom* in the distribution of individuals among the three classes. Generally, the number of degrees of freedom (shown as the different rows of Table 4-1) is the number of classes minus 1. In this case, it is 3 − 1 = 2. Looking along the 2-df line, we find that the χ^2 value places the probability at less than 0.025, or 2.5 percent. This means that, if the hypothesis is true, then deviations from expectations this large or larger are expected approximately 2.5 percent of the time. As mentioned earlier, by convention the 5 percent level is used as the cutoff line. When values of less than 5 percent are obtained, the hypothesis is rejected as being too unlikely. Hence the incomplete dominance hypothesis must be rejected.

Table 4-1 Critical Values of the χ^2 Distribution

	P									
df	0.995	0.975	0.9	0.5	0.1	0.05	0.025	0.01	0.005	df
1	.000	.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	9

Table 4-1

Critical Values of the χ^2 Distribution.

For hypothesis 2, the calculation is set up as follows.

	O	E	$(O-E)^2$	$(O-E)^2/E$
orange	182	180	4	0.02
yellow	61	60	1	0.02
red	77	80	9	0.11
				$\chi^2 = 0.15$

The probability value (for 2 df) this time is greater than 0.9, or 90 percent. Hence a deviation this large or larger is expected approximately 90 percent of the time—in other words, very frequently. Formally, because 90 percent is greater than 5 percent, we conclude that the results uphold the hypothesis of recessive epistasis.

Chromosomal basis of Mendelism - Sutton and Boveri hypothesis with experimental evidences

A pair of papers by Sutton as early as 1902 and 1903 (only two and three years, respectively, after the rediscovery of Mendel's landmark papers initiated the modem age of genetics) clearly pointed the way to a physical basis for the science of heredity. The chromosomal theory of inheritance was proposed independently by Sutton and Boveri in 1902.

Sutton was a graduate student of Wilson at Columbia University and is credited for demonstrating a parallel between meiotic behavior of paired chromosomes and the behavior of pairs of Mendelian factors. He could explain Mendel's principle of segregation on cytological basis.

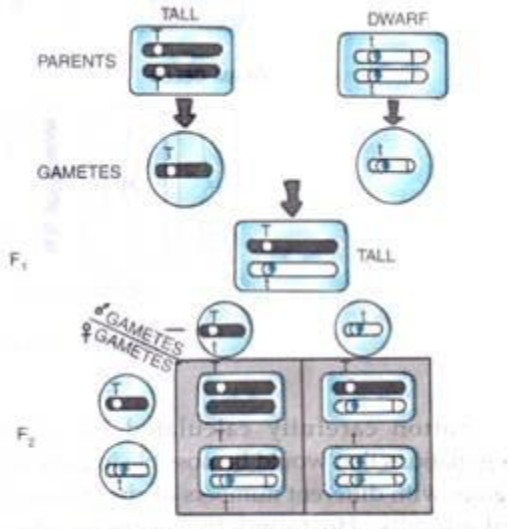


Fig. 5.27. Cytological basis of Mendel's Law of Segregation (Monohybrid cross).

That in meiosis one member of a pair of homologous chromosomes goes to one daughter cell, the other to the second daughter cell. Mendel's principle of Independent Assortment found cytological proof from the fact that members of one pair of homologous chromosomes move to the poles independently of the members of another pair.

Walters S. Sutton argued that these parallels between the behavior of chromosomes and Mendel's factors are striking to be accidental.

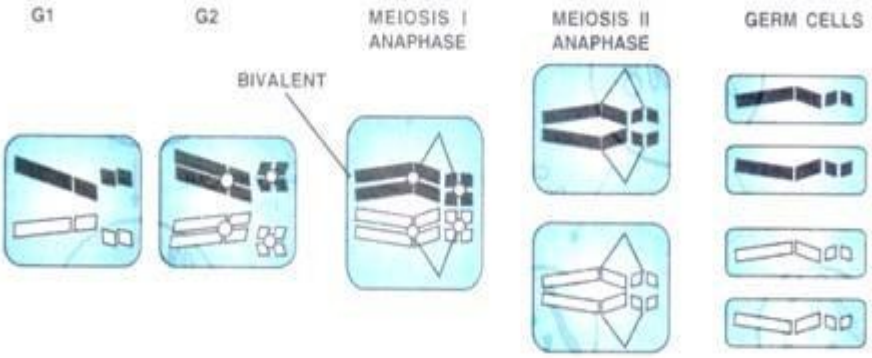


Fig.5.28. Meiosis and germ cell formation in a cell with two chromosomes (represented in black and white). Notice the segregation of chromosomes for formation of germ cells.

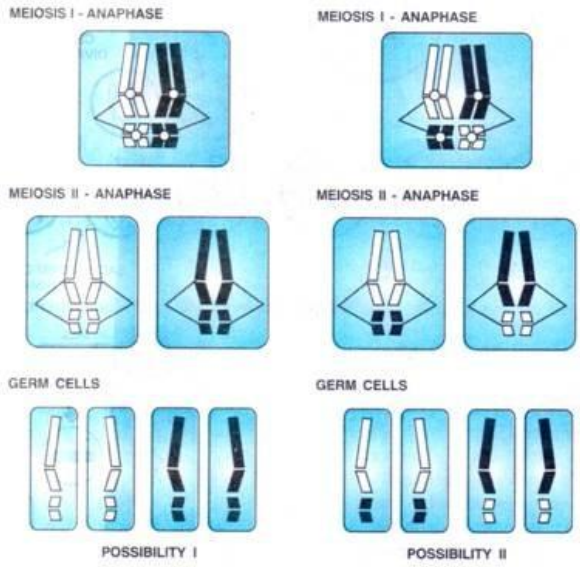


Fig.5.29. Independent assortment of chromosomes.

Sutton carefully calculated the number of chromosome combinations is just the same as the number combinations that would be possible in gametes and in of combinations of factors Mendel postulated in zygotes with different numbers of chromosomes in the explaining the results of crosses with pea plants, diploid cells.

He found that number of possible chromosome combinations is just the same as the number of combinations of factors Mendel postulated in explaining the results of crosses with pea plants.

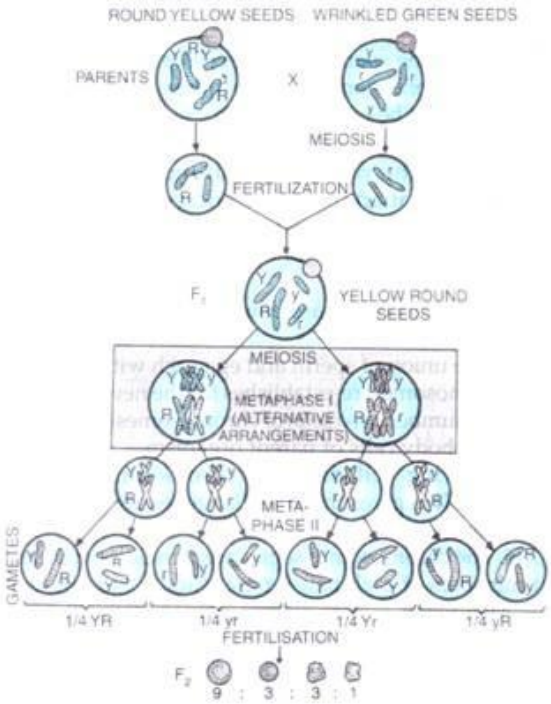


Fig. 1.30. Independent assortment of the chromosomes and the genes they carry occurs because the homologous paired chromosomes can arrange themselves and separate in two different ways during meiosis. This leads to four types of allelic combination in the gametes. The F_2 yields a phenotypic ratio of 9 : 3 : 3 : 1.

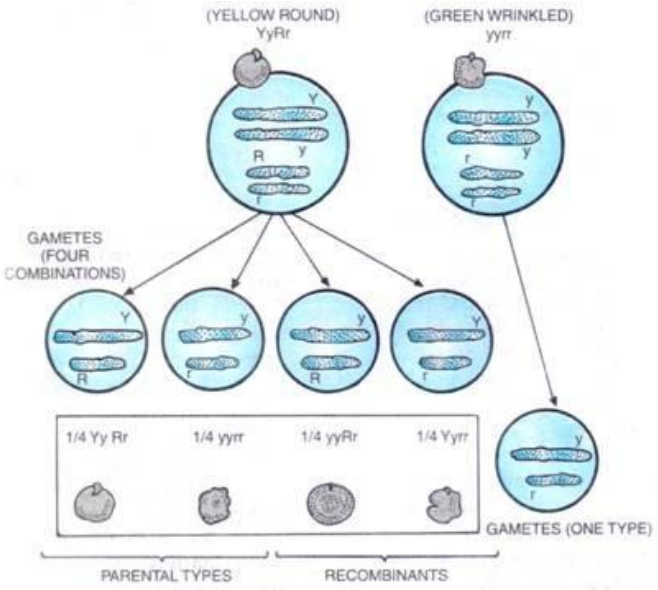


Fig. 1.31. Test cross of heterozygous dihybrid with a double recessive parent. With the genes located on separate chromosomes, four types of progeny are obtained in equal proportions. Two of

these are parental types and two are recombinants. Parental and recombinant frequencies are 50% each.

Sutton and Boveri's arguments for his chromosome theory of heredity were essentially as follows:

1. Since the sperm and egg cells provide the only bridge from one generation to other, all hereditary characters must be carried in them.
2. The sperm cell lose practically all their cytoplasm in the process of maturation, as can be seen by observation under the microscope. Since the sperm contributes as much to heredity as does the egg, the hereditary factors must be carried in the nucleus.
3. Chromosomes are found in pairs; so as the Mendelian factors.
4. The union of sperm and egg each with its single set of chromosomes re-establishes for the new organism the whole number (two sets) of chromosomes previously seen in the body cells of parent organism.
5. During cell division, chromosomes divide accurately. This gives the idea that genes are carried on chromosomes.
6. The chromosomes segregate at meiosis. It means that members of each pair separate and go to different cells. Mendelian factors also segregate at the time of formation of gametes.
7. The members of chromosome pair segregate independently of other chromosome pairs. Mendelian genes also segregate independently.

In other words, chromosomes obey Mendel's laws of inheritance. Sutton concluded his 1902 paper with bold prediction "I may finally call attention to the probability that (the behavior of chromosomes) may constitute the physical basis of the Mendelian laws of heredity". Truly the door was thereby opened to an objective examination of the physical mechanisms of the genetic processes.

POSSIBLE QUESTIONS

2 marks

1. Define heredity.
2. Name the three laws of mendel.
3. Write the equation for binomial expansion.
4. State the laws of probability.
5. State the law of inheritance
6. State the law of independent assortment.
7. State the law of dominance.
8. Name any three characteristic features of model organisms.

8 Marks

1. Explain in detail on formulating and testing genetic hypothesis.
2. Discuss on chromosomal basis of mendelism.
3. Elaborate on Model organisms and its characteristic features.
4. What is dominance and recessive? How it is employed in Mendelian genetics.
5. What are the basic principles of heredity? Explain in detail.
6. Explain the hypothesis of Sutton and Boveri with experimental evidences.
7. Discuss in detail about laws of probability and binomial expansion.
8. Elucidate on monohybrid and dihybrid crosses of Mendel with examples

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
II-B.Sc., BIOCHEMISTRY
CONCEPTS IN GENETICS (17BCU304B)`
MULTIPLE CHOICE QUESTIONS
UNIT I

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	Which one of the following best describes a gene?	A triplet of nucleotide bases.	A specific length of DNA responsible for the inheritance and expression of the character.	A specific length of single stranded RNA.	Both (B) and (C)	**A specific length of DNA responsible for the inheritance and expression of the character.**
2	Mendel's "factors" are in fact	units	chromosomes	genes	none of these	**genes**
3	Who coined the term 'gene' for 'factor'?	Mendel	Morgan	Johannsen	Punnett	**Johannsen**
4	Alleles or allelomorphs occupy	same position on homologous chromosomes.	same position on heterozygous chromosomes.	different position on homologous chromosomes. .	different position on heterozygous chromosomes	**same position on homologous chromosomes.**
5	Who proposed the term 'Allelomorph'?	Hugo De Vries	Morgan	Tschermak	Bateson	**Bateson**
6	Dominant allele means	an allele whose effect is masked by another allele.	an allele that prevents the expression of the other allele.	an allele without any effect.	an allele which cannot express in presence of other.	**an allele that prevents the expression of the other allele.**
7	The external appearance of an individual for any trait is called as	phenotype	karyotype	morphology	physique	**phenotype**
8	Genotype is	genetic constitution of an organism.	genetic constitution of somatic cells.	genetic constitution of plastids.	genetic constitution of germ cells.	**genetic constitution of an organism.**
9	Homozygous individuals	breed true to the trait.	does not breed true to the trait.	produce only one type of gamete.	both (a) and (c)	**both (a) and (c)**
10	Which of the following term indicates a pair of dissimilar alleles?	Homozygous	Heterozygous	Homologous	All of these	**Heterozygous**
11	A cross between two pure individuals, differing in atleast one set of characters, is called	monohybrid	polyploid	mutant	variant	**monohybrid**
12	F1 generation means	first flowering generation	first fertile generation	first filial generation	first seed generation	**first filial generation**
13	Filial means	offsprings produced in sexual reproduction.	offsprings produced in asexual reproduction.	offsprings produced in vegetative	reproduction. both (b) and (c)	**offsprings produced in sexual reproduction.**
14	F2generation is produced by	crossing F1 progeny with one of the parents.	selfing the heterozygous progeny.	selfing the parents.	a cross between recessive parents	**selfing the heterozygous progeny.**
15	In genetics, the use of checkerboard was done by	Mendel	Correns	Punnet	Darwin	**Punnet**
16	Mendel, in his experiments	maintained qualitative records.	maintained quantitative records.	conducted ample crosses and reciprocal crosses.	all of the above	**all of the above**
17	To eliminate chance factor, Mendel performed	monohybrid cross	dihybrid cross	reciprocal cross	trihybrid cross	**reciprocal cross**
18	Mendel always started his experiment (Monohybrid and Dihybrid cross) with	any pea plant	a heterozygous plant	a pure line plant	a fresh new plant	**a pure line plant**
19	Mendel carried out artificial cross by	emasculation of selected female parent plant	emasculation of selected male parent plant	dusting of pollen grains from selected male plant over selected female plant	both (a) and (c)	**both (a) and ©**

20	F3generation was obtained by	selfing F1hybrids	selfing F2hybrids	crossing F1with either parent	none of these	**selfing F2hybrids**
21	What result did Mendel obtained after monohybrid cross between tall and dwarf pea plant?	All new plants were dwarf.	All new plants were tall.	50% plants were dwarf and 50% plants were tall.	75% plants were tall and 25% plants were dwarf.	**All new plants were tall.**
22	When Mendel allowed natural selfing of F1 hybrids during monohybrid cross between pure tall and pure dwarf pea plant, he found	all plants were tall.	all plants were dwarf.	dwarfness reappeared in some plants.	tallness reappeared in some plants.	**dwarfness reappeared in some plants.**
23	During monohybrid cross experiments, Mendel performed reciprocal cross by selecting	tall plant as male and dwarf plant as female.	tall plant as female and dwarf plant as male.	both male and female plant as tall.	both male and female plant as dwarf.	**tall plant as male and dwarf plant as female.**
24	After performing reciprocal cross between tall and dwarf plants, the ratio of tall and dwarf plants obtained was	1:02	3:1	1:03	2:1	**3:1**
25	The first scientific explanation regarding inheritance was given by	William Bateson	Gregor Johann Mendel	Griffith	Johannsen	**Gregor Johann Mendel**
26	Mendel selected pea plant because of	its short life span.	it produced many seeds and large flowers	many contrasting characters.	all of these	**all of these**
27	The botanical name of garden pea is	Pisum sativum	Lathyrus odoratus	Mangifera indica	Solanum tuberosum	**Pisum sativum**
28	F2 generation is produced by	crossing F1 progeny with one of the parents	selfing the heterozygous progeny	selfing the parents.	a cross between recessive parents.	**selfing the parents.**
29	The phenomenon of ‘like begets like’ is due to	genetics	heredity	germplasm	variation	**heredity**
30	Organisms produced by asexual reproduction are called	clones	offsprings	factors	both (A) and (B)	**clones**
31	Filial means	offsprings produced in sexual reproduction.	offsprings produced in asexual reproduction.	offsprings produced in vegetative reproduction.	both (B) and ©	**offsprings produced in sexual reproduction.**
32	The phenotypic ratio of F2 progeny in a dihybrid cross is	9 : 3 : 3 : 1: 1	9 : 3 : 3 : 1	9 : 1 : 3 : 3 : 1	1 : 2 : 2 : 4 : 1 : 2 : 1 : 2 : 1	**9 : 3 : 3 : 1**
33	Transmission of characters from one generation to the next or from parents to offsprings is called	heredity	variation	recombination	mutation	**heredity**
34	The genotypic ratio obtained in a Mendelian dihybrid cross is	1 : 2 : 2 : 4 : 1 : 2 : 1 : 2 : 1	9 : 3 : 3 : 1	1 : 4 : 4 : 1 : 2 : 2 : 1 : 1	9:07	**1 : 2 : 2 : 4 : 1 : 2 : 1 : 2 : 1**
35	Law of independent assortment can be explained by	monohybrid cross and monohybrid ratio	dihybrid cross and dihybrid ratio	trihybrid cross and trihybrid ratio	all of the above	**dihybrid cross and dihybrid ratio**
36	A cross between two pure individuals differing in two sets of characters is called	dihybrid cross	monohybrid cross	trihybrid cross	reciprocal cross	**dihybrid cross**
37	Variation is	differences between parents and offsprings.	differences between individuals of same species.	differences among the offsprings of the same parents.	all of the above.	**all of the above.**

38	Who is known as “Father of Genetics”?	Theophrastus	Stephen Hales	Mendel	Aristotle	**Mendel**
39	The term “genetics” was coined by	Morgan	William Bateson	Johannsen	Karl Correns	**William Bateson**
40	Offsprings are	exactly identical to either of their parents.	not exactly identical to either of their parents.	show intermediate characters inherited from both the parents.	both (B) and (C)	**both (B) and (C)**
41	(3 : 1) ´ (3 : 1) = 9 : 3 : 3 : 1 This signifies	trihybrid ratio	two monohybrid ratio	dihybrid ratio is a product of two monohybrid ratios	none of the above	**dihybrid ratio is a product of two monohybrid ratios**
42	Dihybrid ratio is defined as	phenotypic ratio obtained in F2 generation of dihybrid cross.	phenotypic ratio obtained in F1 generation of dihybrid cross.	genotypic ratio obtained in F2 generation of dihybrid cross.	genotypic ratio obtain	**phenotypic ratio obtained in F2 generation of dihybrid cross.**
43	Mendel was a	physiologist	mathematician	cytologist	taxonomist	**mathematician**
44	Mendel’s laws were first published in the year	1875	1890	1928	1866	**1866**
45	The year 1900 A.D. is highly significant for geneticists due to	chromosome theory of heredity	discovery of genes	rediscovery of Mendelism	principle of linkage	**rediscovery of Mendelism**
46	The Mendelian principles of inheritance were rediscovered by	Sutton and Boveri	Hugo de Vries, Tschermark and Correns	Lederberg and Tatum	Morgan	**Hugo de Vries, Tschermark and Correns**
47	Mendel’s work was rediscovered by three biologists from which of the following countries?	Holland, France and England	Holland, England and Austria	Germany, France and England	Austria, Holland and Germany	**Austria, Holland and Germany**
48	Which of the following is a dominant character in pea?	Wrinkled seeds	Inflated pod	Terminal flower	Dwarf plant	**Inflated pod**
49	Which of the following character was not considered by Mendel?	Seed coat colour	Wrinkled or round leaves	Tallness or dwarfness	Position of flower	**Wrinkled or round leaves**
50	An inherited character and its detectable variant is called	allele	trait	gene	both (A) and (B)	**trait**

UNIT II
SYLLABUS

Extensions of Mendelism, genetics of a gene, bacteria and viruses
Allelic variation and gene function- dominance relationships, multiple alleles, lethal alleles and null alleles. Pleiotropy gene interaction- epistatic and non-epistatic, interaction between gene(s) and environment. Penetrance and expressivity, norm of reaction and phenocopy. Complementation test, limitations of *cis-trans* test, intragenic complementation, rII locus of phage T4 and concept of cistron. Mechanism of genetic exchange - conjugation, transformation and transduction. Gene mapping in bacteria.

ALLELIC VARIATION AND GENE FUNCTION

DOMINANCE RELATIONSHIPS

GENETIC DOMINANCE

Why you have that particular eye color or hair type? It's all due to gene transmission. As discovered by Gregor Mendel, traits are inherited by the transmission of genes from parents to their offspring. Genes are segments of DNA located on our chromosomes. They are passed on from one generation to the next through sexual reproduction. The gene for a specific trait can exist in more than one form or allele. For each characteristic or trait, animals cells typically inherit two alleles. Paired alleles can be homozygous (having identical alleles) or heterozygous(having different alleles) for a given trait.

When the allele pairs are the same, the genotype for that trait is identical and the phenotype or characteristic that is observed is determined by the homozygous alleles. When the paired alleles for a trait are different or heterozygous, several possibilities may occur. Heterozygous dominance relationships that are typically seen in animal cells include complete dominance, incomplete dominance, and co-dominance.

- Complete Dominance
- Incomplete Dominance
- Co-dominance
- Differences Between Incomplete Dominance and Co-dominance

COMPLETE DOMINANCE

In **complete dominance** relationships, one allele is dominant and the other is recessive. The dominant allele for a trait completely masks the recessive allele for that trait. The phenotype is determined by the dominant allele. For example, the genes for seed shape in pea plants exists in two forms, one form or allele for round seed shape **(R)** and the other for wrinkled seed shape **(r)**.

In pea plants that are heterozygous for seed shape, the round seed shape is dominant over the wrinkled seed shape and the genotype is **(Rr)**.

- Genetic Dominance
- Incomplete Dominance
- Co-dominance
- Differences Between Incomplete Dominance and Co-dominance

INCOMPLETE DOMINANCE

In incomplete dominance relationships, one allele for a specific trait is not completely dominant over the other allele. This results in a third phenotype in which the observed characteristics are a mixture of the dominant and recessive phenotypes. An example of incomplete dominance is seen in hair type inheritance. Curly hair type **(CC)** is dominant to straight hair type **(cc)**. An individual who is heterozygous for this trait will have wavy hair **(Cc)**. The dominant curly characteristic is not fully expressed over the straight characteristic, producing the intermediate characteristic of wavy hair. In incomplete dominance, one characteristic may be slightly more observable than another for a given trait. For example, an individual with wavy hair may have more or fewer waves than another with wavy hair. This indicates that the allele for one phenotype is expressed slightly more than the allele for the other phenotype.

- Genetic Dominance
- Complete Dominance
- Co-dominance
- Differences Between Incomplete Dominance and Co-dominance

CO-DOMINANCE

In **co-dominance** relationships, neither allele is dominant, but both alleles for a specific trait are completely expressed. This results in a third phenotype in which more than one phenotype is observed. An example of co-dominance is seen in individuals with the sickle cell trait. Sickle cell disorder results from the development of abnormally shaped red blood cells. Normal red blood cells have a biconcave, disc-like shape and contain enormous amounts of a protein called hemoglobin. Hemoglobin helps red blood cells bind to and transport oxygen to cells and tissues of the body. Sickle cell is a result of a mutation in the hemoglobin gene. This hemoglobin is abnormal and causes blood cells to take on a sickle shape. Sickle-shaped cells often become stuck in blood vessels blocking normal blood flow. Those that carry the sickle cell trait are heterozygous for the sickle hemoglobin gene, inheriting one normal hemoglobin gene and one sickle hemoglobin gene. They do not have the disease because the sickle hemoglobin allele and normal hemoglobin allele are co-dominant with regard to cell shape. This means that

both normal red blood cells and sickle-shaped cells are produced in carriers of the sickle cell trait. Individuals with sickle cell anemia are homozygous recessive for the sickle hemoglobin gene and have the disease.

- Genetic Dominance
- Complete Dominance
- Incomplete Dominance
- Differences Between Incomplete Dominance and Co-dominance

DIFFERENCES BETWEEN INCOMPLETE DOMINANCE AND CO-DOMINANCE

The pink tulip color is a mixture of the expression of both alleles (red and white), resulting in an intermediate phenotype (pink). This is incomplete dominance. In the red and white tulip, both alleles are completely expressed. This shows co-dominance. Pink / Peter Chadwick LRPS/Moment/Getty Images - Red and white / Sven Robbe/EyeEm/Getty Images

INCOMPLETE DOMINANCE VS CO-DOMINANCE

People tend to confuse incomplete dominance and co-dominance relationships. While they are both patterns of inheritance, they differ in gene expression. Some differences between the two are listed below:

1. ALLELE EXPRESSION

- **Incomplete Dominance:** One allele for a specific trait is not completely expressed over its paired allele. Using flower color in tulips as an example, the allele for red color **(R)** does not totally mask the allele for white color **(r)**.
- **Co-dominance:** Both alleles for a specific trait are completely expressed. The allele for red color **(R)** and the allele for white color **(r)** are both expressed and seen in the hybrid.

2. ALLELE DEPENDENCE

- **Incomplete Dominance:** The effect of one allele is dependent upon its paired allele for a given trait.
- **Co-dominance:** The effect of one allele is independent of its paired allele for a given trait.

3. PHENOTYPE

- **Incomplete Dominance:** The hybrid phenotype is a mixture of the expression of both alleles, resulting in a third intermediate phenotype. Example: Red flower **(RR)** X White flower **(rr)** = Pink flower **(Rr)**

- **Co-dominance:** The hybrid phenotype is a combination of the expressed alleles, resulting in a third phenotype that includes both phenotypes. (Example: Red flower **(RR)** X White flower **(rr)** = Red and white flower **(Rr)**

4. OBSERVABLE CHARACTERISTICS

- **Incomplete Dominance:** The phenotype may be expressed to varying degrees in the hybrid. (Example: A pink flower may have lighter or darker coloration depending on the quantitative expression of one allele versus the other.)
- **Co-dominance:** Both phenotypes are fully expressed in the hybrid genotype.

OVERVIEW

In **incomplete dominance** relationships, one allele for a specific trait is not completely dominant over the other allele. This results in a third phenotype in which the observed characteristics are a mixture of the dominant and recessive phenotypes. In **co-dominance** relationships, neither allele is dominant but both alleles for a specific trait are completely expressed. This results in a third phenotype in which more than one phenotype is observed.

- Genetic Dominance
- Complete Dominance
- Incomplete Dominance
- Co-dominance

Additional input

In **codominance**, both alleles are expressed and appear in the heterozygote. Human ABO blood types are one example of where this occurs. Blood type in this case is determined by a single locus, where three different alleles are possible (I^A, I^B, I^O). I^O is completely recessive to both I^A and I^B, so an I^AI^O individual is phenotypically A and an IBIO individual is phenotypically B. However, an individual whose genotype is I^AI^B, is phenotypically AB, as the I^A and I^B alleles are codominant.

To understand why this is the case, its necessary to look at ABO gene function. It may all sound a bit "Old MacDonald's Farm" with all these I^As and I^Os, but bear with us on this one.

What are blood types about anyway? Vampire lore says different blood types have different flavors—well, maybe, if you're a vampire with a particularly sensitive palate. But for us regular, everyday mortals, the important point is this: the different blood groups have different kinds of proteins, called antigens, on the surface of their blood cells.

The I^A allele codes for the production of "A antigens," I^B for "B antigens," while I^O doesn't produce any kind of antigen at all (the slacker).

Also, the immune systems of individuals with different blood groups produce different antibodies (proteins that identify and help neutralize foreign objects and cells). Individuals with blood type A produce antibodies against B antigens, individuals with blood type B produce antibodies against A antigens, but individuals with blood type AB don't produce antibodies against either A or B antigens (because it would be a bit of a disaster all round for them if they did). Individuals with blood type O produce antibodies against both A and B antigens. This, by the way, is why people with AB blood type are universal receivers; they can receive blood transfusions from any blood group precisely because they produce no anti-blood group antibodies. On the other hand, people with O blood types are universal donors: anyone can receive their blood because it can't provoke an immune response, thanks to having no antigens on the surface of its blood cells.

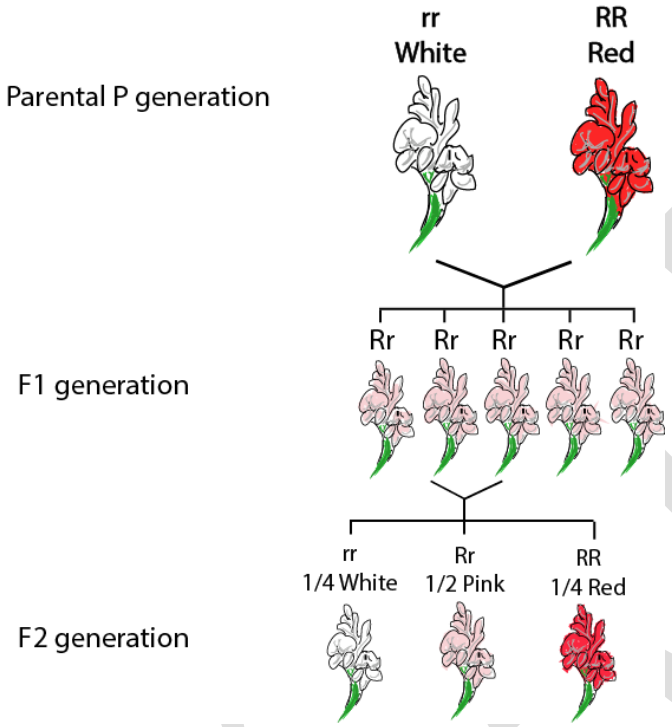
When alleles show **incomplete dominance**, the phenotype of the heterozygote is intermediate in nature between those produced by the homozygous genotypes. This is because the heterozygote has only one functional copy of the gene, and thus produces half the protein. Think of it as a dosage problem: instead of two teaspoons of cough syrup (homozygote for cough syrup dose), you get a teaspoon of cough syrup and a teaspoon of water (heterozygote for cough syrup dose), or two teaspoons of water (no cough syrup at all—ouch!).

Let's look at a real-world example: if a true-breeding red snapdragon (*RR*) is crossed to a true-breeding white snapdragon (*WW*), all of the resulting F1 are pink (*RW*). The pink simply results from producing half the amount of red pigment.

It's tempting, however, to think it works just like paint—red paint mixed with some white paint produces pink paint. In fact, before Mendel's discoveries, many believed in such "blending inheritance": that the phenotype of an offspring resulted from the mixing of the parent's phenotypes, and so is usually in between their traits.

Back it up, though. Taking a closer look at flower color inheritance reveals it's actually consistent with Mendelian principles. If the pink F1 individuals (*RW*) are interbred to produce an F2 generation, 50% of the resulting plants have pink flowers (*RW*), 25% have red flowers (*RR*), and 25% have white flowers (*WW*). The F1 plants produced gametes carrying *R* and *W* in equal numbers that, when mixed up to produce new plants, generate the same genotypic ratios as with the traditional Mendelian traits.

Just like it was in the case of codominance, however, because the heterozygote produces a different phenotype than the homozygous dominant, the phenotypic ratios exactly match the genotypic ratios (which they don't do in a standard dominant/recessive relationship).



Multiple Alleles

The word allele is a general term to denote the alternative forms of a gene or contrasting gene pair that denote the alternative form of a gene is called allele. These alleles were previously considered by Bateson as hypothetical partner in Mendelian segregation.

In Mendelian inheritance a given locus of chromosome was occupied by 2 kinds of genes, i.e., a normal gene (for round seed shape) and other its mutant recessive gene (wrinkled seed shape). But it may be possible that normal gene may show still many mutations in pea besides the one for wrinkledness. Here the locus will be occupied by normal allele and its two or more mutant genes.

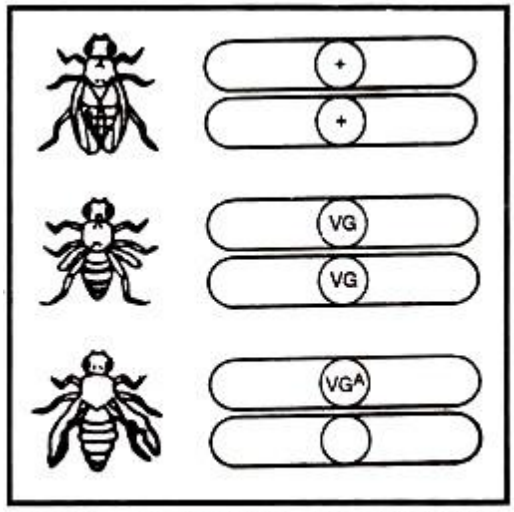


Fig. 57. Multiple alleles. Different form of wings in *Drosophila*.

Thus, three or more kinds of genes occupying the same locus in individual chromosome are referred to as multiple alleles. In short many alleles of a single gene are called multiple alleles. The concept of multiple alleles is described under the term "multiple allelism".

Dawson and Whitehouse in England proposed the term panallele for all the gene mutations at a given locus in a chromosome. These differ from the multiple factor in one respect that multiple factors occupy different loci while alleles occupy same locus.

"Three or more kinds of gene which occupy the same locus are referred to as multiple alleles." Altenburg

Characteristics of Multiple Alleles:

1. The study of multiple alleles may be done in population.
2. Multiple alleles are situated on homologous chromosomes at the same locus.
3. There is no crossing over between the members of multiple alleles. Crossing over takes place between two different genes only (inter-generic recombination) and does not occur within a gene (intragenic recombination).
4. Multiple alleles influence one or the same character only.
5. Multiple alleles never show complementation with each other. By complementation test the allelic and non-allelic genes may be differentiated well. The production of wild type phenotype in a trans-heterozygote for 2 mutant alleles is known as complementation test.
6. The wild type (normal) allele is nearly always dominant while the other mutant alleles in the series may show dominance or there may be an intermediate phenotypic effect.

7. When any two of the multiple alleles are crossed, the phenotype is of a mutant type and not the wild type.
8. Further, F_2 generations from such crosses show typical monohybrid ratio for the concerned character.

Examples of Multiple Alleles:

Wings of Drosophila:

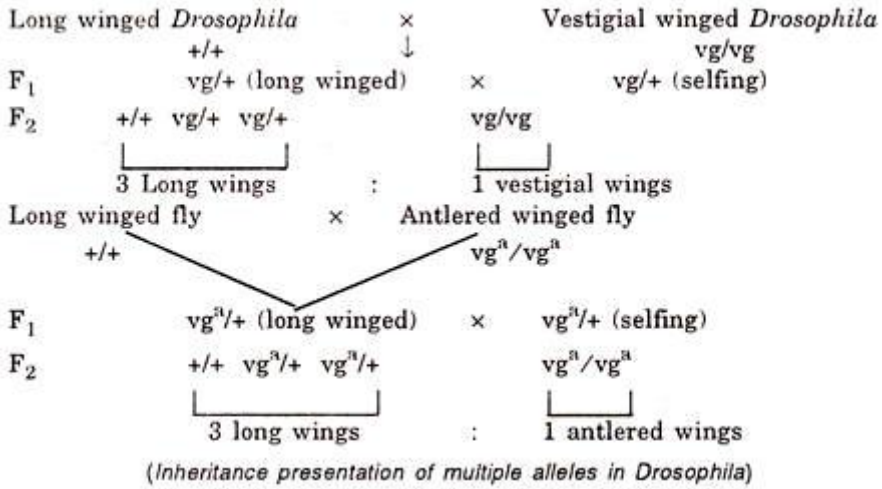
In Drosophila wings are normally long. There occurred two mutations at the same locus in different flies, one causing vestigial (reduced) wings and other mutation causing antlered (less developed) wings. Both vestigial and antlered are alleles of the same normal gene and also of each other and are recessive to the normal gene.

Suppose vestigial is represented by the symbol 'vg' and antlered wing by 'vg^a'. The normal allele is represented by the symbol +.

Thus, there are three races of Drosophila:

(i) Long ++ (+/+)
(ii) Vestigial vg vg (vg/vg)
(iii) Antlered vg^a vg^a (vg^a/vg^a)

A cross between a long winged normal fly and another having vestigial wings or antlered wings is represented below:



(Inheritance presentation of multiple alleles in Drosophila)

When a fly with vestigial wing is crossed with another fly having antlered wings, the F_1 hybrids are intermediate in wing length showing that none of the mutated gene is dominant over the

other. This hybrid is some times said as the vestigial antlered compound and contains two mutated genes at the same locus. They show Mendelian segregation and recombination.

	Vestigial winged fly	×	Antlered winged fly
	vg/vg		vg^{a}/vg^{a}
F_1	vg/vg^{a} (intermediate)	×	vg/vg^{a} (selfing)
F_2	vg/vg,	vg/vg^{a}, vg/vg^{a}	vg^{a}/vg^{a}
	(vestigial)	(intermediate)	(antlered)
	1 :	2 :	1

(Cross representation between a vestigial and an antlered winged fly)

Besides the vestigial and antlered wing described above there are several other mutations occurring at the same locus and resulting in nicked wings, strap wings or no wings etc. These are all multiple alleles.

Close Linkage Versus Allelism:

If we assume that these mutant genes, vestigial and antlered are not allelic located at different loci in place of locating at same locus in different chromosomes so closely linked that there is no crossing over between them, the mutant gene will suppress the expression of adjacent normal allele to certain extent.

These closely linked genes are called pseudo alleles and this suppression is the result of position effect. Thus, visible or apparent cases of allelism may be explained on the assumption of close linkage.

Another example of multiple alleles is the eye colour in Drosophila. The normal colour of the eye is red. Mutation changed this red eye colour to white. Other mutations at white locus took place changing the red eye colour to various lighter shades like cherry, apricot, eosin, creamy, ivory, blood etc., are also visible and are due to multiple alleles.

A cross between the two mutant forms, produces intermediate type in the F_1 except white and apricot races which are not alleles but closely linked genes.

Coat Colour in Rabbit:

The colour of the skin in rabbits is influenced by a series of multiple alleles. The normal colour of the skin is brown. Besides it there are white races called albino and Himalayan as the mutant races. The Himalayan is similar to albino but has darker nose, ear, feet and tail. The mutant genes albino (a) and Himalayan (a^{h}) occupy the same locus and are allelic. Both albino and Himalayan are recessive to their normal allele (+).

A cross between an albino and Himalayan produces a Himalayan in the F_1 and not intermediate as is usual in the case of other multiple alleles.

	Albino	×		Himalayan			
	a/a	↓		a^h/a^h			
F_1		a^h/a (Himalayan)					
	a^h/a	×	a^h/a	(Selfing)			
F_2	a^h/a^h	:	a^h/a	:	a^h/a	:	a/a
			3	:			1
			Himalayan	:			Albino

(*Representing a cross between an albino and Himalayan races of rabbits*)

Blood Groups in Man:

Several genes in man produce multiple allelic series which affect an interesting and important physiological characteristic of the human red blood cells. The red blood cells have special antigens properties by which they respond to certain specific components (antibodies) of the blood serum.

The antigen-antibody relationship is one of the great specificity like that between lock and key. Each antigen and its associated antibody has a peculiar chemical configuration. Landsteiner discovered in 1900 that when the red cells of one person are placed in the blood serum of another person, the cells become clumped or agglutinated.

If blood transfusions were made between persons of two such incompatible blood groups, the transfused cells were likely to clump and shut out the capillaries in the recipient, sometimes resulting in death.

However, such reactions occurred only when the cells of certain individuals were placed in serum from certain other persons. It was found that all persons could be classified in to four groups with regard to the antigen property of the blood cells.

Large number of persons have been classified in to these four groups by means of the agglutination test and the distribution of blood groups in the offspring of parents of known blood groups has been studied. The evidence shows that these blood properties are determined by a series of three allelic genes I^A, I^B and i, as follows:

Blood groups	Genotype
AB	I^AI^B
B	I^BI^B or I^Bi
A	I^AI^A or I^Ai
O	ii

I^A is a gene for the production of the anti-gin A. I^B for antigen B, and i for neither antigen. The existence of these alleles in man and the case with which the blood groups can be identified have

obvious practical applications in blood transfusion, cases of disputed percentage and description of human populations.

The alleles of these genes which affect a variety of biochemical properties of the blood, act in such a way that in the heterozygous compound I^AI^B, each allele exhibits its own characteristics and specific effect. The cells of the heterozygote contain both antigens A and B. On the other hand, I^A and I^B both show complete dominance over i, which lacks both antigens.

PARENTS		CHILDREN	
Phenotypes	Genotypes	Phenotypes	Genotypes
O X O	ii X ii	O	ii
O X A	ii X I^AI^A or I^Ai	O, A	ii, I^Ai
O X B	ii X I^BI^B or I^Bi	O, B	ii, I^Bi
O X AB	ii X I^AI^B	A, B	I^Ai, I^Bi
A X A	I^AI^A or I^Ai X I^AI^A or I^Ai	A, O	I^AI^A, ii
A X B	I^AI^A or I^Ai X I^BI^B or I^Bi	A, AB, O, B	I^Ai, I^AI^B, ii, I^Bi
A X AB	I^AI^A or I^Ai X I^AI^B	A, B, AB	I^AI^A, I^Bi, I^AI^B
B X B	I^BI^B or I^Bi X I^BI^B or I^Bi	B, O	I^BI^B, ii
B X AB	I^BI^B or I^Bi X I^AI^B	A, B, AB	I^AI^A, I^BI^B, I^AI^B
AB X AB	I^AI^B X I^AI^B	A, B, AB	I^AI^A, I^BI^B, I^AI^B

Table showing possible blood types of children from parents of various blood groups.

Table showing possible blood types of children from parents of various blood groups.

The 'Rhesus' Blood Group in Man:

A very interesting series of alleles affecting the antigens of human blood has been discovered through the work of Landsteiner, Wiener, Race, Levine, Sanger, Mourant & several others.

The original discovery was that the red cells are agglutinated by a serum prepared by immunizing rabbits against the blood of Rhesus monkey. The antigen responsible for this reaction was consequently called as Rhesus factor and the gene that causes this property was denoted as R-r or Rh-rh.

Interest in this factor was stimulated by Levine's study of a characteristic form of anemia, known as Erythroblastosis foetalis, which occurs occasionally in new born infants.

It was found that the infants suffering from this anemia are usually Rh-positive and so are their fathers; but their mothers are Rh-negative. The origin of the disease was explained as follows: The Rh^+foetus developing in the uterus of an Rh^- mother causes the formation of mother's blood stream of anti Rh antibodies.

These antibodies, especially as a result of a succession of several Rh^+pregnancies, gain sufficient strength in the mother's blood so that they may attack the red blood cells of the foetus. The reaction between these antibodies of the mother and the red cells of her unborn child provokes

haemolysis and anemia; this may be serious enough to cause the death of the newborn infant or abortion of the fetus.

The blood stream of a mother who has had an erythroblastotic infant is a much more potent and convenient reagent than sera of rabbits, immunized by blood of rhesus monkeys for testing the blood of other persons to distinguish Rh^+ from Rh^- individuals using such sera from woman who had erythroblastotic infants, it was discovered that: there exist not one but several kinds of Rh^+ and Rh- persons. There are several different Rh antigens which are detected by specific antisera.

Thus, an Rh^- woman immunized during pregnancy by the Rh^+children may have in her blood serum antibodies, that agglutinate not only Rh+ red cells but also cells from a few persons known to be Rh^-.

By selective absorption two kinds of antibodies may be separated from such a serum, one known as anti-D which agglutinates (= coagulates) only Rh^+ cells, the other known as anti-C which agglutinates particular rare types of Rh^-. Another specific antibody, known as anti-c agglutinates all cells that lack C.

With these three antisera, six types of blood can be recognized. Studies of parent and children show that persons of type Cc are heterozygous for an allele C determining C anti-gena. CC persons are homozygous for C and cc are homozygous for c. There is obviously no dominance, each allele producing its own antigen in the heterozygote as in the AB blood type.

No anti serum is available for detecting d, the alternative to D. D^+persons may be heterozygous or homozygous. However, the genotypes of such persons may be diagnosis from their progeny; for example D^+person who has a d^- child is thereby shown to be Dd.

Two other specific antibodies, anti-E and anti-c have been found. These detect the antigens E and e determined by a pair of alleles E and e. The three elementary types of antigens C-c, D and E-e, occur in fixed combinations that are always inherited together as alleles of a single gene. Wiener and Fisher showed the existence of a series of eight different alternative arrangements of these three types of Rh antigens and expressed them by means of following symbols.

The Rh System of Alleles:

Fisher's symbols	Wiener's symbols	
CDE	R^z	Rh-Positive
CDe	R^1	
cDE	R^2	
cDe	R^0	
CdE	r^Y	Rh-Negative
Cde	r'	
cdE	r''	
cde	r	

Thus, allelism is determined by cross-breeding experiments. If one gene behaves as dominant to another the conclusion is that they are alleles and that they occupy identical loci in homologous chromosomes when two genes behave as dominant to other gene. They should occupy identical loci in the chromosome. When more than a pair of alleles occur in respect of any character in inheritance the phenomenon is known as multiple allelism.

There is not much difference between the two theories of Wiener and Fisher. Wiener opinion is that there are multiple variations of one gene whereas according to the view of Fisher three different genes lying very close together are responsible for differences.

Theories of Allelism:

Various theories have been put forward to explain the nature of allelism origin and occurrence.

1. Theory of Point Mutation:

According to this theory multiple alleles have developed as a result of mutations occurring at same locus but in different directions. Hence all the different wing lengths of Drosophila are necessarily the result of mutations which have occurred at same long normal wing locus in different directions.

2. Theory of Close Linkage or Positional Pseudoallelism:

According to this view the multiple alleles are not the gene mutations at same locus but they occupy different loci closely situated in the chromosome. These genes closely linked at different loci are said to as pseudo alleles and affect the expression of their normal genes i.e., position effect.

3. Heterochromatin Theory of Allelism:

Occasionally heterochromatin becomes associated with the genes as a result of chromosomal breakage and rearrangement. These heterochromatin particles suppress the nature of genes in question due to position effect.

In maize the position effect are sometimes due to transposition (act of changing place or order) of very minute particles of heterochromatin. There are also sign or token that particles of different kinds of heterochromatin suppress the expression of normal gene to different degrees.

In Drosophila the apricot might be a partially suppressed red (normal) and white completely suppressed red while apricot and white hybrid may give rise to red or intermediate by unequal crossing over. The above theories in some way or other do not explain clearly the particular case of allelism and it is possible that all the three theories are applicable in different cases.

Importance of Multiple Allelism:

The study of multiple alleles has increased our knowledge of heredity. According to T.H. Morgan a great knowledge of the nature of gene has come from multiple alleles. These alleles suggest that a gene can mutate in different ways causing different effects. Multiple allelism also put forward the idea that different amounts of heterochromatin prevent the genes to different degree or space.

1. Pseudo alleles:

Alleles are different forms of the same gene located at the corresponding loci or the same locus. Sometimes it has been found that non-homologous genes which are situated at near but different loci affect the same character in the same manner as if they are different forms or alleles of the same gene. They are said as pseudo alleles. These pseudo alleles which are closely linked show re-combinations by crossing over unlike the alleles.

2. Penetrance and Expressivity:

Simply a recessive gene produces its phenotypic effect in homozygous condition and a dominant gene produces its phenotypic effect whether in homozygous or heterozygous condition. Some genes fail to produce their phenotypic effect when they should. The ability of a gene to produce its effect is called penetrance.

The percentage of penetrance may be altered by changing the environmental conditions such as moisture, light intensity, temperature etc. A gene that always produces the expected effect is said to have 100 percent penetrance. If its phenotypic effect is produced only 60 percent of the individuals that contains it then it is said to show 60 percent penetrance.

In Gossypium a mutant gene produces crinkled leaf. While all the leaves produced in the normal season are crinkled but some of the leaves which are produced late in the season do not show this character and are normal. It represents that penetrance is zero or in other words the gene is non-penetrant. Sometimes there is great variation in the manner in which a character is expressed in different plants.

In Lima beans there is a variety named venturra where a dominant gene is responsible for tips and margins of the leaves of the seedlings to be partially deficient in chlorophyll. Sometimes only the margins are effected and sometimes only the tips. In other words, this single gene may express itself in a variety of ways that may resemble a number of characters. This gene is then to exhibit variable expressivity.

Whether a gene is expressed at all is denoted by the term penetrance whereas the term expressivity denotes the degree of its expression.

3. Lsoalleles:

Sometimes, a dominant gene occurs in two or more forms. These multiple dominant alleles will produce the same phenotypic effect in homozygous condition but their effect will show a small difference in heterozygous state.

In Drosophila, thus, the gene for red eye colour is dominant over white. The red gene will produce dark red colour in the homozygous condition but in combination with the white allele the gene for red colour produces a dark red colour in flies from Soviet Russia but the same combination in the flies coming from the U.S.A. produces a light red colour. It does mean that dominant gene for red colour occurs in two forms. These are said as isoalleles.

4. Phenocopy:

Characters are the result of interaction between the genotype and the environment. When a gene mutates, its phenotypic effect also changes. Some times, a change in the environment produces a visible change in the phenotype of the normal gene which resembles the effect as already known mutant.

The effect of the normal gene under the changed environment is a mimic or imitation of the mutant gene. Such an imitation induced by environmental changes has been termed as phenocopy by Goldschmidt.

In fowls, a mutant gene is responsible for the character, ruinplessness, in which the caudal vertebrate and tail feathers do not develop. Rumplessness is also induced as a phenocopy when normal eggs which do not have the gene for rumplessness, are treated with insulin before incubation.

Phenocopies of other mutant genes are also produced in Drosophila by high temperature treatment of the larvae for short periods. It has also been found that different or non- allelic genes can produce the same phenotype. This phenomenon is said as genetic mimic or genocopy.

5. Xenia and Metaxenia:

The immediate effect of foreign pollen on visible characters of the endosperm is called xenia. The ‘xenia’ term was given by Focke (1800). This has been studied in maize plant. If a white

endosperm variety is open pollinated in the field where there are also plants of the yellow endosperm variety then the cobs that develop will contain a mixture of yellow and white seeds.

The yellow colour of the endosperm in the yellow seeds is the result of fertilization by pollen from the yellow variety. The yellow colour indicates that the seeds are hybrids and the white seeds are homozygous.

The yellow colour of the endosperm is dominant over white and when the plants raised from the yellow seeds are self-pollinated, yellow and white seeds are produced in the ratio of 3:1. Another example of xenia may be exemplified. If a sweet corn (maize) is pollinated by a starchy variety, the endosperm is starchy because the starchy gene introduced by the pollen is dominant over its sugary allele.

6. Metaxenia:

It is the term used to describe the effect of foreign pollen on other tissues belonging to the mother plant, outside the endosperm and embryo. It is sometimes evident in the fruit and seed coats.

In cucurbitaceous fruits, the skin colour is affected by the pollen grains; in oranges, the colour and flavor of the fruit is influenced by the pollen parent. The same is true of fuzziness and hair length in cotton. It has been suggested that metaxenia effects may be due to certain hormones secreted by the endosperm and embryo.

Lethal Allele

Alleles that cause an organism to die are called lethal alleles, where the gene involved is an essential gene. When Mendel's Laws were rediscovered, geneticists believed that mutations would only alter the appearance of a living organism. However, it was discovered that a mutant allele could cause death. When an essential gene is mutated, it can result in a lethal phenotype. If the mutation is caused by a dominant lethal allele, the homozygote and heterozygote for the allele will show the lethal phenotype. If the mutation is caused by a recessive lethal allele, the homozygote for the allele will have the lethal phenotype. Most lethal genes are recessive.

Examples of diseases caused by recessive lethal alleles are cystic fibrosis, Tay-Sachs disease, sickle-cell anemia, and brachydactyly. Huntington's disease is caused by a dominant lethal allele and even though it is not described as lethal, it is invariably lethal in that the victim experiences gradual neural degeneration and mental deterioration for some years before death occurs. One coat color of ranch foxes is caused by a recessive lethal gene. This gene causes a death if both recessive alleles are possessed by the same individual.

Recessive lethal alleles do not cause death in the heterozygous form because a certain threshold of protein output is maintained. In the homozygous form, the protein output does not meet the threshold, causing death.

Null allele

A null allele is a mutant copy of a gene at a locus that completely lacks that gene's normal function. This can be the result of the complete absence of the gene product (protein, RNA) at the molecular level, or the expression of a non-functional gene product. At the phenotypic level, a null allele is indistinguishable from a deletion of the entire locus.

A mutant allele that produces no protein is called a protein null (shown by western analysis), and one that produces no RNA is called an RNA null (shown by Northern analysis or by DNA sequencing of a deletion allele). A genetic null or amorphic allele has the same phenotype when homozygous as when heterozygous with a deficiency that disrupts the locus in question. A genetic null allele can be a protein and RNA null, but can also express normal levels of a gene product that is non-functional due to mutation.

Another definition of null allele concerning molecular markers, refers to such a marker in the case it can no longer be detected because of a mutation. For example, microsatellites(i.e. a repetitive sequence of DNA, in which the repeat is rather short) are used as molecular markers amplifying them through PCR. To do so, a primer or oligonucleotide aligns with either of ends of the locus. If a mutation occurs in the annealing site, then the marker can no longer be used and the allele is turned into a null allele.

One example of a null allele is the 'O' blood type allele in the human A, B and O blood type system. The alleles for the A-antigen and B-antigen are co-dominant, thus they are both phenotypically expressed if both are present. The allele for O blood type, however, is a mutated version of the allele for the A-antigen, with a single base pair change due to genetic mutation. The protein coded for by the O allele is enzymatically inactive and therefore the O allele is expressed phenotypically in homozygous OO individuals as the lack of any blood antigen. Thus we may consider the allele for the O blood type as a null allele.

Null alleles can have lethal effects. Mice homozygous for a null allele for insulin die 48 - 72 hours after birth.

Pleiotropy gene interaction

The opposite of polygene effect is known as pleiotropism i.e., a single gene influence or govern many characters. For example, gene for vestigial wing influence the nature of halters (modified balancers of Drosophila). The halters are not normal but reduced in flies with vestigial wings.

The vestigial gene also affects position of dorsal bristles which instead of being horizontal turn out to be vertical.

This gene also affects the shape of spermatheca i.e., the shape of spermatheca is changed ; the number of egg strings in the ovaries is decreased compared to normal when the vestigial larvae are well fed but relatively increased when they are poorly fed ; length of life and fruitfulness or fertility are lowered, and there are still other differences.

Epistatic and Non-Epistatic interaction

Due to the phenomenon of dominance a recessive allele remains obscure in the hybrid. But when two different genes which are not alleles, both affect the same character in such a way that the expression of one masks, inhibits or suppresses the expression of the other gene, it is called epistasis. The gene that suppresses is said to be epistatic, and the gene which remains obscure is hypostatic.

Types of Epistasis:

i. Dominant Epistasis:

In poultry white birds belong to two different varieties namely white leghorns or white Wyandotte's. Experiments reveal that the gene for white plumage of white leghorns is dominant over the gene for coloured plumage of coloured varieties. But the gene for white plumage of white Wyandotte's is recessive to the gene for coloured plumage of coloured varieties.

Therefore the gene which produces white plumage in white leghorns is different from the gene for white plumage in white Wyandotte's. A cross between a white leghorn and a white Wyandotte gives an F_1 of white birds with small dark flecks. When such birds are inbred, the F_2 progeny segregates in the ratio of 13 white to 3 coloured birds.

The experiment is explained below by postulating two genes C and I for the white leghorns:

P:	White Leghorn	×	White Wyandotte	
	IICC		*iicc*	
Gametes:	*IC*		*ic*	
	White with small			
F_1:	dark flecks	×	Inbred	
	IiCc			
F_2:	White	:	Coloured	
	13		3	

A checkerboard for the 16 phenotypes and genotypes of the F_2 birds indicates that only three out of sixteen genotypes, that is iiCC, iiCc, iiCc produce coloured birds. The white leghorns obviously contain a gene I, which in the dominant state inhibits or suppresses the expression of the dominant colour gene C, resulting in white plumage.

The recessive alleles of the inhibitor gene (ii) produce coloured birds due to expression of gene C. In other words gene I is epistatic to gene C. This is a case of dominant epistasis because even one dominant allele of gene I is able to express itself.

ii. Recessive Epistasis:

Epistasis due to recessive genes is called recessive epistasis. In mice albinism (white coat) is produced by a recessive gene aa. There is a different gene B which in the dominant state (BB and Bb) produces grey coat colour called agouti, and when recessive (bb) leads to black coat colour. The recessive gene for albinism (aa) is found to be epistatic to the gene for agouti (BB and Bb), and also to its recessive, homozygous allele (bb) for black. The presence of the dominant allele (AA) of the epistatic gene allows expression of gene B so that agouti (BB and Bb) and black (bb) coat colours can be produced (Fig. 2.2).

P: Agouti *AABB* × Albino *aabb*

Gametes: *AB* *ab*

F_1: Agouti *AaBb* × Agouti *AaBb*

Gametes: *AB*, *Ab*, *aB*, *ab*

	AB	*Ab*	*aB*	*ab*
A B	*AABB* Agouti	*AABb* Agouti	*AaBB* Agouti	*AaBb* Agouti
Ab	*AABb* Agouti	*AAbb* Black	*AaBb* Agouti	*Aabb* Black
aB	*AaBB* Agouti	*AaBb* Agouti	*aaBB* Albino	*aaBb* Albino
ab	*AaBb* Agouti	*Aabb* Black	*aaBb* Albino	*aabb* Albino

9 Agouti : 3 Black : 4 Albino

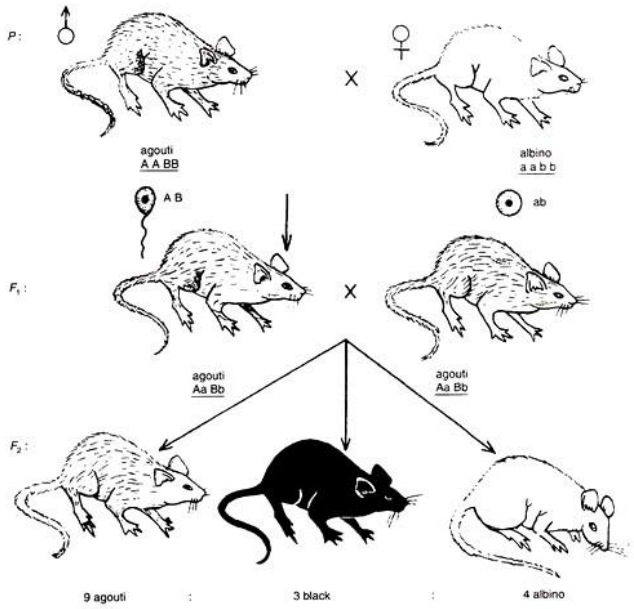


Fig. 2.2 Inheritance of the coat colour gene in mouse illustrating recessive epistasis.

The 9: 3: 4 ratio obtained is a modification of the classical 9: 3: 3 : 1 in which the last two classes (3: 1) are phenotypically identical and are therefore added up together. In human beings also the recessive gene for albinism shows epistasis in a similar manner. Epistatic effect is usually only in one direction, from one particular gene pair to another.

Epistasis in Drosophila:

There are two recessive wing mutants in Drosophila: apterous (ap) which produces small stubby wings instead of the normal transparent ones; the other called cubitus interrupts (ci) which causes a small interruption in the fourth longitudinal vein (Fig. 2.3). When the two mutants are used in a dihybrid cross, the F_2 progeny segregates in the ratio of 9 normal: 3 interrupted vein: 4 apterous wing.

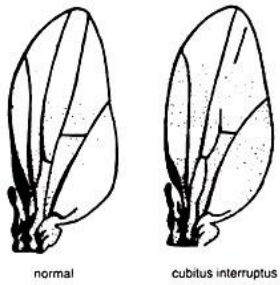


Fig. 2.3 Normal and interrupted veins in lateral wings of *Drosophila*.

P:	apterous	×		interrupted vein
	ap/ap+/+			*+/+Ci/Ci*
Gametes:	*ap+*			*+Ci*
F_1:		Wild type	×	Inbred
		ap/+Ci/+		
F_2:	9 Wild type	: 3 interrupted vein	:	4 apterous wing.

In the cross above, the presence of the homozygous recessive mutant gene for apterous wing (ap/ap) marks the expression of the gene for interrupted vein.

Epistasis and Blood Groups in Man:

In the ABO blood group system, a person with blood group A has antigen A on the surface of red blood cells; a blood group B person has B antigen: an AB person has both A and B antigens, whereas type O has neither A nor B antigen. The antigens are controlled by an autosomal gene I (iso-haemaglutinin) which has multiple alleles. Thus I^A controls antigen A, I^B controls antigen B. The recessive allele i produces no antigen and results in phenotype O.

Certain proteins present in the blood serum of a person show a precise relationship with the red cell antigen. These proteins are the agglutinins or antibodies. Thus the serum of an individual with blood group A has antibodies against antigen B, whereas the serum of a blood group B person is anti-A. In persons with blood group AB where both antigens A and B are present, the serum has neither antibody A nor B; individuals with blood group O have both types of antibodies.

It has been found that there is another gene H which controls production of a precursor substance H. In persons of blood type A, having gene I^A, precursor H gets converted into antigen A under control of gene I^A. Similarly when gene I^B is present, H gets converted to antigen B.

In the presence of the recessive allele i in type O persons, precursor H is not converted and remains as such in the blood. It can be agglutinated by anti-H or by such substances as an extract of the seeds of a plant Ulex europeus or blood of eel. In persons of heterozygous blood groups such as I^A/j or I^B/i only part of the H substance which is unconverted is agglutinated.

In AB persons there is likewise no agglutination of H substance. It is found that most human beings are homozygous (HH) for gene H, few are heterozygous (Hh), and very rarely a homozygous recessive (hh) individual is met with that cannot produce any H substance at all.

Such a person may have any blood group phenotype from the ABO system, but his red cells are not agglutinated by anti-A, anti-S, anti-H. First discovered in Bombay, India, a homozygous recessive (hh) person is said to show the Bombay phenotype. It demonstrates epistatic action of hh gene over I gene.

The secretor trait in humans also shows epistatic effect of genes. Just as red blood cells carry antigens on their surface, most persons have these antigens in a water soluble form in some body secretions such as saliva, gastric juice, fluids from nose, eyes and mammary glands of females. Such individuals are known as secretors. In a few persons regardless of blood group present, there are no antigens in body secretions. They are known as non-secretors.

The secretor trait is due to a dominant autosomal gene Se, whose recessive allele sd results in a non-secretor. The dihybrid inheritance pattern of secretor trait and blood groups has been studied in a large number of families. Marriages between two double heterozygotes for both blood groups and secretor trait, i.e., I^AiSese x I^AiSese or I^BiSese x I^BiSese show progeny phenotypes in the ratio 9:7 indicating epistatic interaction of genes.

$I^AiSese \times I^AiSese$

Gametes: I^ASe, I^Ase, iSe, ise

	I^ASe	I^Ase	iSe	ise
I^ASe	I^AI^ASeSe	I^AI^ASese	I^AiSeSe	I^AiSese
I^Ase	I^AI^ASese	I^AI^Asese	I^AiSese	I^Aisese
iSe	I^AiSeSe	I^AiSese	$iiSeSe$	$iiSese$
ise	I^AiSese	I^Aisese	$iiSese$	$iisese$

The checkerboard shows that there are 9 individuals showing both dominant phenotypes (blood group A secretors), 3 type 0 secretors and 4 non-secretors. The data indicate epistatic interaction of genes.

Interaction between gene(s) and environment

Most common diseases are a result of both your genes and your environment. Your environment can include personal choices, such as what foods you eat and how much you exercise, and external factors, such as stress, clean water, and air quality. Only a small number of diseases are a result of just a single mutation in a gene. Examples of these single-gene disorders are Huntington disease and Tay Sachs . Most diseases, especially common diseases, are a combination of your genetic risk and your environment. It is becoming difficult to group diseases into either purely 'genetic' or 'environmental' because most diseases are a little bit of both. For example, emphysema can be the result of both smoking and a disorder called alpha-1-AT deficiency . The field of research looking at gene-environment interactions (GxE) is growing.

It is important to understand that most times your genes do not determine your health. Small differences in your genetic makeup mean that two people can respond differently to the same environmental exposure. Here are some ways that your genes and your environment can interact:

- **Mutagens** – Mutagens are pollutants in the environment that enter the body and directly change your DNA sequence. Example: The chemicals in cigarette smoke can cause cancer.
- **Gene-gene interactions** – Gene-gene interactions occur when pollutants in the environment do not change your DNA sequence, but rather cause a chain reaction that affects the functioning of one gene that then affects the functioning of another gene. Example: Regularly drinking way too much alcohol can cause a specific gene, TACE, not to produce enough of its protein. TACE protein is supposed to help the MTHFR gene make enough of its protein. Too little MTHFR protein changes the level of folate (another protein) in our blood, and low folate levels may cause depression.
- **Transcription factors** – Pollutants in the environment can indirectly affect the DNA sequence by altering transcription factors, which are responsible for starting the process of using genes to make proteins that are needed for different functions in the body. Example: Stress can change the amount of proteins made by genes involved in your immune system and therefore, you may get sick more easily when you're stressed.
- **Epigenetics** – The environment can alter your health by affecting the proteins that turn genes on or off. Continue reading for more information on epigenetics. Example: half the genes that cause familial or inherited cancer are turned off when pollutants in the environment affect these proteins. Because they are turned off, these genes cannot suppress tumor formation or repair DNA.

Epigenetics

The epigenome is the primary location of gene-environment interactions and can be altered by the environment both directly and indirectly. It literally means "on top of or in addition to genetics," or basically factors outside of the genetic sequence. Epigenetic factors (most famously histone modification and DNA methylation) can

switch genes on or off and determine what proteins are transcribed. They are involved in many normal cellular processes and epigenetic changes are a natural part of human development. Some changes, however, can lead to disease. Some of these abnormal changes can lead to diseases such as:

- Cancer
- Mental retardation
- Neurodevelopmental disorders
- Cardiovascular diseases
- Type-2 diabetes
- Obesity
- Infertility

Epigenetics and the Environment

Some environmental exposures and dietary factors can lead to abnormal changes in epigenetic pathways. Because epigenetic changes are subtle and cumulative, it is difficult to know the true causal relationship between epigenetics and the environment.

Penetrance and Expressivity

Simply a recessive gene produces its phenotypic effect in homozygous condition and a dominant gene produces its phenotypic effect whether in homozygous or heterozygous condition. Some genes fail to produce their phenotypic effect when they should. The ability of a gene to produce its effect is called penetrance.

The percentage of penetrance may be altered by changing the environmental conditions such as moisture, light intensity, temperature etc. A gene that always produces the expected effect is said to have 100 percent penetrance. If its phenotypic effect is produced only 60 percent of the individuals that contains it then it is said to show 60 percent penetrance.

In Gossypium a mutant gene produces crinkled leaf. While all the leaves produced in the normal season are crinkled but some of the leaves which are produced late in the season do not show this character and are normal. It represents that penetrance is zero or in other words the gene is non-penetrant. Sometimes there is great variation in the manner in which a character is expressed in different plants.

In Lima beans there is a variety named venturra where a dominant gene is responsible for tips and margins of the leaves of the seedlings to be partially deficient in chlorophyll. Sometimes only the margins are effected and sometimes only the tips. In other words, this single gene may express itself in a variety of ways that may resemble a number of characters. This gene is then to exhibit variable expressivity.

Whether a gene is expressed at all is denoted by the term penetrance whereas the term expressivity denotes the degree of its expression.

Norm of reaction and phenocopy

The phenotypic distribution of a trait, as we have seen, is a function of the average differences between genotypes and of the variation between genotypically identical individuals. But both are in turn functions of the sequence of environments in which the organisms develop and live. For a given genotype, each environment will result in a given phenotype (for the moment, ignoring developmental noise). Then a distribution of environments will be reflected biologically as a distribution of phenotypes. The way in which the environmental distribution is transformed into the phenotypic distribution is determined by the norm of reaction, as shown in Figure 25-5. The horizontal axis is environment (say, temperature) and the vertical axis is phenotype (say, plant height). The norm of reaction curve for the genotype shows how each particular temperature results in a particular plant height. So, the dashed lines from the 18°C point on the temperature axis is reflected off the norm of reaction curve to a corresponding plant height on the vertical phenotype axis, and so forth for each temperature. If a large number of individuals develop at, say, 20°C, then a large number of individuals will have the phenotype that corresponds to 20°C, as shown by the dashed line; and, if only small numbers develop at 18°C, few plants will have the corresponding plant height. Then the frequency distribution of developmental environments will be reflected as a frequency distribution of phenotypes as determined by the shape of the norm of reaction curve. It is as if an observer, standing at the vertical phenotype axis, were seeing the environmental distribution, not directly, but reflected in the curved mirror of the norm of reaction. The shape of the curvature will determine how the environmental distribution is distorted on the phenotype axis. So, the norm of reaction in Figure 25-5 falls very rapidly at lower temperatures (the phenotype changes rapidly with small changes in temperature) but flattens out at higher temperatures, so the plant height is much less sensitive to temperature differences at the higher temperatures. The result is that the symmetric environmental distribution is converted into an asymmetric phenotype distribution with a long tail at the larger plant heights, corresponding to the lower temperatures.

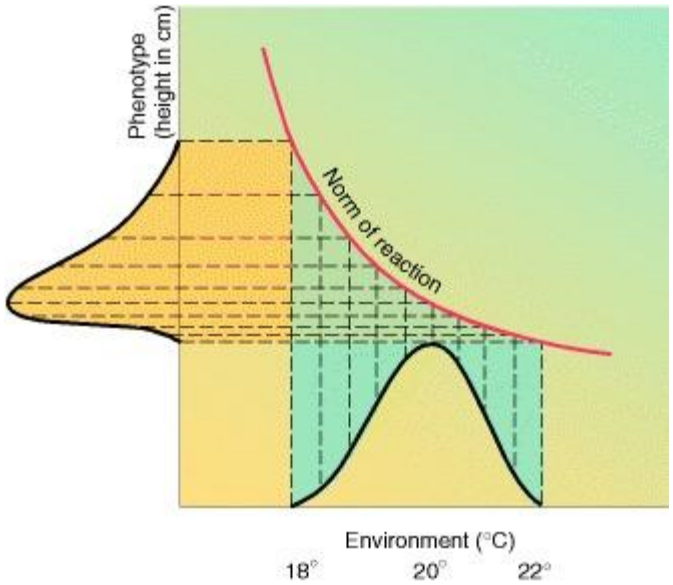


Figure 25-5

The distribution of environments on the horizontal axis is converted into the distribution of phenotypes on the vertical axis by the norm of reaction of a genotype.

By means of the same analysis, Figure 25-6 shows how a population consisting of two genotypes with different norms of reaction has a phenotypic distribution that depends on the distribution of environments. If the environments are distributed as shown by the black distribution curve, then the resulting population of plants will have a unimodal distribution, because the difference between genotypes is very small in this range of environments compared with the sensitivity of the norms of reaction to small changes in temperature. If the distribution of environments is shifted to the right, however, as shown by the gray distribution curve, a bimodal distribution of phenotypes results, because the norms of reaction are nearly flat in this environmental range but very different from each other.

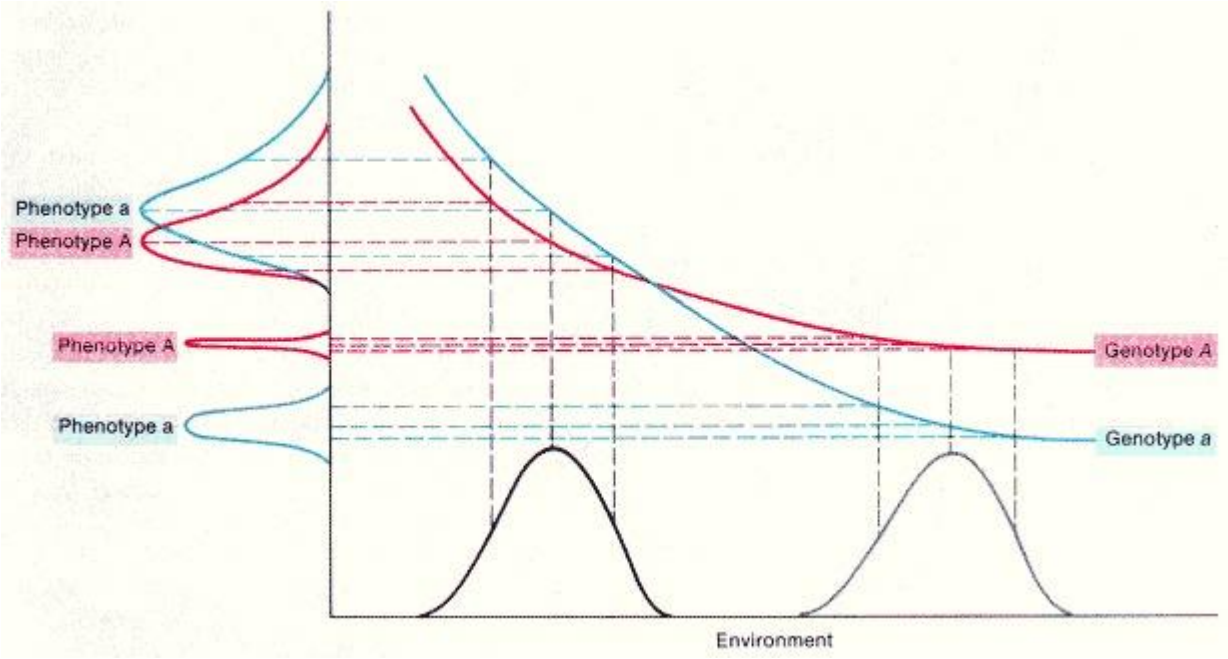


Figure 25-6

Two different environmental distributions are converted into different phenotypic distributions by two different genotypes.

Cis-Trans Test

a method of genetic analysis that determines whether recessive mutations occur in a single gene or in different genes. Thetest was devised by the American geneticist E. Lewis in 1951.
In conducting the cis-
trans test, the mutations under study are combined in the trans and cis positions. When the mutati -ons are combined in the trans position, two individuals, each with one of the mutations under analysis, are crossed. When themutations are combined in the cis position, an individual with both mutations is crossed with an individual of a wild (normal)type. In the trans test, that is, the functional test for allelism, if the mutations combined in the trans position belong to different genes, a hybrid organism is automatically obtained from an intact copy of each gene.
In this case the recessive mutationsare not manifested and the hybrid has a normal phenotype, that is, the mutations are complementary. If the combinedmutations belong to the same gene, both copies of the given gene in the hybrid are negatively affected and a mutant phenotype appears, that is, the mutations are non-complementary.
Lewis' trans test may be improved by studying mutations in the process of combining in the cis position as well, thuseliminating artifacts (atypical structures) caused by gene interaction at the

level of gene products. If in the cistrans test thehybrid's phenotype is the same in the cis and trans positions, that is, if there is no cistrans effect, the mutations under studyoccur within different genes. If, however, the hybrid's phenotype is different in the cis and trans positions, that is, if there is a cistrans effect, the mutations occur within the same gene.
In 1957 the American geneticist S. Benzer proposed that the unit determined by the cis-trans test be called the cistron.
Since it is difficult to combine in the cis position mutations that are situated closely together with in a chromosome, the cistrans test is rarely conducted. Research has revealed that in some cases several mutations of the same gene arecomplementary, and also that the cistrans effect may occur in mutations of different genes in a single operon. These
discoveries have reduced the theoretical value of both the trans test and the cis-trans test.
The trans test is still widely usedin applied genetic analysis to determine whether given mutations occur within a single gene (cistron).

Complementation Test

The production of wild type phenotype in a trans-heterozygote for two mutant alleles is termed as complementation and such a study is known as complementation test. The results obtained from complementation tests are highly precise and reliable and they permit an operational demarcation of gene.
Mutant alleles present in the same gene do not show complementation, while those located in different genes show complementation. Actually, this concept is generally true in prokaryotes but in eukaryotes several noteworthy exceptions are known.
The basis of complementation test (Fig. 15.3) may be simply described as follows. A gene produces its effect primarily by directing the production of an active enzyme or polypeptide. On the other hand, a mutant allele of this gene directs the production of an inactive form of the enzyme as a result of which it produces the mutant phenotype.
In the cis heterozygote, one of the two homologous chromosomes has the wild type allele(s) of the gene(s). This wild type allele will direct the synthesis of active enzyme—thereby producing the wild type phenotype.

Limitations of Cis-Trans Test:

This is an indirect experimental evidence to prove that a gene is sub-divisible. The standard phenotype, i.e., parental form, without any mutation, is called wild type. The genes present in the wild type organism is generally designated by '+' sign for comparison with mutant gene.
Before going to discuss the cis-trans test, it is reasonable to understand the meaning of cis and trans arrangement of gene. Cis arrangement means the condition in which a double heterozygote has received two linked mutations from one parent and their wild type alleles from the other parent, e.g., ab/ab x ++/.++-produces heterozygotes ab/++ (Fig. 15.1).

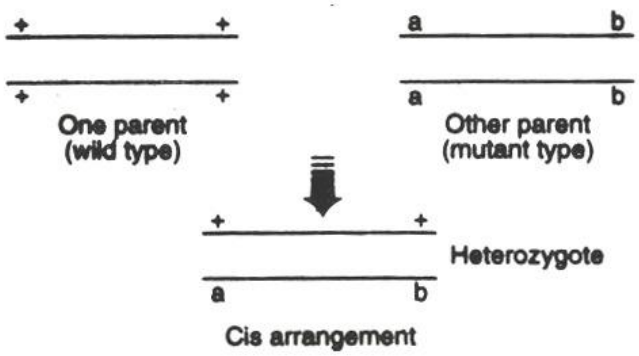


Fig. 15.1: Cis arrangement of gene.

Trans arrangement means the condition in which a double heterozygote has received a mutant and a wild type allele from each parent— for example a+/a+x+b/-l-b produces a+/+b (Fig. 15.2).

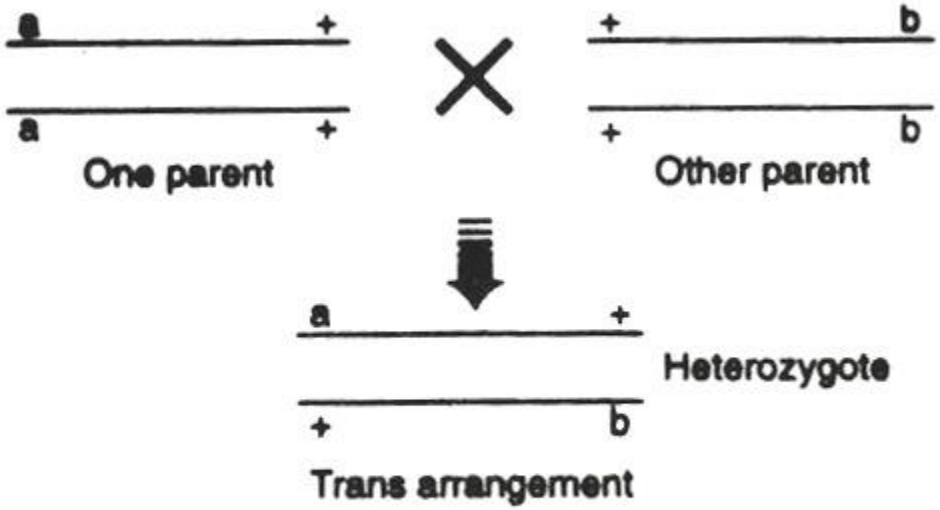


Fig. 15.2: Trans arrangement of gene.

In a cis-trans test the phenotypes produced in cis and trans heterozygotes for two mutant alleles are compared with each other. In a cis heterozygote, both mutant alleles are located in the same chromosome and their wild type alleles are present in the homologous chromosome, i.e., mutant alleles are linked in the coupling phase.

Thus it is expected to produce the wild type phenotype (unless the mutant alleles are dominant or co-dominant) irrespective of whether the two mutant, alleles are located in the same gene or in two different genes.

On the other hand, in case of trans heterozygotes one, mutant alleles are located in the homologous chromosome—they are linked in the repulsion phase. Hence, in trans heterozygotes, it is expected to produce the mutant phenotype if the two alleles are located in the same gene. But if they are located in two different genes, the wild type phenotype would be produced.

Hence simply by comparing the phenotypes for any two mutant alleles it is possible to determine if they are located in the same gene or in two different genes.

They are located in the same gene if their cis heterozygotes produce the wild type phenotype, while their trans heterozygotes have the mutant phenotype. But if both their cis and trans-heterozygotes have the wild type phenotype they are considered to be located in two different genes.

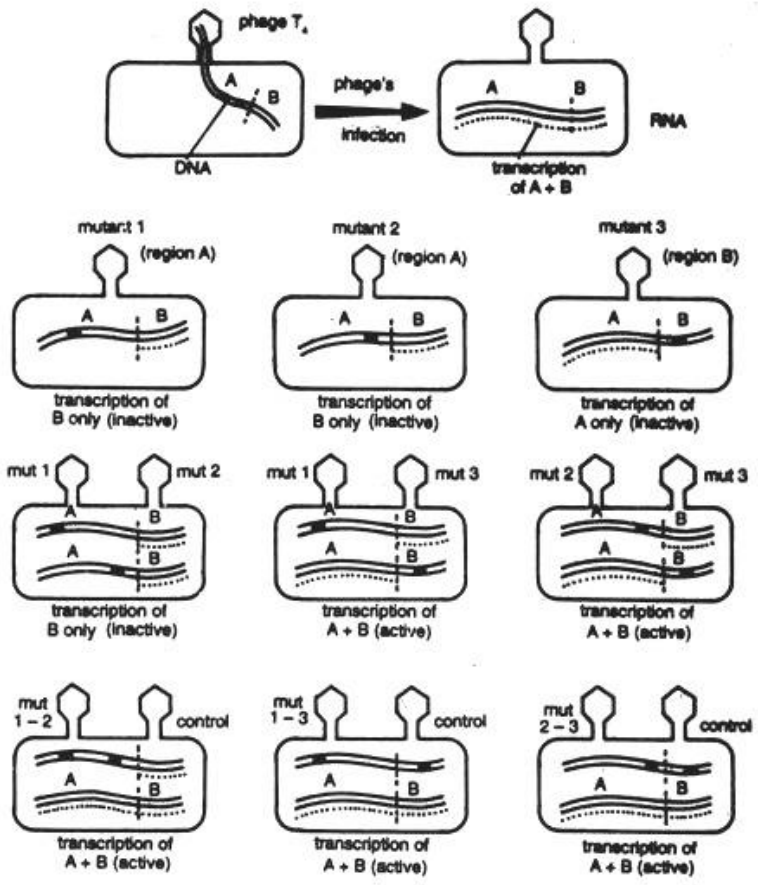


Fig. 15.3: Showing complementation test between mutants to find out if they belong to same or different cistrons.

In trans heterozygotes, if the mutant alleles are present in the same gene, the enzyme molecules produced by them will be inactive and capable of producing only the mutant phenotype. But if two mutant alleles are located in two different genes, one chromosome of trans heterozygote will have the wild type allele of the other gene.

Therefore, the trans heterozygote will have functional product of both the genes and the wild type phenotype will be produced by complementation. The complementation test has proven to be useful in delimiting genes.

But, in many cases, this test does not provide evidences to delimit gene.

These are:

a. Dominant or co-dominant mutation.

b. Genes in which mutations occur that show intragenic complementation.

c. Polar mutation, i.e., mutation that affects the expression of adjacent genes.

d. The gene in question does not produce a diffusible gene product, e.g., proteins.

Some other genes—such as operator and promotor genes which generally occur in the operon—do not code for a polypeptide or an enzyme. Hence they can act only in the cis position and they cannot show complementation. Therefore, such genes are called 'cis- acting gene.'

Intragenic complementation and rII locus of phage T4 and concept of cistron

The most extensive study on the fine structure of gene was undertaken by Seymour Benzer for a locus in T_4 bacteriophage infecting E. coli. This locus is known as rπ locus.

T4 bacteriophage contains a linear molecule of DNA of about 200,000 base pair long which is packed within its head (Fig. 15.4). When T_4bacteriophage infects E. coli the bacterial cell lyses in about 20-25 minutes releasing 200-300 progeny phage particles.

When the inoculum of E. coli cells are plated in a petridish containing semi-solid nutrient medium, it will produce an uniform confluent growth or lawn on the surface of nutrient medium after certain period of incubation at the appropriate temperature [Fig. 15.5(a)], If the isolated T_4 bacteriophage particles are placed at different sites on the surface of bacterial lawns, T4 bacteriophage infect the bacterial cell and all the E. coli cells in the immediate surrounding vicinity of phage will be destroyed.

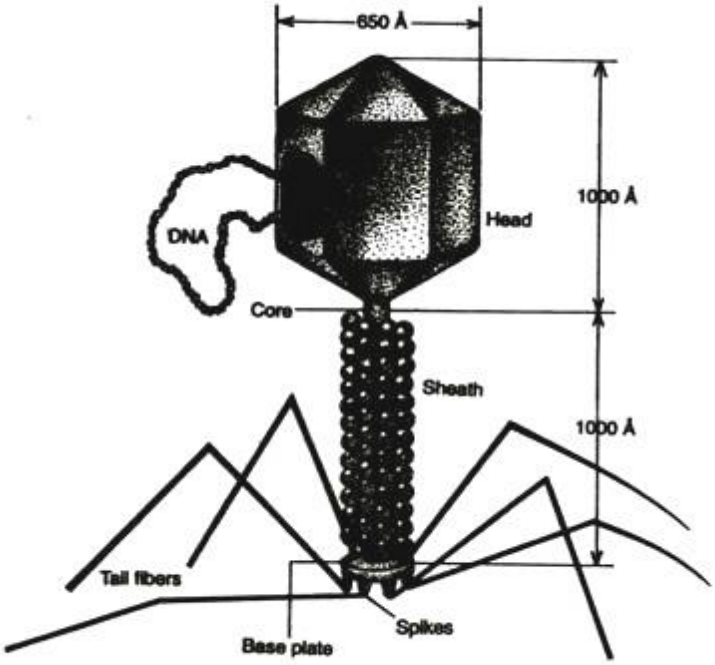


Fig. 15.4: Diagram of the morphology of T_4 bacteriophage.

This leads to the development of a clear area in the bacterial lawn. The clear areas are called plaques which indicate the areas of infection and lysis of bacterial cell due to infection by phage and is characteristic of phage.

The plaques are surrounded by a fuzzy or turbid margin called halos which are produced due to a phenomenon called lysis inhibition [Fig. 15.5(b)], It is a delay in lysis of T_4 infected E. coli cells as a consequence of a subsequent infection by another T_4 particle. The ability of T_4 phage to cause lysis of bacterial cell is controlled by gene(s) present in a specific locus called 'r' locus (r = rapid lysis).

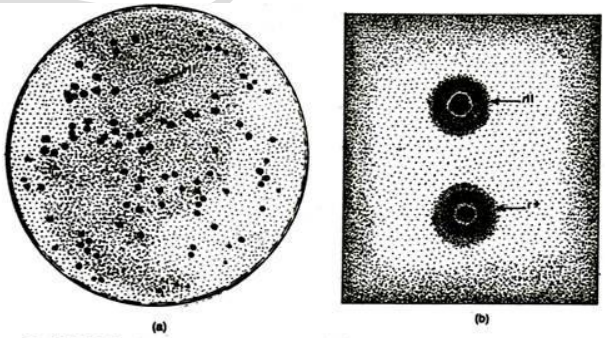


Fig. 15.5: (a) Showing the confluent growth or lawn of E. coli on the surface of nutrient medium; (b) Showing the sharp margin on the r^{II} plaque (above) and the fuzzy margin on the r^{+} plaque.

S. Benger isolated over thousands of independent mutant strains carrying mutation in the r locus (Fig. 15.6). Most of the r mutants map and classify into three distinct loci called $r^{I}r^{II}$ and r^{III}. The mutants can be recognised to some extent on the basis of the morphology of plaques, the ability of mutants to cause lysis of bacterial cell.

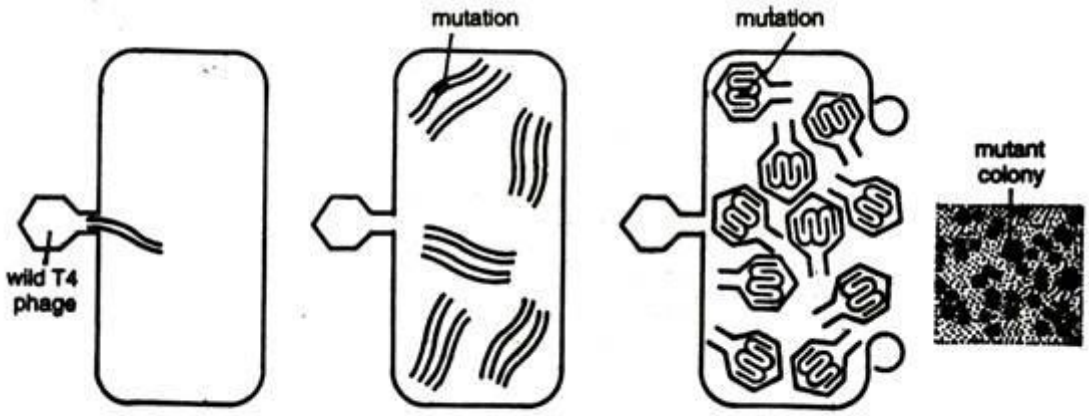


Fig. 15.6: Showing the production of a r^{II} mutant of T_4 phage after infection of E. coli by wild T_4 phage.

Mutants in the r^{II} locus are easily recognised due to their inability to multiply in E. coli strain $K_{12}(\lambda)$ which has the chromosome of phage A integrated in its chromosome. However, r^{II} mutants grow rapidly in other strains of E. coli such as strain B and strain K_{i2} lacking the λ chromosome.

The wild phage T_4 r^{II+} makes small and fuzzy plaques both on B and K strains, whereas the r^{II} mutants make large sharp plaques on E. coli strain B and K strains (Fig. 15.7). These distinguishable properties enabled Benzer to distinguish mutants and wild type phage with high efficiency. The r^{II} mutants axe conditional lethals unable to grow in $K_{12}(\lambda)$; this property was exploited by Benzer for a fine genetic analysis of the r^{II} locus.

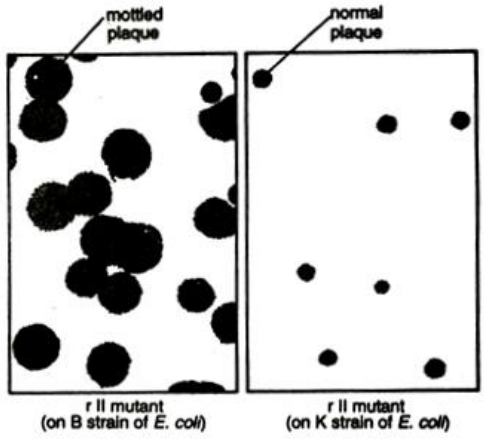


Fig. 15.7: Showing plaques formed by r^{II} mutant strain on B and K_{12} (lacking the λ chromosome) strain of E. coli.

Benzer isolated over 3,000 independent mutants of the r^{II} locus and subjected them to complementation test. Phage carrying r^{II} mutation can be easily identified by sterile toothpick transfers of phage from individual plaques growing on E. coli strain B (r^{II}-permissive)
"Lawns" to lawn of E. coli strain K_{12} (λ) (r^{II} restrictive) and lawns of E. coli strain B (Fig. 15.8). Each plaque to be tested (left side of Fig. 15.8) is stabled with a sterile toothpick which is subsequently touched to maxked axea in a petridish with a K_{12}(λ) lawn (in the center of Fig. 15.8) and then to an identically marked area in a dish with an E. coli B lawn (right side of Fig. 15.8).
Mutants that fail to grow (are lethal) on K_{12}(λ) (left side of the centre plate) can be recovered from the plaques on the E. coli B plates (right side of the Fig. 15.8).

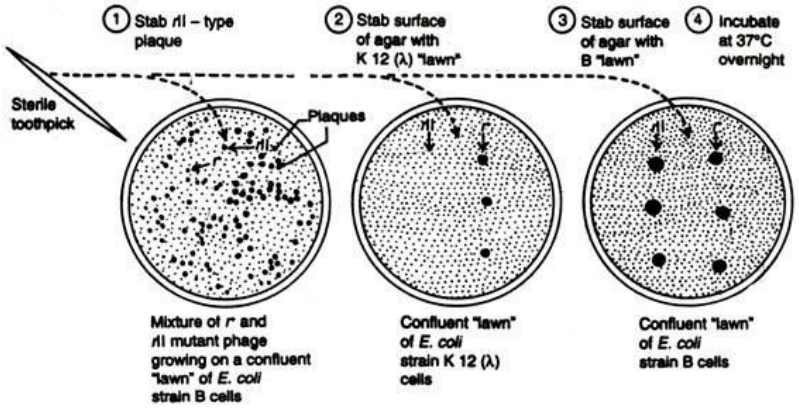


Fig. 15.8: Identification of r^{II} mutants by serially transferring inocula with sterile toothpicks from r-type plaques on E. coli B lawn to lawn of E. coli K_{12}(λ) and E. coli B. All R^{II} mutants produces r-type plaques on E. coli B, but will not grow on E. coli K_{12}(λ).

If plaques develop on the E. coli B lawn, it indicates complementation between the two r^{II} mutants used for co-infection, while an absence of plaques signifies a lack of complementation. Mutants at the r^{I} and r^{III} loci as well as r^{+} phage (right side of the central plate) will grow on both K_{12}(λ) and B. Benzer placed all r^{II} mutants in two arbitrary groups named be A and B.
All the r" mutations were found to located in one of the two genes of cistron. Benzer designated these two genes r^{II}A and r^{II}B (Fig. 15.9).

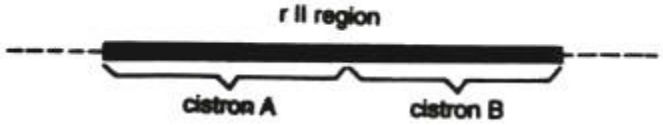


Fig. 15.9: r^{II} region of T_4 phage showing cistron A and cistron B.

The r^{II}A region appears to consist of about 2,000 deoxyribonucleotide pairs. The A region transcribes a messenger RNA that translates an A polypeptide; the B region is similarly

responsible for a B polypeptide. B polypeptides are needed for lysis of K type E. coli cells. The wild type (r^+) phages produces both A and B polypeptides. A mutant produces normal B polypeptide but not A, and vice versa.

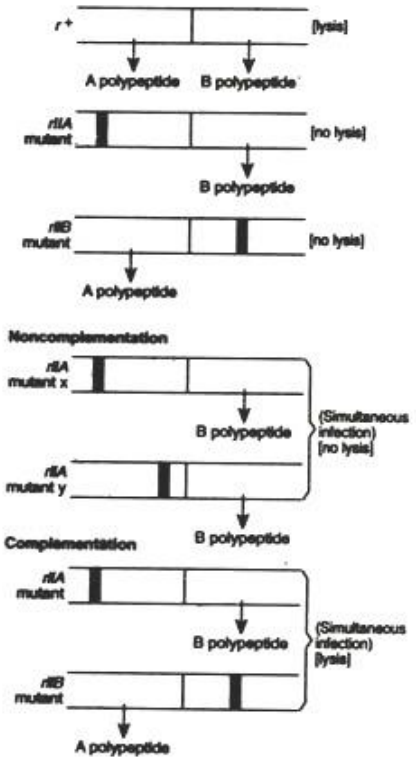


Fig. 15.10: Complementation occurs when Escherichia coli is infected simultaneously with an rIIA and an rIIB mutant phage; complementation does not occur if two different A mutants, for example, simultaneously infect the host. For reference, R^+, rIIA, and rIIB mutants are diagrammed also. Diagrams are not to any scale.

Hence infection only by identical r^{II} A mutants or by identical r^{II} B mutant alone can cause lysis of the host cells, because none of the phages can produce both A and B polypeptide (Fig. 15.10). On the other hand, infection by two different mutants (one an r^{II} A mutant and the other an r^{II} B mutant) on the same host cell does result in lysis (Fig. 15.11). It indicates that regions A and B are functionally different and show complementation.

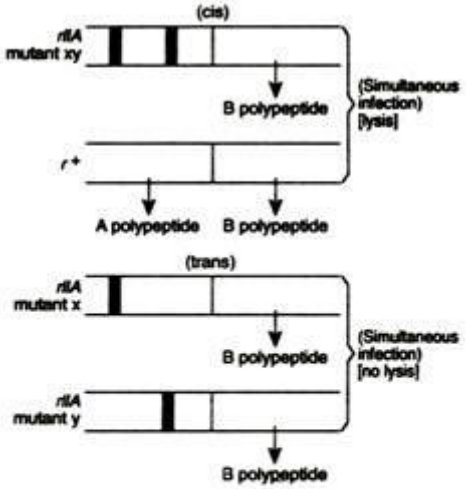


Fig. 15.11: Simultaneous infection of Escherichia coli by rIIA (or rIIB) double mutants and wild type r^+ phages (top diagram), where the cistron defects of the mutant are in the cis position (i.e., in the same DNA molecule), results in lysis. When the mutations are in the transpositions (lower diagram), i.e., in different DNA molecules, no lysis takes place.

Benzer observed that with infection by two phages—one the wild type (r^+) and the other mutant in either A or B region, i.e., with mutation in the cis position—lysis occurred. But the lysis did not occur when the mutation A or B were in the trans configuration.

Thus, it was clear that mutation in one functional region (A or B) Eire complementary only to mutations in the other region and complementation is detectable by cis-trans test.

Each functional region is responsible for the production of a given polypeptide chain. Benzer defined the functional unit as cistron and conformed operationally more closely to what we commonly think as gene. This cistron, therefore, may be thought of as the gene at the functional level. There can be over a hundred points within a functional unit wherein a mutation can take place and cause a detectable phenotypic effect.

This means that a cistron is over hundred nucleotide pairs in length and there is some evidence that some cistrons may be as long as 30,000 nucleotide pairs. Actually each cistron represents a part of a gene which is responsible for coding of only one polypeptide chain of an enzyme that has two or more different polypeptide chains in its complete enzymatic unit.

A cistron also includes initiating, terminating and any un-transcribed nucleotides.

(a) The Muton:

It is the smallest unit of DNA which, when altered, can give rise to a mutation. Study of the genetic code makes it clear that an alteration of a single nucleotide pair in DNA may result in a missense codon in transcribed mRNA (e.g., AGC—>AGA) or nonsense (e.g., UGC—> UGA).

So a cistron may be expected to consist of many mutable units or mutons. The term muton was given by Benzer.

(b) The Recon:

It is the smallest part of DNA which is interchangeable through crossing over and recombination. Extremely delicate studies of recombination in E. coli indicate that a recon consists of not more than two pairs of nucleotides, may be only one.

A recon may occur within a cistron. Thus a gene of classical concept is made up of a number of functional units—the cistrons— which consist of a number of recons and mutons (Fig. 15.12).

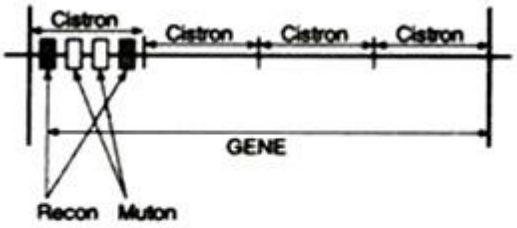


Fig. 15.12: Schematic representation of a fine structure of gene.

(i) Recombination Frequency:

The complementation test shows that all the r^{II} mutants were located within A and B cistrons. In order to estimate the frequency of recombination between r^{II} mutants, E. coli strains B cells are infected with a mixture of the two r^{II} mutants.

If the crossing occurs between two chromosomes of mutant strain it yields one wild type and one double mutant type for each crossing over event (Fig. 15.13). Therefore, some of the progeny phage present in the lysate of the B strain (infected by a mixture of two r^{II} mutant) would be of wild type.

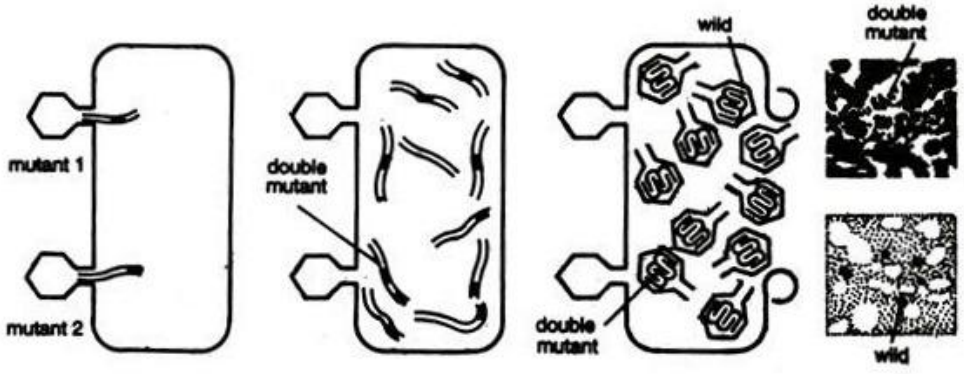


Fig. 15.13: Production of double mutant of rII locus due to simultaneous infection of E. coli by two different mutants.

The frequency of the wild type phage in the lysate is determined by plating lysate on the lawn of $K_{12}(\lambda)$ strain. Each wild type phage would produce a plaque on this lawn. This is a highly efficient selection system for wild type phage and as many as 10® progeny phage may be examined in a single petridish.

The number of plaques produce on $Kj_2(\lambda)$ represents the number of wild type phage particles in the lysate. A equal number of phage would have the double mutant produced due to recombination. Therefore, the number of recombination phage in the lysate would be twice the number of plaques produced on $Ki_2(\lambda)$.

$$\text{Frequency of the recombination} = \frac{2 \times \text{the number of plaques of } K_{12}(\lambda)}{\text{No. of plaques on B}}$$

Thus, the frequency of recombination may be measured as follows:

However, to map 3,000 mutations by only standard recombination test is a highly laborious task and is practically impossible because the number of all possible two-point crosses only (infection of E. coli cells by mixture of two mutants at a time) will be about 4 ½ illion, i.e., 3,000 x 2,999/2. Hence Benzer was able to avoid such a laborious undertaking by developing a shortcut method of mapping that used overlapping deletion mutation. This technique is known as deletion mapping. It permits the deletion of recombination value of 0.0001 or even 0.00001%.

(ii) Deletion Mapping:

Benzer first mapped a number of r^{II} mutants using the data of recombination test. He noted that some of these mutants did not show recombination with some other r^{II} mutants. These mutants also failed to undergo reverse mutation, i.e., mutation to wild type r^+.

Benzer classified these non-reverting, non-recombining r^{II} mutations as deletion mutation. Benzer also proposed that these deletion mutations (multisite mutations) resulted from the deletion or loss of segments of DNA. These deletions were arranged in sets of overlapping deletions representing segments of different sizes in r^{II}regions as shown in Fig. 15.14.

The principle involved in this method was that if a particular mutation presents in the region of a deletion represented by a r^{II} mutant, then, on mixed infection with this deletion mutant, the point mutation will not be able to give rise to wild type, but if it falls outside the deletion regions it will be able to give rise to wild type and recombinant type.

The extents of the deleted segments can be analysed by crossing the deletion mutants to a set of reference point mutations which are previously mapped. Once a set of overlapping deletion has been mapped, their end- point will divide the region resolved by the longest deletion in a set of intervals A, B, C, D (Fig. 15.15).

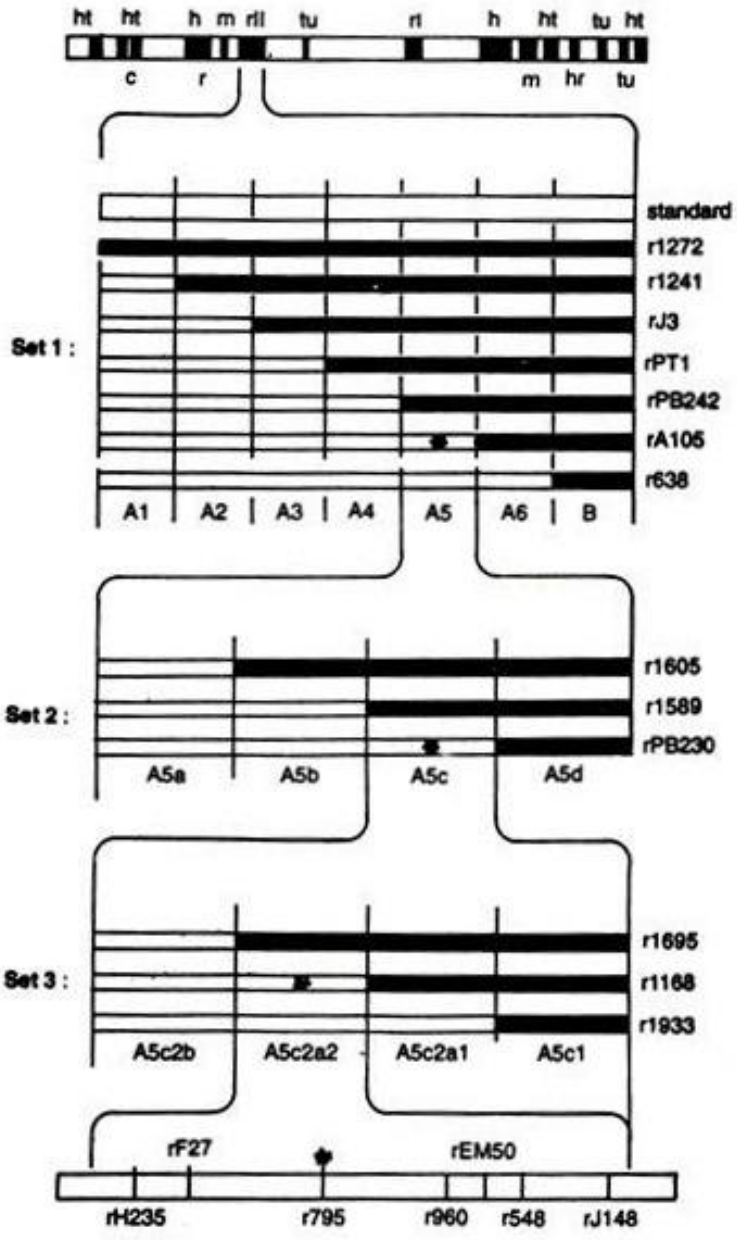


Fig. 15.14: Showing different overlapping deletion (non-revertant) mutants at rII locus in T_4 phage and their use in locating the position of revertant (point) mutation (marked as *).

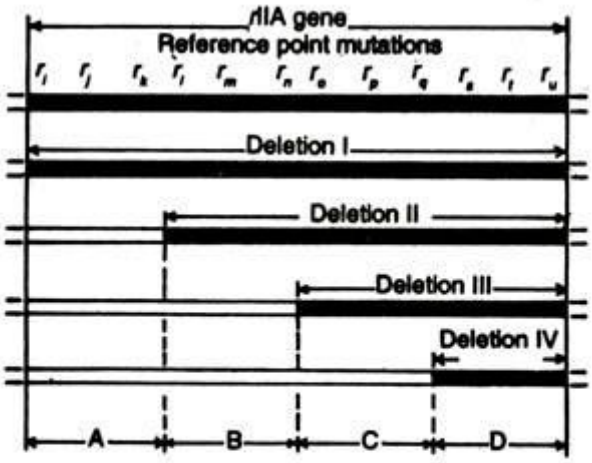


Fig. 15.15: Schematic diagram showing the use of overlapping deletion in genetic fine structure mapping.

When an unknown new mutant carrying a point mutation is isolated, the mutant can immediately be mapped to a defined interval by crossing the mutant with each of the overlapping deletion mutants. A mutant in interval D will not produce any wild type recombinant progeny in any of the four crosses. A mutation in interval C will recombine with deletion IV (Fig. 15.15) but not with the other three deletions, and so on.

Mutants in the region (Fig. 15.17)	Recombination with the deletion mutant of the r^{II} locus						
	1272	1241	J3	PT 1	PB 242	A 105	638
A_1	–	+	+	+	+	+	+
A_2	–	–	+	+	+	+	+
A_3	–	–	–	+	+	+	+
A_4	–	–	–	–	+	+	+
A_5	–	–	–	–	–	+	+
A_6	–	–	–	–	–	–	+
B	–	–	–	–	–	–	–

Note: + denotes recombination, which—signifies a lack of recombination. A_1, A_2 etc. represent segments of the cistron A of r^{II} locus, while B is the cistron B of this locus.

In this manner, Benzer characterized with deleted segments of a large number of r^{II} deletion mutants. This permitted him to divide the entire r^{II} locus into 47 small segments (Fig. 15.16). A set of seven of these deletion mutants permitted him to divide the r^{II} locus into 7 regions like $A_1 – A_6$ and B.

Each new r^{II} mutant to be mapped was crossed pair wise with each of these seven deletion mutants and the presence of wild type (recombinants) phage particles counted in the progeny.

On the basis of this data a new r^{II} mutant was localized in one of the- seven segments. (Table 15.1 and Fig. 15.17.) Once an unknown r^{II}mutant is pointed in a segment, it is crossed to another set of deletion mutant's which allows its localization in a smaller sub-division of that segment.

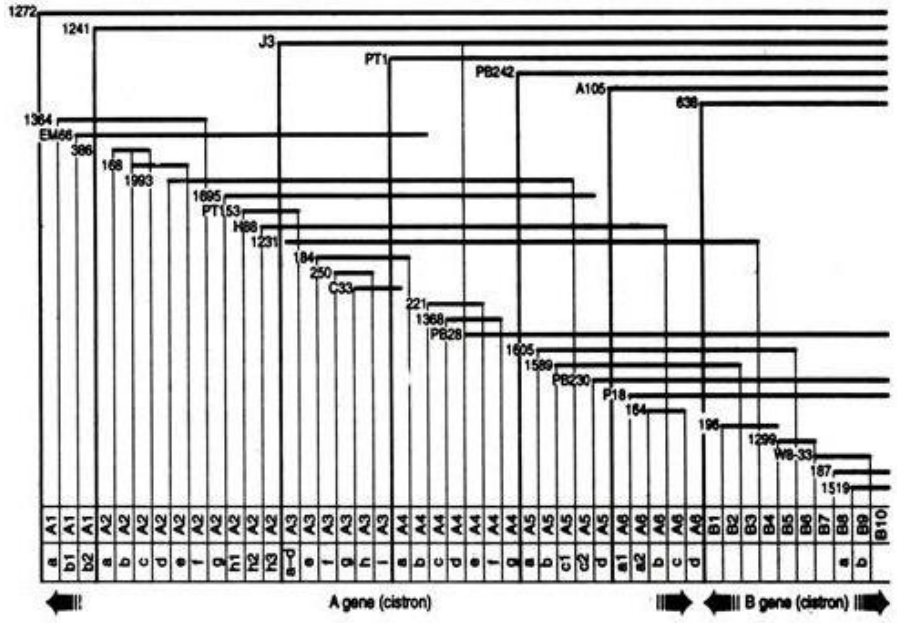


Fig. 15.16: The overlapping deletions used by Benzer and coworkers to divide the rII locus into 47 smaller intervals. Benzer's designations for each of the 47 segments are given in the boxes near the bottom. The limits of the A and B genes, ad defined by complementation tests, are shown by the arrows at the bottom. (After S. Bener, Proc. Natl. Acad. Sci. U.S.A. 47: 410, 1961.)

The final mapping of r^{II} mutants is done on the basis of recombination data from two and point crosses among mutants located within the concerned subsection of the r^{II} locus. Benzer et al identified more than 300 sites of mutation that were separable by recombination. The progeny of mutation at different sites is highly variable.

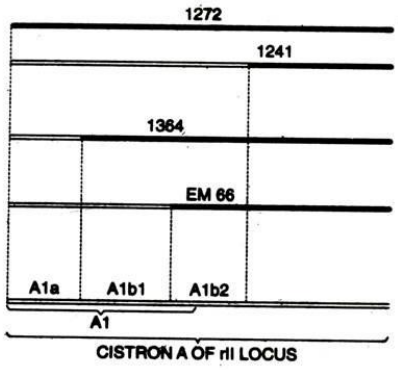


Fig. 15.17: Subdivision of the section A1 (delineated by mutants 1272 and 1247) of the r^{II} locus of phage T_4 (shaded areas show the regions of deletion in the chromosomes).

Electron Microscope Heteroduplex Mapping:

The presence of genetically well-defined deletion mutations at the r^{II}locus can also be determined using a technique called heteroduplex mapping. A DNA heteroduplex is a DNA molecule in which the two strands are not complementary.

One strand of a DNA double helix may contain one allele of the gene and the other strand may not be totally complementary and may carry of different alleles of the gene. The non complementary portions of DNA then form a heteroduplex which may vary in size from one mismatched base pair to large segments of the molecule.

Heteroduplex mapping involves in vitro preparation of DNA hetero-duplexes and their analysis by electron microscope. The heteroduplex may be prepared by mixing the denatured single-stranded DNA segments of wild type and mutant type followed by DNA renaturation.

The prepared hetero-duplexes between DNA from T_4r^+ phage and DNA from each of several genetically well characterised r^{II} deletion mutants. Thereafter they are analyzed by electron microscope.

The results obtained estimates of 1,800±70 nucleotide pairs and 845 ± 50 nucleotide-pairs for the sizes of the r^{II} A and r^{II} B genes, respectively. These results combined with the extensive genetic data of Benzer et al provide a fairly clear picture of the fine structure of the r^{II} locus.

(c) Overlapping Genes:

The presence of overlapping gene provides an interesting information for the study of fine structure of a gene. It is generally accepted that the boundaries of neighbouring genes do not overlap.

The study of nucleotide sequences of φ x 174 bacteriophage has clearly resolved that out of total 10 genes of φ x 174, two are located entirely within the coding sequences of two different genes. A third overlaps the sequences of three different genes. This surprising result has important genetic implication for the study of fine structure of gene.

(d) Fine Structure of Genes in Eukaryotes:

Complementation and recombination study have been used to prepare fine structure maps of several eukaryotic genes. By this technique, genetic fine structure maps have now been constructed for many genes of Drosophila; maps have also been worked out for several other higher animals and higher plants.

One of the best examples of such a gene is the rosy (ry) eye locus of Drosophila which codes for the enzyme xanthine dehydrogenase. The different alleles of ry locus map at 10 different sites on the basis of recombination frequency (Fig. 15.18).

Many of the ry mutants do not show complementation (shown in upper line of the Fig. 18.18) while several others show complementary (shown in the lower line of the figure). The complementary allele may be located at the same site or at a site very close to one where non-complementary alleles are located.

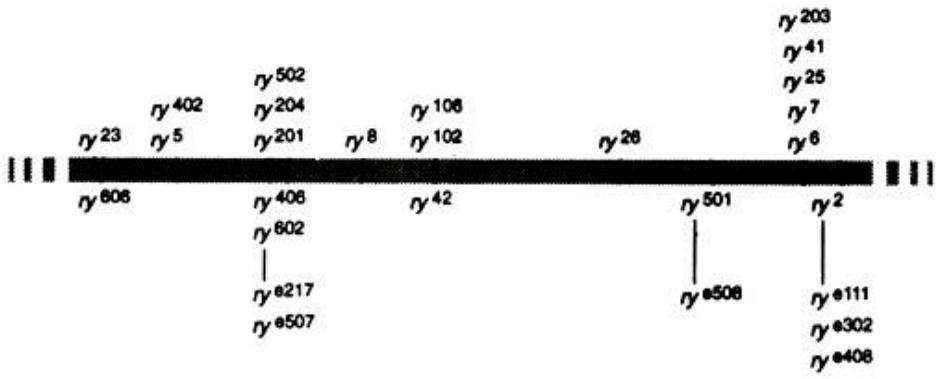


Fig. 15.18: Map of the rosy locus of Drosophila melanogaster, showing the map position of mutations in the structural gene for the enzyme xanthine dehydrogenase. Noncomplementing xanthine dehydrogenase-deficient ry mutations are shown above the line. Mutations that exhibit intragenic complementation are shown directly below the line.

The complementation of ry alleles is a case of intragenic complementation. The recovery of wild type recombination is very easy because rosy mutants are conditional lethals. As a result, wild type recombination produced from the heterozygotes for ry^2 ry^3alleles can be easily isolated and counted by growing their progeny on a purine supplemented medium, on this medium only wild type progeny would survive.

Besides rosy locus, fine structure maps of gene of many other eukaryotes have been prepared like white (w) eye, notch (N) wing, lozenge (lz) eye, zetse (another eye colour locus near the white locus) etc. loci of Drosophila, waxy (wx) and other loci of maize, some loci of yeast.

some limitations due to:

i. Examining enough progeny of a cross to detect rare intragenic recombination in eukaryotes is a laborious job.

ii. In many cases, determining how many genes are present at a locus has proven difficult in eukaryotes. This problem arises due to the presence of complex loci.

iii. Complementation tests have often yielded ambiguous results—due to the occurrence of intragenic complementation. Delimiting genes of eukaryotes by complementation test should be done, whenever possible, using amorphic or null mutation (mutation resulting in no gene product) to minimize the possibility of confounding effects of intragenic complementation.

In eukaryotes, however, some genes have interesting structural features which are not found in most prokaryotes. Therefore, our view of fine structure of any gene as discussed earlier may be partly ambiguous due to use of a specific recombination system. Further, the distances between genes on a genetic map may not correspond to the distances between them in the DNA molecule of which they are a part at the molecular level. There may also be present gaps or a genetic map due to non-availability of mutants in that region. At the molecular level the fine structure of a gene can be resolved by modern genetic mapping through determination of nucleotide sequence of the concerned DNA segment. Alternatively we can prepare a genetic map by breaking the

DNA at specific sites with the help of restriction endonucleases which are specific in recognizing very short DNA sequences and cutting the DNA at these specific sites. These sites of breakage can be identified and mapped in eukaryotes. This modern technique for in prokaryotes to give rise to a restriction genetic mapping at the molecular level has been map.

Mechanism of genetic exchange

The process of transfer of genetic material and recombination is very interesting bacterial recombination is given (PPT. an overview of bacterial recombination).The three main mechanisms by which bacteria acquire new DNA are transformation, conjugation, and transduction. Transformation involves acquisition of DNA from the environment, conjugation involves acquisition of DNA directly from another bacterium, and transduction involves acquisition of bacterial DNA via a bacteriophage intermediate.

Transformation

Transformation is the process by which bacteria pick up DNA from their environment. The DNA may come from a variety of sources, but most likely it is the remnants of DNA from dead bacterial cells. In order to become successfully transformed, bacteria must be competent. This means that the bacteria are expressing the appropriate enzymes (the 'transformation machinery') required to transport the exogenous DNA into the cell. Therefore, the correct genes must be expressed in order to carry out transformation. Expression of these genes depends on the growth conditions: bacteria most likely to be competent are dividing rapidly, but nutrients in the environment are becoming limited. (For more on the control of gene expression, see the module on bacterial gene regulation. In transformation, a cell surface receptor binds to DNA in the environment. After binding, the DNA is transported across the membrane by the transformation machinery. As this occurs, one strand of the DNA is digested away by an exonuclease, so that the DNA that enters the cell is single stranded. This promotes recombination, as long as the DNA taken up is sufficiently homologous to the host DNA to allow recombination to occur. The recombination that occurs is one-way (non-reciprocal); unlike the exchange of strands diagrammed in the module on recombination, in this case the new DNA will simply replace a strand of the host DNA. The replaced segment of host DNA will be degraded. If the new DNA is of a different allelic nature than the host DNA, a gene conversion event can occur. This is what happened in the example mentioned above: the avirulent strain of S. pneumoniae had a mutation in a gene required for production of the bacterial capsule. Heat killing the virulent cells (which contained the wild-type capsule gene) caused the release of fragments of the dead cells' genomes.

Some of the avirulent cells picked up a piece of DNA containing the wild-type capsule gene, and underwent gene conversion so that they were wild type for that gene, causing them to become virulent.

Conjugation

Conjugation is a mating process involving bacteria. It involves transfer of genetic information from one bacterial cell to another, and requires physical contact between the

two bacteria involved. The contact between the cells is via a protein tube called an F or sex pilus, which is also the conduit for the transfer of the genetic material. Basic conjugation involves two strains of bacteria: F+and F-. The difference between these two strains is the presence of a Fertility factor (or F factor) in the F+ cells. The F factor is an episome that contains 19 genes and confers the ability to conjugate upon its host cell.

Genetic transfer in conjugation is from an F+ cell to an F-cell, and the genetic material transferred is the F factor itself. Here is an overview of the process: Basic conjugation occurs between an F+ cell and an F-cell. The difference between these two types of cells is the presence or absence of the F (fertility) factor, which is a circular DNA molecule independent of the bacterial chromosome (the larger circular molecule. The F+ cell initiates conjugation by extending an F pilus toward the F-cell. Among the genes present on the F factor are the genes encoding the proteins required for pilus construction. The F pilus, when finished, temporarily connects the two cells. On strand of the F factor is nicked, and begins unwinding from the other strand. The nicked strand begins to transfer through the F pilus to the F-cell. As it does so, this strand begins to be replicated, as does circular strand remaining behind in the F+ cell. Eventually, the nicked strand completely passes through to the recipient cell, and is completely replicated. This process produces a new F factor in the recipient cell. The pilus is broken, severing the connection between the two cells. Since both cells now contain an F factor, both cells are F+. The new F+ cell (which was the F-cell, can now initiate conjugation with another F-cell. Recombination rarely occurs with this kind of conjugation. This is because the F factor is not homologous to the DNA in the bacterial chromosome. As we will see, however, there are variations of this basic conjugation process that allow recombination to occur.

Conjugation Involving Hfr Bacteria

Occasionally, the F factor integrates into a random position in the bacterial chromosome. When this happens, the bacterial cell is called Hfr instead of F+. Hfr bacteria are still able to initiate conjugation with F-cells, but the outcome is completely different from conjugation involving F+ bacteria: As mentioned above, Hfr cells are formed when the F factor integrates into the bacterial chromosome. This integration occurs at a random location. The Hfr cell is still able to initiate conjugation with an F-cell. When DNA transfer begins, the Hfr cell tries to transfer the entire bacterial chromosome to the F-cell. The first DNA to be transferred is chromosomal DNA, and the last DNA to be transferred will be the F factorDNA.

Transfer of the bacterial chromosome is almost never complete. Pili are fairly fragile structures, and shear forces tend to break the pilus, disrupting DNA transfer before the entire chromosome can be transferred. As a result, the F factor itself is almost never transferred to the recipient cell. This cell will remain F-. This cell will receive new DNA from the Hfr cell however, and this new DNA can undergo recombination at a high frequency with the host chromosome, because the DNA sequences will be homologous. In fact, Hfr is short for 'high frequency recombination'. This recombination can result in gene conversion events, if the transferred DNA and the corresponding region of host DNA contain different alleles of the same gene.

Mapping Genes on Bacterial Chromosomes

Bacteria, since they are usually haploid, cannot have their chromosomes mapped by the same techniques as eukaryotes (For a reminder of how this works, see the module on linkage and mapping). They can, however, be mapped by using Hfr bacterial conjugation. For example, imagine that an F-cell has mutant alleles of two genes, a and b(the F-would therefore be a-, b-). If this cell undergoes conjugation with an Hfr cell that is a+, b+ (in other words, wild type), the F-cell should undergo gene conversion to a+, b+ when both of those genes have been transferred by conjugation. By determining how long it takes the b gene to transfer after the agene has transferred, it is possible to get a relative idea of how far apart the two genes are on a chromosome. The experiment would be done this way: a+, b+ Hfr cells would be mixed with a-,bF-cells. The time of mixing would be designated 'time zero'. At regular intervals, a small amount of the mixture would be removed and conjugation would be disrupted using a blender (the shear force of the blender would cause any pili to break). These bacteria would then be tested for gene conversion (for example, if the mutations rendered the F- bacteria auxotrophic, the bacteria could be tested by growing them on minimal medium, or minimal medium supplemented with the necessary nutrient required because of one or the other mutation). If the agene was converted to wild type at 8 minutes after time zero, and the b gene was converted to wild type at 19 minutes after time zero, then the distance between the two genes would be '11 minutes' (because that was the difference in time required to transfer the b gene compared to the agene). Bacterial map distances are always expressed in minutes, because of this technique.

F' Conjugation

Just as F factors can occasionally integrate into the bacterial chromosome (producing an Hfr cell from an F+ cell), integrated F factors can occasionally excise themselves from the bacterial chromosome. If this excision occurs properly, the Hfr cell becomes an F+ again. The excision is sometimes sloppy, however, and the F factor takes a small segment of the bacterial chromosome with it. Some of the chromosomal DNA has therefore become associated with the episome. When this happens, the cell is called an F'. Conjugation involving F' cells allows for the possibility of recombination, as shown below:

The F' cell has a full complement of chromosomal genes; however, some of those genes are now on the episome. F' cells are able to initiate conjugation with F-cells because of the presence of the F factor. When the F factor begins to transfer its DNA to the recipient cell, it will transfer the small segment of chromosomal DNA as well. Just as in the F+/F-mating, both cells wind up with a copy of the episome. The cell that was F-now has the F factor (along with the piece of chromosomal DNA) and is therefore now F'.

This cell, however, also has a complete chromosome, so it will be diploid for the segment of chromosomal DNA on the episome. Such a partially diploid bacterial cell is called a merozygote. The chromosomal DNA on the episome can undergo recombination at high frequency with its homologous sequence on the chromosome.

Transduction

Transduction involves the exchange ofDNA between bacteria using bacterial viruses (bacteriophage) as an intermediate. There are two types of transduction, generalized transduction and specialized transduction, which differ in their mechanism and in the DNA that gets transferred. Before we can address these processes, however, we need to understand the life cycle of a bacteriophage. When a phage infects a bacterial cell, it injects its DNA into the cell. The viral DNA is replicated numerous times, and viral genes are expressed, producing the proteins that make up the viral capsid (or protein coat) and nucleases that digest the host genome into fragments. The newly replicated viral DNA molecules are packaged into viral capsids, and the bacterial cell is lysed (burst, and therefore killed), releasing hundreds of viral progeny, which then go on to infect other cells.

Sometimes, during bacteriophage replication, a mistake is made, and a fragment of the host DNA gets packaged into a viral capsid. The resulting phage would be able to infect another cell, but it would not have any viral genes, so it would not be able to replicate. The cell infected by this phage would survive, and would have an extra piece of bacterial DNA present, which could undergo recombination with the host chromosome, and perhaps cause a gene conversion event. Because it is a random fragment that gets packaged into the viral capsid, any segment of the bacterial DNA can be transferred this way (hence the name 'generalized').

Specialized Transduction

Specialized transduction occurs only with certain types of bacteriophage, such as phage lambda. Lambda has the ability to establish what is called a lysogenic infection in a bacterial cell. In a lysogenic infection, the viral DNA becomes incorporated into the host chromosome, much as the F factor did in Hfr cells. In a lysogenic infection by lambda, the DNA integrates into a very specific spot in the host chromosome.

The integrated viral DNA can remain integrated for long periods of time, without disturbing the cell. Under the appropriate conditions (the regulation of this is very complex, so don't worry about it), the viral DNA will excise itself from the chromosome, and enter the lytic phase, in which the virus replicates just as described above. The cell gets lysed, and new bacteriophage particles are released to infect other cells. As with excision of the F factor (when Hfr cells become F'), sometimes the excision of lambda is sloppy, and some bacteria DNA is excised along with it. When the resulting virus infects another cell, it will pass that bacterial DNA into the cell, along with its own DNA. If the infected cell survives (it can happen; there are bacterial defenses against viral infection), it will contain a new piece of bacterial DNA, which can undergo recombination and possibly cause gene conversion. Because the viral DNA integrates into a specific location, when it excises, the bacterial DNA removed with it will be the same in all cases. Therefore, the DNA transferred to the second cell will be the same segment of the bacterial chromosome. This is why this process is called 'specialized' transduction.

Bacterial Recombination: Summary

Bacteria can pick up loose DNA in their environment through the process of transformation. The newly acquired DNA is rendered single stranded, and can recombine with the host chromosome.

- Bacteria can exchange DNA through the process of conjugation. The F factor confers the ability to initiate conjugation. If the F factor alone is transferred, no recombination will occur. Under certain circumstances, chromosomal DNA can be transferred to the recipient cell. In these cases, recombination will occur.
- Bacteria can receive bacterial DNA from viruses through the process of transduction. Bacterial viruses can accidentally pick up pieces of bacterial DNA. When they subsequently infect a cell, they transfer the piece of bacterial DNA, which can undergo recombination with the host bacterial chromosome.
- The result of recombination in the above cases may be gene conversion, in which a mutant allele becomes wild-type or vice versa.
- Conjugation involving Hfr bacteria can be used to map genes along the bacterial chromosome. This done by determining in what order genes are transferred during conjugation, what the time difference is between the transfer of genes. Bacteria do not reproduce sexually but can acquire new DNA through transformation, transduction or conjugation. These natural processes have been modified so that DNA can be deliberately incorporated into host microbes-even genes that would normally never be transferred this way.

Genetic Mapping in Bacteria:

Bacteria are haploid organisms and do not undergo meiosis. So for creating their genetic maps geneticists made use of other methods to induce crossovers between homologous segments of bacterial DNA.

They used the three methods of recombination that occurs in bacteria:

(a) Conjugation-Two bacteria come into physical contact and one bacterium (the donor) transfers DNA to the second bacterium (the recipient). The transferred DNA can be a copy of some or possibly the donor cell's entire chromosome, or it could be a segment of chromosome DNA up to 1 mb in length integrated in a plasmid. The latter is called episome transfer,

(b) Transduction- It involves transfer of a small segment of DNA up to 50 kb or so, from donor to recipient via a bacteriophage,

(c) Transformation- The recipient cell takes up from its environment a fragment of DNA, rarely longer than 50 kb, released from a donor cell.

In bacteria, the phenotype studied are the biochemical characteristics like ability to synthesize tryptophan in the dominant or wild type strain and inability to synthesize tryptophan in other strain, which is the recessive allele. The gene transfer is usually set up between a donor strains that possesses dominant gene to the recipient strain that possesses recessive gene. The transfer into the recipient is monitored by looking for attainment of the biochemical function specified by the gene being studied. This can be understood by (Fig. 21.1). Here, the functional gene for tryptophan synthesis from a wild strain is being transferred to recipient that lacks the functional copy of that gene (trp^{-}).

This recipient is called as auxotroph (bacteria which can survive only if provided with tryptophan). The wild strain (trp^{+}) does not require tryptophan for its survival. After the transfer, two crossovers are needed to integrate the transferred gene into the recipient cell's chromosome, converting the recipient from trp" to trp^{+}.

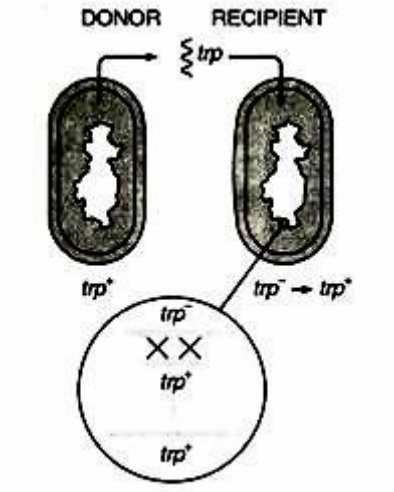


Fig. 21.1. Transfer of DNA between donor and recipient bacteria.

The precise detail of the map depends on the method of gene transfer being used. During conjugation, DNA is transferred from donor to recipient in the same way that a string is pulled through a tube. The relative positions of markers on the DNA molecule can therefore be mapped by determining the times at which the markers appear in the recipient cell. For example in Fig.

21.2, markers A, B and C are transferred after 8, 20 and 30 minutes of beginning of conjugation. The entire E. coli DNA takes approx. 100 minutes to transfer.

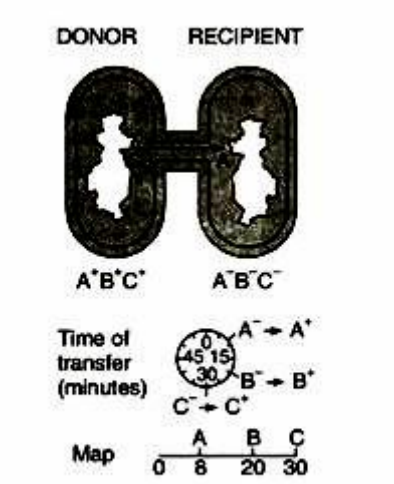
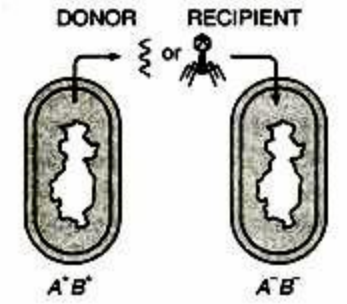


Fig. 21.2. Sequential transfer of markers during conjugation.

In case of transformation and transformation mapping enable genes that are relatively close together to be mapped, because the transferred DNA segment is short (<50kb), so the probability of two genes being transferred together depends on how close together they are on the bacterial DNA (Fig. 21.3).



Frequency with which $A^-B^- \rightarrow A^+B^+$ depends on how close together *A* and *B* are on the chromosome

Fig. 21.3. Co-transfer of closely linked markers during transduction or transformation.

Elie Wollman and Francois Jacob (1950s) conducted first genetic mapping experiments in bacteria. They studied linear transfer of genes in conjugation experiments between Hfr (Hfr-High frequency of recombination) and F^- (F^- fertility factor) strains of E. coli. During the experiment they interrupted conjugation between bacteria at specific times termed as

"Interrupted mating". They noticed that the time it takes for genes to enter a recipient cell is directly related to their order along the chromosome.

This experiment derived them to give the hypothesis that (a) The chromosome of the Hfr donor is transferred in a linear manner to the F- recipient cell (b) The order of genes along the chromosome can be deduced by determining the time required for various genes to enter the recipient.

Conjugation studies have been used to map over 1,000 genes along the circular E. coli chromosome. The genetic maps are scaled in minutes e.g., E. coli chromosome is 100 minutes long, conjugative transfer of the complete chromosome takes approximately 100 minutes (Fig. 21.4).

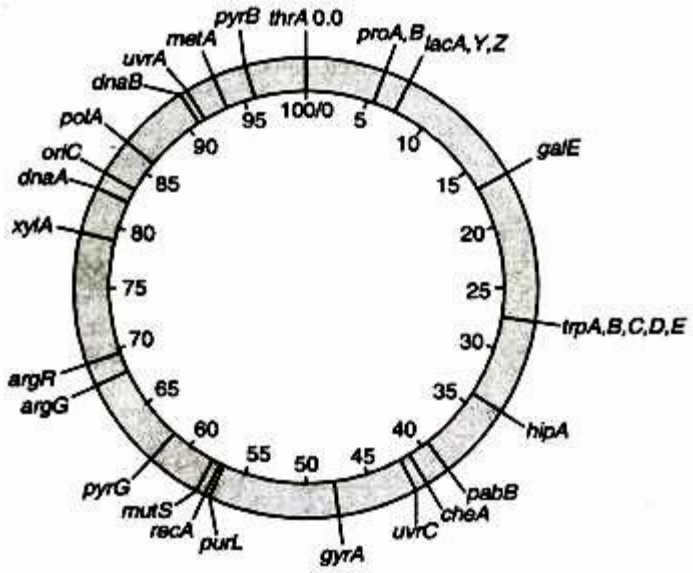


Fig. 21.4. Genetic map of *E. coli*; Positions of genes indicate the relative times at which they are transferred.

POSSIBLE QUESTIONS

2 marks

1. Define multiple allele.
2. Define allele.
3. What are the methods implicated in gene mapping in *E.coli* ?
4. Define conjugation.
5. Define lethal allele
6. Define Null allele.
7. Define transformation.

8 Marks

1. Write short notes on multiple allele, null allele with suitable examples.
2. Discuss about genetic exchange mechanism in detail.
3. Comment on interaction between gene and environment.
4. Explain in detail on gene mapping in bacteria.
5. Comment on epistatis and non-epistatis in detail.
6. Explain the mechanism of genetic exchange.
7. Compare incomplete dominance with co-dominance.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
II-B.Sc., BIOCHEMISTRY
CONCEPTS IN GENETICS (17BCU304B)
MULTIPLE CHOICE QUESTIONS
UNIT II

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	Incomplete dominance is also termed as	partial dominance	half dominance	co-dominance	none of above	**partial dominance**
2	In incomplete dominance, one could get 1 : 2 : 1 ratio in	test cross	F2 generation	F1 generation	R cross	**F2 generation**
3	ABO blood grouping is based on	codominance	incomplete dominance	epistasis	multiple allelism	**multiple allelism**
4	A mutation in which most of the protein coding sequence of the gene is removed is most likely to be which type of allele?	Lethal	Dominant	sex-limited	loss of function	**loss of function**
5	Which of the following statements about a lethal allele is NOT correct?	Lethal alleles are always recessive	Lethal alleles may have a late age of onset	Lethal alleles may be caused by mutations in essential genes	None of these	**Lethal alleles are always recessive**
6	Co-dominance differs from incomplete dominance as in co-dominance	the hybrid is intermediate	both the genes are expressed equally	dominant gene is expressed in F1 generation	genotypic ratio is 1 : 1	**both the genes are expressed equally**
7	Multiple alleles of a gene always occupy	the same locus on a chromosome	the same position on different chromosome	different loci on a chromosome	different loci on different chromosomes	**the same locus on a chromosome**
8	AaBbCc is the genotype of	fair	mulatto	pure black (negro)	albino	**mulatto**
9	The plasmid mediated properties are	fermentation of lactose	production of entrotoxin	resistance to antibiotics	all of above	**all of above**
10	In extracelluar medium, DNA degrading enzymes would likely to prevent transfer of DNA by	conjugal transfer by a self transmissible plasmid	generalized phage transduction	natural transformation	none of the above	**natural transformation**
11	What is the term used for a segment of DNA with one or more genes in the centre and the two ends carrying inverted repeat sequences in nucleotides	plasmid	transposon	insertion sequence	none of these	**transposon**
12	The plasmids can be eliminated from a cell by the process known as	curing	breaking	fixing	expulsion	**curing**
13	Recombination of virus genomes occurs	by transduction	by transcription	simultaneous infection of a host cell by two viruses with homologous chromosomes	by transformation	**simultaneous infection of a host cell by two viruses with homologous chromosomes**
14	How is genetic transfer different from sexual reproduction?	In genetic transfer, genetic material is transferred from one individual to another.	In sexual reproduction, genetic material from two individuals is mixed equally.	In both processes, the individual donating genetic material retains their original complement of genetic material.	None of these.	**In sexual reproduction, genetic material from two individuals is mixed equally.**
15	Which method of gene transfer involves direct contact between the bacteria?	Conjugation.	Transduction.	Transformation.	All of these.	**Conjugation.**
16	How might the results of Bernard Davis's U tube experiment have been different if each strain were	There would be no difference.	Transformation by small pieces of DNA could restore	Conjugation would not be required to	b and c	**b and c**

	only mutant for a single gene?		gene function.	change the phenotype of the bacterial strains.		
17	The end result of F factor mediated conjugation	is that both strains are F+.	involves transfer of the entire bacterial chromosome.	converts the recipient strain to F+ and the donor to F-.	a and b	**is that both strains are F+.**
18	Which of the following is not a protein or group of proteins required for conjugation to occur?	The relaxosome.	Pilin protein.	The origin of transfer.	Coupling protein.	**The origin of transfer.**
19	Which of the following is a description of generalized transduction?	Inaccurate excision of prophage can occasionally remove a small portion of the bacterial chromosome.	Random pieces of bacterial DNA can become incorporated into a phage coat.	Pieces of bacterial DNA may be released into the culture medium when a cell is lysed.	All of these.	**All of these.**
20	What is the best explanation for why Bernard Davis didn't observe generalized transduction in his U tube experiments?	He used genes that could not be packaged into a bacteriophage head.	He used bacterial strains that could not fit through the filter.	He used bacterial strains that did not host a prophage.	He used virulent bacteriophage.	**He used bacterial strains that did not host a prophage.**
21	A researcher would like to map the location of galE and trpA genes in a new species of bacterium that appears to be closely related to E. coli. He decides to use cotransduction, and generates appropriate donor and recipient strains. He is disappointed when cotransduction is not seen in his experiement. What is the most reasonable explanation for this situation?	His new bacterial species does not have galE or trpA genes.	His new bacterial species cannot survive galE or trpA mutation.	These two genes are too far apart to be mapped by cotransduction.	These two genes are too close together to be mapped by cotransduction.	**These two genes are too far apart to be mapped by cotransduction .**
22	Which of the following is NOT a critical function of competence factors?	Transportation of DNA fragments into bacterial cells.	Binding of DNA to a cell surface receptor.	Degradation of one strand of the transforming DNa.	Degradation of the bacterial chromosome in the presence of single standed DNA.	**Degradation of the bacterial chromosome in the presence of single standed DNA.**
23	How might the function of extracellular endonucleases help protect the bacterium from infection by viruses?	They would prevent the virus from binding to the outside of the bacterial cell.	They would destroy the viral coat, so it cannot inject its genetic material.	They would cut the viral DNA into small pieces.	All of the above.	**They would cut the viral DNA into small pieces.**
24	What do transduction, transformation, and conjugation have in common?	All require presence of an outside factor to facilitate gene transfer.	In all three processes, DNA is transferred as a single stranded molecule.	They all transfer large pieces of DNA into recipient cells.	None of the above.	**None of the above.**
25	Which type(s) of genetic transfer lead to incorporation of new DNA into the bacterial chromosome? I. Conjugation mediated by F factor II. Hfr mediated conjugation III. Transduction IV. Transformation	I and II only	III and IV only	II, III, and IV	None of these.	**II, III, and IV**
26	How would you expect a mutation in a gene encoding a tail fiber protein would affect the plaque phenotype of T4?	It could lead to very large plaques.	It would be unlikely to have an affect on plaque formation.	It could lead to smaller than normal plaques.	A and B.	**It would be unlikely to have an affect on plaque formation.**
27	In Benzer's intragenic mapping experiments, what event was required to allow production of infectious phage from rII mutants?	rII mutant strains with mutations in different genes.	Homologous recombination between phage.	Crossover within the mutant gene, between the mutations.	Crossover at one end of the mutant gene, outside the mutation.	**Crossover within the mutant gene, between the mutations.**

28	In deletion mapping, mutations can be localized to a region of the gene because:	the mutation fails to complement other mutations in the gene.	the mutation fails to complement a deletion of a known portion of the gene.	the gene sequence can be determined.	a and b	**the mutation fails to complement a deletion of a known portion of the gene.**
29	A key finding of intragenic mapping was:	that genes consist of indivisible particles, like beads on a string.	that genes do not change over time.	that genes can be modified by crossing over and mutation.	All of these	**that genes can be modified by crossing over and mutation.**
30	In the extracellular medium, DNA-degrading enzymes would likely be to prevent transfer of DNA by	conjugal transfer by a self-transmissible plasmid generalized phage	transduction	natural transformation	none of the above	**natural transformation**
31	The type of recombination that commonly occurs between a pair of homologous DNA sequences is,	mutagenic recombination	site-specific recombination	replicative recombination	general recombination	**general recombination**
32	Which of the following statement describes plasmids?	Another name for a protoplast	A complex membrane structure that covers the chromosome of bacteria	Small, circular DNA molecules that can exist independently of chromosomes commonly found in bacteria	None of the above	**Small, circular DNA molecules that can exist independently of chromosomes commonly found in bacteria**
33	In lysogeny,	a bacteriophage transfers bacterial DNA	bacteria take up double stranded DNA from the environment	DNA-degrading enzymes in the extracellular medium would stop the process	a bacteriophage genome is integrated into the bacterial genome	**a bacteriophage genome is integrated into the bacterial genome**
34	Who discovered transposons (jumping genes)?	Abelson	Harvey	McClintock	Griffith	**McClintock**
35	Which type of plasmid can exist with or without being integrated into the host's chromosome?	Medisome	Lisosome	Lysogen	Episome	**Episome**
36	Which of the following statement(s) is/are true in regards to F+ x F- mating events?	DNA is transferred from F- to F+ cells	DNA is transferred from F+ to F- cells	No DNA is transferred because F- cells are unable to perform conjugation	No DNA is transferred because F+ cells are unable to perform conjugation	**DNA is transferred from F+ to F- cells**
37	What information can be generated by interrupted mating experiments?	Levels of DNA homology	Bacterial genome maps	DNA nucleotide sequences	Proteomics of the bacteria	**Bacterial genome maps**

38	Which of the following transport bacterial DNA to other bacteria via bacteriophages?	Conjugation	Transduction	Transformation	Translation	**Transduction**
39	The term used for acquisition of naked DNA from its environment and its incorporation in their genome by a bacterium is	transformation	lysogenic conversion	conjugation	transduction	**transformation**
40	The expression of gene X (which has promoter Px) is to be monitored. A gene fusion construction for carrying this work will	have Px but not the rest of the X coding region	allow to monitor the expression of all genes with a promoter similar in sequence to Px	give the same information as from a microarray	have the promoter of lacZ or some other reporter gene	**have Px but not the rest of the X coding region**
41	The correct term for the transfer of genetic material between bacteria in direct physical contact is	Conjugation	Transduction	Transformation	Translation	**Conjugation**
42	Interaction between two alleles which are present on the same gene locus of two homologous chromosomes is called	intragenic interaction	interallelic interaction	intergenic interaction	both (a) and (b)	**both (a) and (b)**
43	Skin colour is controlled by	2 pairs of genes	single gene	3 pairs of genes	2 pairs of genes with an intragene	**3 pairs of genes**
44	Some individuals with blood group A may inherit the genes for black hair, while other individuals with blood group A may inherit the genes for brown hair. This can be explained by the principle of	independent assortment	incomplete dominance.	dominance	multiple alleles.	**independent assortment**
45	If gene expresses itself then it's penetrance is	50%	100%	150%	200%	**100%**
46	AaBbCc is the genotype of	fair	mulatto	pure black (negro)	albino	**mulatto**
47	Multiple alleles are present	on different chromosomes.	at different loci on the same chromosome.	at the same locus of the chromosome	on non-sister chromatids	**at the same locus of the chromosome**
48	When red homozygous flower is paired with white homozygous flower then resultant pink snapdragon is result of	partial dominance	complete dominance	co-dominance	all of above	**partial dominance**
49	RR (red) flowered plant of Mirabilis is crossed with rr (white) flowered plant of Mirabilis. All the Rr offsprings are pink. This is an indication that the R gene is	codominant	recessive	incompletely dominant	linked	**incompletely dominant**
50	Blood grouping in humans is controlled by	4 alleles in which A is dominant	3 alleles in which AB is codominant	3 alleles in which none is dominant	3 alleles in which A is dominant	**3 alleles in which AB is codominant**

UNIT 3
SYLLABUS

Genetics of eukaryotes and Human pedigree analysis

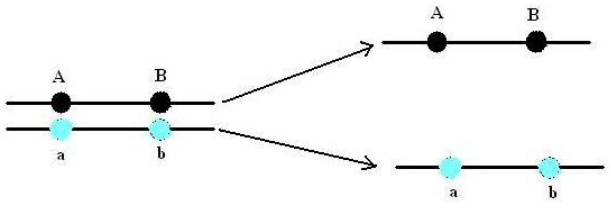
Linkage and crossing over, genetic mapping in eukaryotes, centromere mapping with ordered tetrads, cytogenetic mapping with deletions and duplications in *Drosophila*, detection of linked loci by pedigree analysis in humans and somatic cell hybridization for positioning genes on chromosomes. Pedigree conventions, characteristics of dominant and recessive inheritance. Applications of pedigree analysis.

Linkage and Crossing over

Linked genes

By studying the inheritance of characters in the fruitfly Drosophila, TH Morgan and colleagues (1910) determined that genes are not completely independent as Mendel had thought, but that they tend to be inherited in groups. Since independent assortment does not occur, a dihybrid cross following two linked genes will not produce and F2 phenotypic ratio of 9:3:3:1. They observed that genes in the same chromosome are often transmitted together as a group, but that this was not always so and that 'crossing-over' between chromosomes could occur to disrupt these linkage groups. Genes that are present on the same chromosome, and that tend to be inherited (transmitted to the gametes) together, are termed **linked genes** because the DNA sequence containing the genes is passed along as a unit during **meiosis**. The closer that genes reside on a particular chromosome, the higher the probability that they will be inherited as a unit, since crossing over between two linked genes is not as common. The genes present on same chromosome, thus, would not assort (separated) independently. Such type of genes are called **linked genes** and this phenomenon is called **linkage.**

For example, the "A" **and** "B" alleles which are present in same chromosome will both be passed on together if the chromosome is inherited. "A" **and** "B" are linked due to their occurrence in the same chromosome. Similarly, "**a**" **and** "**b**" are linked in the other chromosome.



Linked genes tend to be inherited together because they are located on the same chromosome

Morgan defined linkage as follows: "that the pairs of genes of homozygous parents tend to enter in the same gametes and to remain together, whereas same genes from heterozygous parent tend to enter in the different gametes and remain apart from each other. He further stated that the tendency of linked genes remaining together in original combination is due to

their location in the same chromosome ".According to him the degree or strength of linkage depends upon the distance between the linked genes in the chromosome.

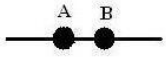
Chromosomes Theory of Linkage:- Morgan along with castle formulated the chromosome theory of linkage which is as follows :-

1. The genes which show the phenomenon of linkage are situated in the same chromosomes and separated during the process of inheritance.
2. The distance between the linked genes determines the strength of linkage. The closely located genes show strong linkage than the widely located genes which show the weak linkage.
3. The genes are arranged in linear fashion in the chromosomes.

Types of linkage :- Linkage may be complete, incomplete, or absent (not detectable)linkage, depending upon the distance between linked genes in a chromosome.

Complete linkage

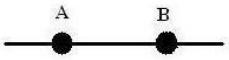
During synapsis, exchange of segments takes place. In such condition the possibility of separation of two genes situated close together is greatly reduced. When genes are closely associated and tend to transmit together, it is called complete linkage.



"A" and "B" are too close to each other.

Incomplete linkage

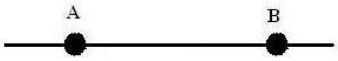
When linked genes are situated at long distance in chromosomes and have chances of separation by crossing over are called incompletely linked genes and phenomenon of their inheritance is called **incomplete linkage**.



"A" and "B" are separated to allow crossing-over between them

Absent (not detectable)linkage

The probability of crossover increases with the physical distance between genes on a chromosome, and genes that are located quite far from each other within a linkage group may not exhibit any detectable linkage



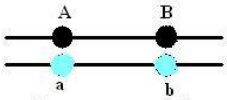
"A" and "B" are too far apart to allow crossing-over between them in all

image407

Linkage groups :- The group of linked genes that are located on the same chromosome, called linkage groups. Because, all the genes of a chromosome have their identical genes (alleomorphs) on the homologous chromosome, therefore linkage groups of a homologous pair of chromosome is considered as one. In any species, the number of linkage groups is equal to the number of pairs of chromosomes. **e.g. Corn (Zea mays)** has 10 pairs of chromosomes and 10 linkage groups.

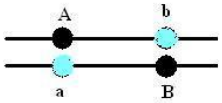
Arrangement of genes

Cis- arrangement of genes : If the dominant alleles(A,B) of two linked genes are present on the same chromosome and their recessive alleles (a,b) are present on the homologous chromosomes the arrangement of genes is called **cis- arrangement**.



Cis- arrangement of genes

Trans- arrangement if one dominant gene and other recessive gene present on one chromosome (A,b) and their allele type (a, B) on the chromosome this type of arrangement is known as **trans-arrangement**.



Trans- arrangement of genes

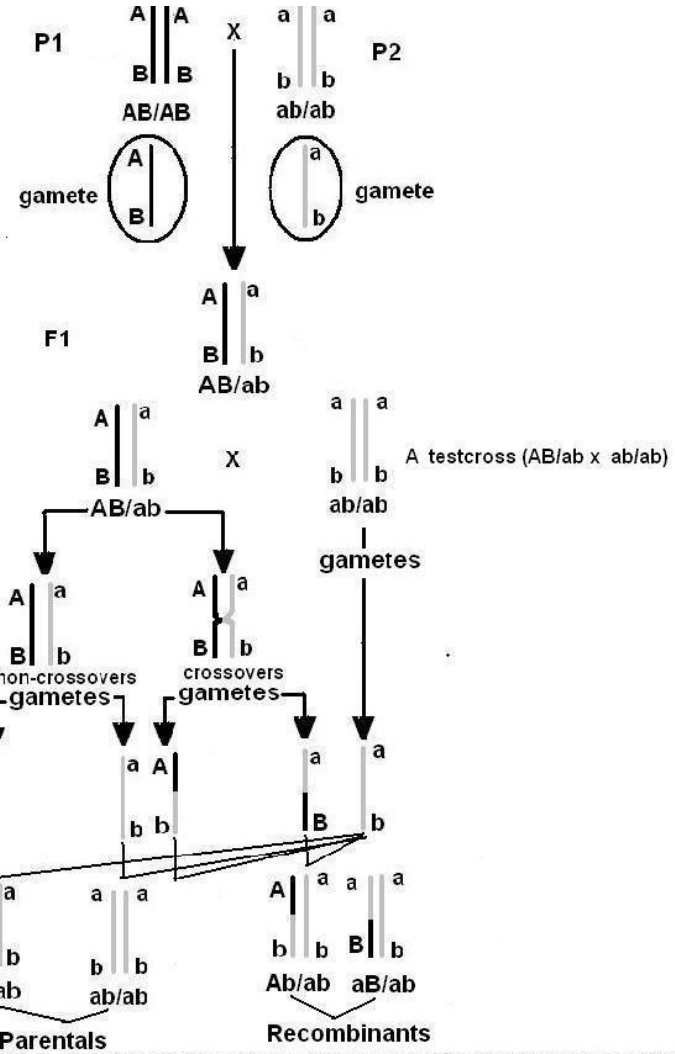


Fig 11 showing crossover and non crossover gametes(AB/ab) and progenies after test cross

Subsections---"B"

Crossing-over:-Crossing-over takes place during prophase I of meiosis. Crossing-over is another name for recombination or physical exchange of equal pieces of adjacent non-sister chromatids. During the process of crossing-over one of the paired chromosome arms exchanged physically at one or more locations .The two homologous chromosomes are connected at a certain point called **chiasma (pl chiasmata).**When crossing-over occurs chromatids break at chiasma and reattached to a different homologous chromosome . The chromatids resulting from the interchange of segments are known as the **cross over recombinants** and the chromatids that remain intact are called **non-crossover parental chromatids.** When these chromosomes segregate in meiosis, they form gametes that have completely new combinations of alleles. Generally, the longer the chromosome, the greater the number of chiasmata. The ability of genes to recombine is called recombination frequency.

For example ,If the alleles on tetrad are :-

- **1. A B C D E F G**
- **2. A B C D E F G**
- **3. a b c d e f g**
- **4. a b c d e f g**

After crossing- over a new associations of genes and alleles are formed if crossing over takes place between 2nd and 3rd chromosomes

- **1. A B C D E F G- non-crossover parental chromatids**
- **2. A B c d e f g - cross over recombinants**
- **3. a b C D E F G - cross over recombinants**
- **4. a b c d e f g - non-crossover parental chromatids**

Types of crossing over crossing over are of many types depending on number of chiasma:-

1. **Single crossing over**:- when the chiasma formation takes place at a single point of The chromosome pair this type of crossing over is known as single crossing over. In this types two crossed over chromatids and two non crossed over chromatids are formed.

Fig 12 Single crossing over

image412

2. **Double crossing over**:- When the chiasmata occur at two places in the same chromosomes known as double crossing over In the double crossing over formation of each chiasma is independent of the other and in it four types of recombination is possible.

Two types of chiasma may be formed in double cross over:-

A. **Reciprocal chiasma** in this type both the chiasma are formed on two same chromatids. So, the second chiasma restores the order which was changed by the first Chiasma, and as a result two non- cross over chromatids are formed.

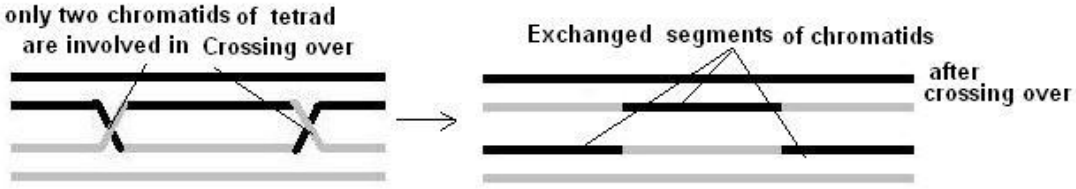


Fig 12 Double crossing over (Reciprocal chiasma)

image413

In this type out of four chromatids only two are involved in the double crossing over.

B. **Complimentary chiasma** When both the chromatids taking part in the second chiasma are different from those chromatids involved in the first. In this type four single cross over are produced but no non cross over. Complimentary chiasma occurs when three or four chromatids of tetrad undergo crossing over.

3. **Multiple crossing over** When crossing over take place at more than two point in the same chromosome pair it is known as multiple crossing over. It occurs rarely.

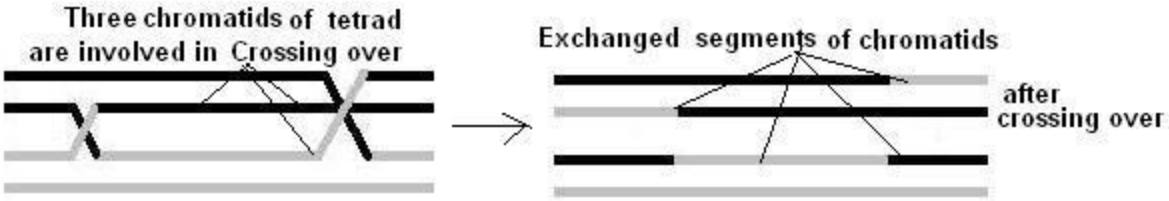


Fig 13 Double crossing over (Complimentary chiasma)

image414

There are two theories on the physical nature of the process:-

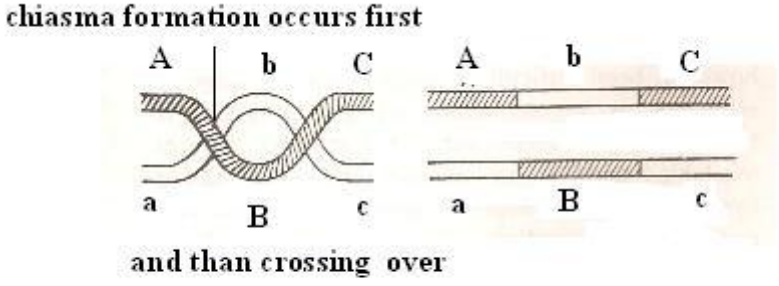


Fig 9 Showing chiasma formation and crossing over - Classical theory

image408

1. Classical theory or two plane theory (L. W. Sharp):- proposes that cross-over and formation of the chiasma occur first, followed by breakage and reunion with the reciprocal homologues. According to this theory, chiasma formation need not be accompanied by chromosome breakage. But this theory was not accepted.

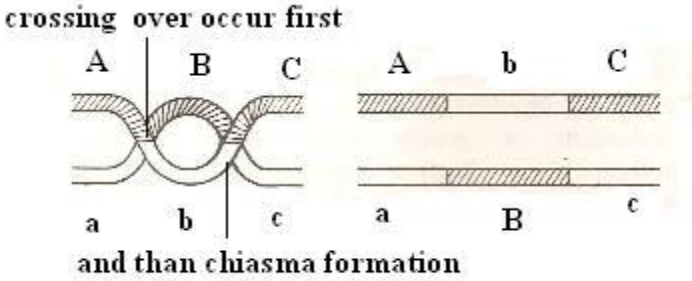


Fig 10 Showing crossing over and chiasma formation-- Chiasmatype theory

2. Chiasma type theory or one plane theory This theory was proposed by F.A. Janssens (1909) breakage occurs first, and the broken strands then reunite. Chiasmata are thus evidence, but not the causes, of a cross-over.

Recombination During Meiosis:- John Belling (1928) suggested that no break was necessary and proposed the copy choice model. He believed that crossing over might occur during duplication of homologous chromosomes and might brought about due to novel attachments formed between newly synthesized genes. While studying meiosis in some plant species. He visualized genes as beads (described as chromomeric), connected by non-genic interchromomeric regions. The newly synthesized daughter chromatids is derived due to copying of one chromosomes upto certain region and then switching on to the other homologous chromosome for copying the remaining portion or region of the chromosomes. The new chromatid would have a new arrangement, but no breaks and rejoining need be involved. This was such an attractive idea that the hypothesis in some form held center stage for nearly thirty years.

New chromosome arrangements were associated with recombinant genes, using chromosomes marked by mutant genes and morphological differences at each end demonstrated in *Drosophila melanogaster* (Stern 1931) and *Zea mays* (H. S. Creighton and B. McClintock 1931). That crossing over was correlated with segmental interchange between homologous chromosomes.

DNA models during 1960s,DNA models had become very popular, the widely accepted model for DNA crossover was first proposed by Robin Holliday in 1964.

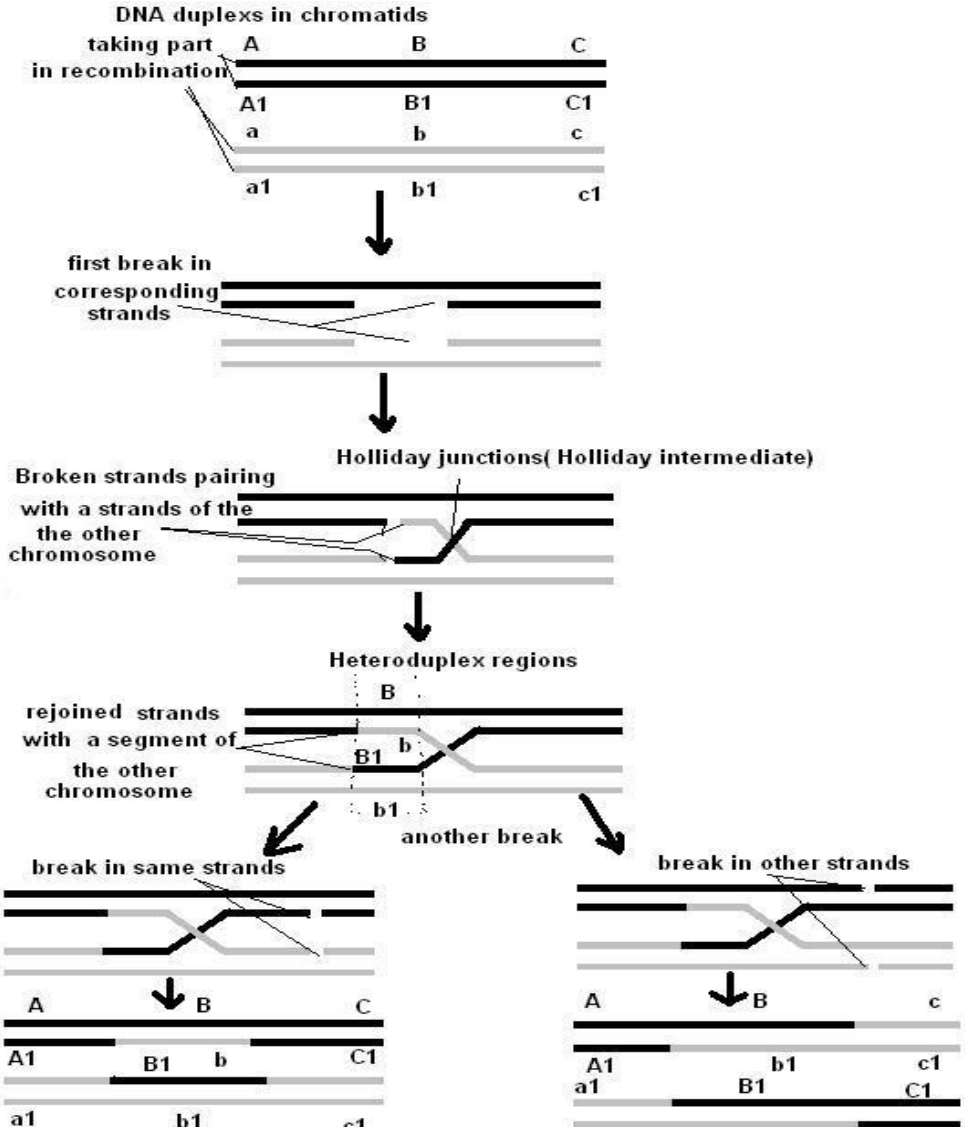


image416

Homologous recombination during meiosis has following important features.

Two homologous DNA molecules line up (e.g., two non-sister chromatids line up during meiosis).

Cuts in one strand of both DNAs, a double-strand break in a DNA molecule is enlarged by an exonuclease, such that the single-strand extension with a free 3‘-hudroxyl group is left at the broken end.

The exposed 3‘ ends invade the intact duplex DNA, and this is followed by branch migration and/or replication to create a pair of crossover structures,(The cut strands cross and join homologous strands)called Holliday junctions(Holliday structure).

Heteroduplex region is formed by branch migration, cleavage of the two crossover creates two complete recombinant products

In this double-strand break repair model for recombination, the 3‘ ends are used to initiate the genetic exchange. Once paired with the complementary strand on the intact homolog, a region of hybrid DNA is created containing complementary strands from two different parental DNA . Each of the 3‘ ends can then act as a primer for DNA replication.

The structure that are formed, called Holliday Intermediates are a feature of homologous genetic recombination pathways in all organisms. Homologous recombination can vary in many ways from one species to another, but most of the process are same.

DNA strands may be cut along either the vertical line or horizontal line and break the Holliday intermediate so that the two recombinant products carry genes in the same linear order as the original, unrecombined chromosomes.

If cleaved occurs on vertical line , the DNA flanking the region containing the hybrid DNA is not recombined; if cleaved occurs on the horizontal line. , the flanking DNA is recombined .

Subsections---"C"

Cytological basis of crossing over Creighton and McClintock first obtained a strain of corn with an abnormal chromosome belonging to homologous pair number nine.

This chromosome carried a distinct knob at one end and a detectable translocation at the other end. The knob has no effect on the appearance of the plants, but the translocation, when heterozygous with normal chromosomes 8 and 9, results in 50% of pollen grains being sterile and empty and 50% of eggs, embryo sacs, and ovules aborting—consequently, the ear becomes only half-filled, and the kernels are irregularly distributed.

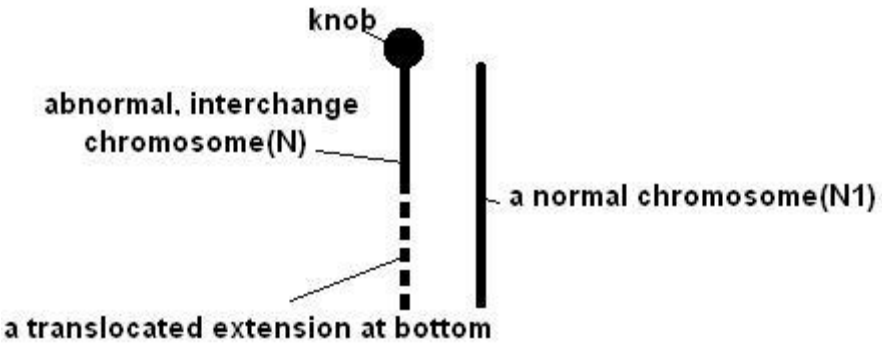


Fig15. An abnormal, interchange chromosome(N), which possesses visible abnormalities at both ends,a knob at the top and a translocated extension at bottom and a normal chromosome(N1)

image417

The interchange between chromosome 9 and chromosome 8 was first recognized by its effects on fertility, with 50% pollen and kernel abortion This semi sterility results from the formation of deficient gametes following synapsis of four chromosomes in a cross-shaped

configuration. This configuration opens out into a ring , from which the centromeres distribute either two alternate chromosomes, with balanced (viable) genomic constitutions, or two adjacent chromosomes, with deficient-duplicate (in viable) constitutions.

By crossing corn carrying only the abnormal chromosomes with corn carrying only normal chromosomes they were able to obtain plants with heteromorphic pairs of chromosomes — that is, with one normal and one abnormal chromosome number nine. Then, they carried out crosses involving plants with heteromorphic chromosome pairs so that they could look for the occurrence of cytological recombination.

The essential components to the demonstration of cytological and genetic crossing over are:-

a. Differential features along the chromosomes that are morphologically (i.e., physically) recognizable and
b. genes in the region of the cytological markers (groups of traits= a distinguishing characteristic)

.

Genetic markers (the seed color alleles and their associated inheritance patterns) and cytological markers (the presence of abnormal sets of sex chromosomes).

Two cytological features were sufficient for the experiment::-

- A dark-staining, heterochromatic ‘‘knob’’ at the end of chromosome 9, and
- A reciprocal interchange (translocation) of a part of chromosome 9 with a part of chromosome 8.

The knob feature is present in some strains and absent in others.

Crossing over and Linkage Maps Recombination due to crossing over can be used to order and determine distances between loci (chromosome positions) by genetic mapping techniques. Loci that are on the same chromosome are all physically linked to one another, but they can be separated by crossing over. Examining the frequency with which two loci are separated allows a calculation of their distance: The closer they are, the more likely they are to remain together. The percentage of recombinants formed by F1 individuals can range from a fraction of 1% up to the 50% always seen with gene loci on separate chromosomes (independent assortment).

Example :- Normally, when a heterozygous colored, full kernel(CcSHsh) corn plant crossed with homozygous(ccshsh)recessive, expected result is:- CcSHsh X ccshsh

Phenotype	Genotype	Observed	Expected
CcSHsh	color, full	4032	1/4
Ccshsh	color, shrunken	149	1/4
ccSHsh	colorless, full	152	1/4
ccshsh	colorless, shrunken	4035	1/4

The two types with the higher numbers represent the two parent-type gametes, since not all cells will crossover. The other two represent the single crossovers.Since it is observed that crossing over has occurred, therefore it is possible to map the distance

between these two traits on the chromosome. The closer the traits are on the chromosome, the less likely crossing over will occur.

To figure map distance, the numbers are observed to determine the parent crosses, and whether or not the genes are in trans or cis formation. Then this equation is possible:-

$$\text{Map distance} = \frac{\#\text{ (Number) of crossover gametes}}{\text{Total gametes}} \times 100$$

This percent can be used as the number of map units apart the two genes are. If the two genes are further than 50 units, crossing over will not be a factor. Therefore, the percentage of recombinant gametes (reflected in the percentage of recombinant offspring) correlates with the distance between two genes on a chromosome. By comparing the recombination rates of multiple different pairs of genes on the same chromosome, the relative position of each gene along the chromosome can be determined. This method of ordering genes on a chromosome is called a linkage map.

The higher the percentage of recombinants for a pair of traits, the greater the distance separating the two loci. By definition, one map unit (m.u.) is equal to one percent recombinant phenotypes. In honor of the work performed by Morgan, one m.u. is also called one centimorgan (cM).

Interference and Coincidence: Crossing over does not occur uniformly along a chromosome. For example, fewer crossovers occur in the area around the centromere than in other areas of the chromosome (making the loci appear closer together than they actually are). Recombination is detected in heterozygotes only. It also occurs in homozygotes with the same frequency but not detected. Because Maximum frequency of recombination between two linked genes is 50% and cannot exceed 50%. Due to Even if every meiocyte had a crossing-over between two linked genes, only 2 of 4 chromatids in a bivalent (tetrad) was involved in crossing-over. (2/4 = 50%) Double or multiple crossovers decrease the number of detectable recombinants (cancellation effect).

Also, the formation of one chiasma typically makes it less likely that a second chiasma will form in the immediate vicinity of the first. This seems to may be due to the inability of the chromatids to bend back upon themselves within a certain minimum distance. This lack of independence is called interference and results in the observation of fewer double crossover types than would be expected according to true map distance. Interference varies in different sections of the chromosome and is measured by the Coefficient of Coincidence (C.C.) which is the ratio of observed to expected double crossover types.

C.C. = (observed DCO) / (expected DCO)

Interference = 1 - C.C.

Genetic mapping in eukaryotes

Genetic mapping is based on the principles of inheritance as first described by Gregor Mendel in 1865 and genetic linkages.

Genetic maps are created to locate the genes or characters on the chromosome for their utilization in genetic studies. Physical maps are created to identify certain markers to detect or diagnose the specific character.

i. Genetic Linkage:

Genetic linkage occurs when particular genetic loci or alleles for genes are inherited jointly. Genetic linkage was first discovered by the British geneticists William Bateson and Reginald Punnett shortly after Mendel's laws were rediscovered. Genetic loci on the same chromosome are physically connected and tend to stay together during meiosis, and are thus genetically linked. For example, in fruit flies the genes affecting eye color and wing length are inherited together because they appear on the same chromosome.

Alleles for genes on different chromosomes are usually not linked, due to independent assortment of chromosomes during meiosis. Because there is some crossing over of DNA when the chromosomes segregate, alleles on the same chromosome can be separated and go to different daughter cells. There is a greater probability of this happening if the alleles are far apart on the chromosome, as it is more likely that a cross-over will occur between them. The relative distance between two genes can be calculated using the offspring of an organism showing two linked genetic traits, and finding the percentage of the offspring where the two traits do not run together.

The higher the percentage of descendants that does not show both traits, the further apart on the chromosome they are. Among individuals of an experimental population or species, some phenotypes or traits occur randomly with respect to one another in a manner known as independent assortment.

Today scientists understand that independent assortment occurs when the genes affecting the phenotypes are found on different chromosomes or separated by a great enough distance on the same chromosome that recombination occurs at least half of the time. But in many cases, even genes on the same chromosome that are inherited together produce offspring with unexpected allele combinations. These results from a process called crossing over.

At the beginning of normal meiosis, a chromosome pair (made up of a chromosome from the mother and a chromosome from the father) intertwine and exchange sections or fragments of chromosome. The pair then breaks apart to form two chromosomes with a new combination of genes that differs from the combination supplied by the parents. Through this process of recombining genes, organisms can produce offspring with new combinations of maternal and paternal traits that may contribute to or enhance survival.

ii. Genetic Map:

A genetic map is a linkage map of a species or experimental population that shows the position of its known genes and/or genetic markers relative to each other in terms of recombination frequency during crossover of homologous chromosomes. The greater the frequency of recombination (segregation) between two genetic markers, the farther apart they are assumed to

be. Conversely, the lower the frequency of recombination between the markers, the smaller the physical distance between them.

Historically, the markers originally used were detectable phenotypes (enzyme production, color, shapes etc.) derived from coding DNA sequences. Now, non-coding DNA sequences such as microsatellites or those generating restriction fragment length polymorphisms (RFLPs) have been used. Genetic maps help researchers to locate other markers, such as other genes by testing for genetic linkage of the already known markers. A genetic map is not a physical map or gene map.

To be useful in genetic analysis, a gene must exist in at least two forms, or alleles; each specifying a different phenotype. Earlier only those genes could be studied whose specifying phenotypes were distinguishable by visual observation. This approach soon became outdated as in many cases a single phenotypic character could be affected by more than one gene. For example, in 1922, 50 genes had been mapped onto the four fruit fly chromosomes, but nine of these genes were for eye color.

The observations by Thomas Hunt Morgan that the amount of crossing over between linked genes differs (partial linkage) led to the idea that crossover frequency might indicate the distance separating genes on the chromosome. Morgan's student Alfred Sturtevant developed the first genetic map, also called a linkage map.

iii. Recombination Frequency:

Sturtevant assumed that crossing over was a random event, there being an equal chance of it occurring at any position along a pair of lined-up chromatids. He proposed that the greater the distance between linked genes, the greater the chance that non-sister chromatids would cross over in the region between the genes. By working out the number of recombinants it is possible to obtain a measure for the distance between the genes. This distance is called a genetic map unit (m.u.), or a centimorgan and is defined as the distance between genes for which one product of meiosis in 100 is recombinant.

A recombinant frequency (RF) of 1% is equivalent to 1 m.u. A linkage map is created by finding the map distances between a numbers of traits that are present on the same chromosome, ideally avoiding having significant gaps between traits to avoid the inaccuracies that will occur due to the possibility of multiple recombination events.

Recombination frequency is the frequency that a chromosomal crossover will take place between two loci (or genes) during meiosis. Recombination frequency is a measure of genetic linkage and is used in the creation of a genetic linkage map. During meiosis, chromosomes assort randomly into gametes, such that the segregation of alleles of one gene is independent of alleles of another gene. This is stated in Mendel's second law and is known as the law of independent assortment.

The law of independent assortment always holds true for genes that are located on different chromosomes, but for genes that are on the same chromosome, it does not always hold true. As an example of independent assortment, consider the crossing of the pure-bred homozygote parental strain with genotype AABB with a different pure-bred strain with genotype aabb. A and a and B and b represent the alleles of genes A and B. Crossing these homozygous parental strains will result in F1 generation offspring with genotype AaBb.

The F1 offspring AaBb produces gametes that are AB, Ab, aB, and ab with equal frequencies (25%) because the alleles of gene A assort independently of the alleles for gene B during meiosis. Note that 2 of the 4 gametes (50 %) Ab and aB-were not present in the parental generation. These gametes represent recombinant gametes. Recombinant gametes are those gametes that differ from both of the haploid gametes that made up the diploid cell. In this example, the recombination frequency is 50% since 2 of the 4 gametes were recombinant gametes.

The recombination frequency will be 50% when two genes are located on different chromosomes or when they are widely separated on the same chromosome. This is a consequence of independent assortment. When two genes are close together on the same chromosome, they do not assort independently and are said to be linked. Linked genes have a recombination frequency that is less than 50%.

As an example of linkage, consider the classic experiment by William Bateson and Reginald Punnett. They were interested in trait inheritance in the sweet pea and were studying two genes-the gene for flower color (P- purple and p- red) and the gene affecting the shape of pollen grains (L- long and I- round). They crossed the pure lines PPLL and ppll and then self-crossed the resulting PpLl lines.

According to Mendelian genetics, the expected phenotypes would occur in a 9:3:3:1 ratio of PL:P1:pL:p1. To their surprise, they observed an increased frequency of PL and pi and a decreased frequency of P1 and pL (Table 21.1).

Table 21.1. Bateson and Punnett experiment.

Phenotype and genotype	Observed	Expected from 9:3:3:1 ratio
Purple, long (PPLL)	284	216
Purple, round (PPll)	21	72
Red, long (ppLL)	21	72
Red, round (ppll)	55	24

Their experiment revealed linkage between the P and L alleles and the p and l alleles. The frequency of P occurring together with L and with p occurring together with I is greater than that of the recombinant PI and pL. The recombination frequency cannot be computed directly from this experiment, but it is less than 50%. The progeny in this case received two dominant alleles linked on one chromosome (referred to as coupling or cis arrangement).

However, after crossover, some progeny could have received one parental chromosome with a dominant allele for one trait (e.g., Purple) linked to a recessive allele for a second trait (eg round) with the opposite being true for the other parental chromosome (e.g., red and long). This is referred to as repulsion or a Trans-arrangement.

The phenotype here would still be purple and long but a test cross of this individual with the recessive parent would produce progeny with much greater proportion of the two crossover phenotypes. While such a problem may not seem likely from this example, unfavorable repulsion linkages do appear while breeding for disease resistance in some crops.

When two genes are located on the same chromosome, the chance of a crossover producing recombination between the genes is directly related to the distance between the two genes. Thus, the use of recombination frequencies has been used to develop linkage maps or genetic maps.

v. Gene Mapping in Humans by Pedigree Analysis:

To map human chromosomes, obviously one cannot perform controlled mating experiments. However, it is possible to estimate map positions by examining linkage in several generations of relatives. This means that only limited data are available, and their interpretations is often difficult because a human marriage rarely results in a convenient test cross, and often the genotypes of one or more family members are unobtainable because those individuals are dead or unwilling to cooperate.

For example, blood samples from several large Mormon families in Utah, where all the members of at least three generations were alive to be sampled, have been collected and stored. These have already been used to establish genetic linkage relationships and will be available in the years ahead to study other human genes as they are identified.

2. Molecular Markers in Physical Mapping:

Different types of molecular markers are used to understand and ascertain relationship in different organisms/individuals as well as to detect or diagnose character. These markers are to locate certain characteristics on the gel (banding pattern) which can be used to detect a specific character/defect in the genome. Unlike genetic mapping, physical mapping is not to locate the genes/characters on a genome, but to create a unique pattern by processing the genomic DNA.

There are several molecular markers available which are used depending upon the objective of the work and facilities available at the centre. Use of these markers to create maps (e.g., electrophoretic patterns) of an organism is known as 'physical mapping'. Molecular markers used in physical mapping are described below. New technologies are also developed simultaneously to resolve biological problems and help legal proceedings.

Restriction Fragment Length Polymorphism (RFLP):

RFLP is a method used by molecular biologists to follow a particular sequence of DNA as it is passed on to other cells. RFLPs can be used in many different settings to accomplish different objectives. RFLPs can be used in paternity cases or criminal cases to determine the source of a DNA sample.

RFLPs can be used to determine the disease status of an individual. RFLPs can be used to measure recombination rates which can lead to a genetic map with the distance between RFLP loci measured in centiMorgans.

RFLP, as a molecular marker, is specific to a single clone/restriction enzyme combination. It is a difference in homologous DNA sequences that can be detected by the presence of fragments of different lengths after digestion of the DNA samples in question with specific restriction endonucleases. Most RFLP markers are co-dominant (both alleles in heterozygous sample will be detected) and highly locus-specific.

An RFLP probe is a labeled DNA sequence that hybridizes with one or more fragments of the digested DNA sample after they were separated by gel electrophoresis, thus revealing a unique blotting pattern characteristic to a specific genotype at a specific locus. Short, single- or low-copy genomic DNA or cDNA clones are typically used as RFLP probes. The RFLP probes are

frequently used in genome mapping and in variation analysis (genotyping, forensics, paternity tests, hereditary disease diagnostics, etc.) (Fig. 21.5).

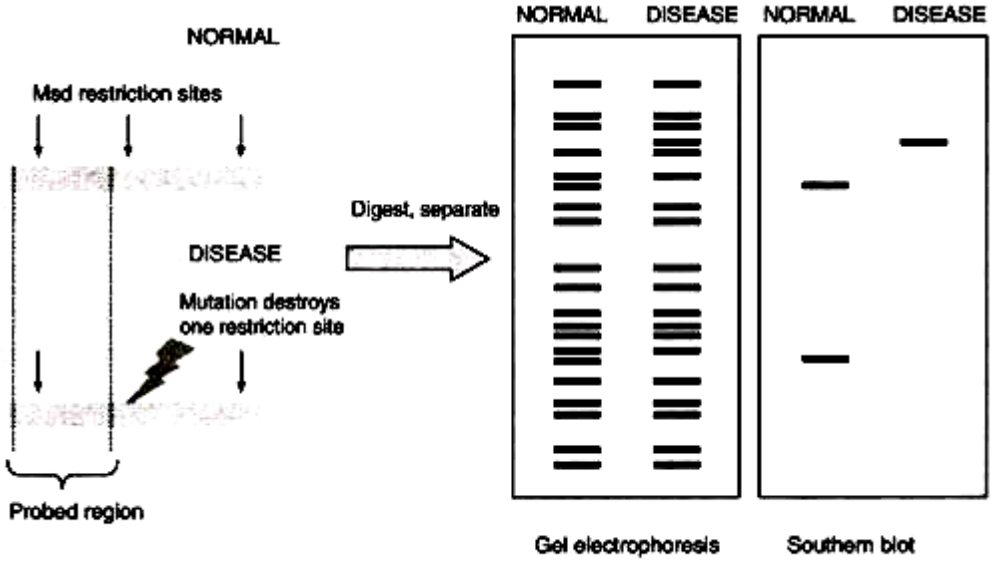


Fig. 21.5. Disease diagnosis by RFLP.

i. Procedure:

Usually, DNA from an individual specimen is first extracted and purified. Purified DNA may be amplified by polymerase chain reaction (PCR), The DNA is then cut into restriction fragments using suitable endonucleases, which only cut the DNA molecule where there are specific DNA sequences, termed recognition sequence or restriction sites that are recognized by the enzymes.

These sequences are specific to each enzyme, and may be either four, six, eight, ten or twelve base pairs in length. The more base pairs there are in the restriction site, the more specific it is and the lower the probability that it will find a place to be cut. The restriction fragments are then separated according to length by agarose gel electrophoresis. The resulting gel may be enhanced by Southern blotting. Alternatively, fragments may be visualized by pre-treatment or post-treatment of the agarose gel, using methods such as ethidium bromide staining or silver staining respectively.

RFLPs have provided valuable information in many areas of biology, including: screening human DNA for the presence of potentially deleterious genes (Fig. 21.6). Providing evidence to establish the innocence of or a probability of the guilt of, a crime suspect by DNA "fingerprinting". The distance between the locations cut by restriction enzymes (the restriction sites) varies between individuals, due to insertions, deletions or trans-versions.

This causes the length of the fragments to vary, and the position of certain amplicons differs between individuals (thus polymorphism). This can be used to genetically tell individuals apart. It can also show the genetic relationship between individuals, because children inherit genetic elements from their parents. Mitochondrial DNA RFLP analyses can lead to the determination of maternal relationships.

Fragments may also be used to determine relationships among and between species by comparison of the resulting haplotypes (abridged for 'haploid genotype'). RFLP is a technique used in marker assisted selection. Terminal Restriction Fragment Length Polymorphism (TRFLP or sometimes T-RFLP) is a molecular biology technique initially developed for characterizing bacterial communities in mixed-species samples. The technique has also been applied to other groups including soil fungi.

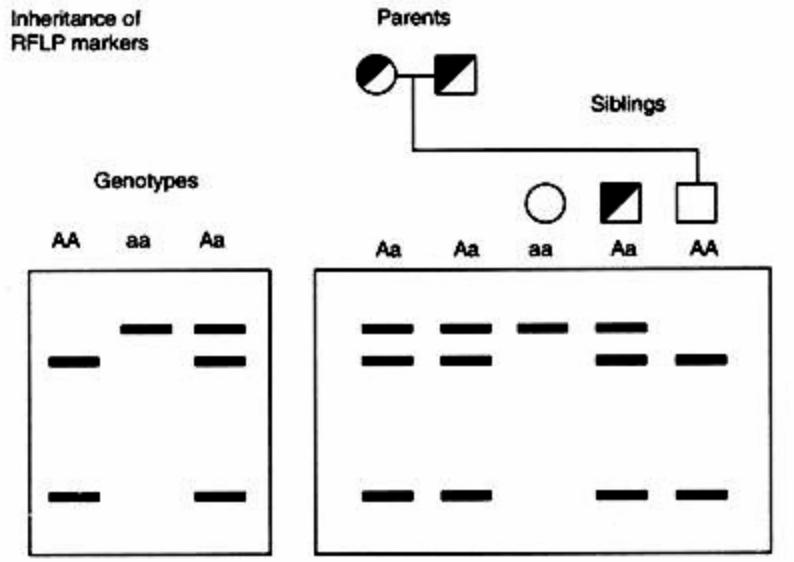


Fig. 21.6. Inheritance of RFLP markers.

The technique works by PCR amplification of DNA using primer pairs that have been labelled with fluorescent tags. The PCR products are then digested using RFLP enzymes and the resulting patterns visualized using a DNA sequencer. The results are analyzed either by simply counting and comparing bands or peaks in the TRFLP profile, or by matching bands from one or more TRFLP runs to a database of known species.

ii. Measurement of distance between two RFLP loci:

To calculate the genetic distance between two loci, you need to be able to observe recombination. Traditionally, this was performed by observing phenotypes but with RFLP analysis, it is possible to measure the genetic distance between two RFLP loci whether they are a part of genes or not. Let's look at a simple example in fruit flies. Two RFLP loci with two RFLP bands possible at each locus (Fig. 21.7).

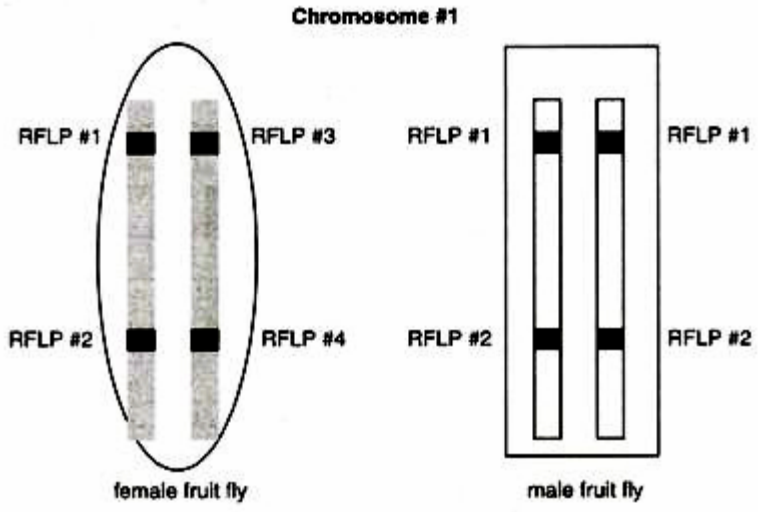


Fig. 21.7. RFLP loci on male and female chromosome.

These loci are located on the same chromosome for the female (left) and the male (right). The upper locus can produce two different bands called 1 and 3. The lower locus can produce bands called 2 or 4. The male is homozygous for band 1 at the upper locus and 2 for the lower locus. The female is heterozygous at both loci. Their RFLP banding patterns can be seen on the Southern blot below (Fig. 21.8).

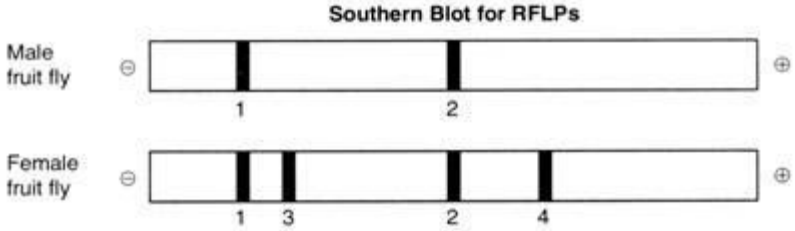


Fig. 21.8. RFLP loci on gel of male and female chromosome.

The male can only produce one type of gamete (1 and 2) but the female can produce four different gametes. Two of the possible four are called parental because they carry both RFLP bands from the same chromosome; 1 and 2 from the left chromosome or 3 and 4 from the right chromosome. The other two chromosomes are recombinant because recombination has occurred between the two loci and thus the RFLP bands are mixed so that 1 is now linked to 4 and 3 is linked to 2.

Type of chromatid	Alleles
Parental	RFLP 1 and 2
Parental	RFLP 3 and 4
Recombinant	RFLP 1 and 4
Recombinant	RFLP 3 and 2

When these two flies mate, the frequency of the four possible progeny can be measured and from this information, the genetic distance between the two RFLP loci (upper and lower) can be determined (Fig. 21.9).

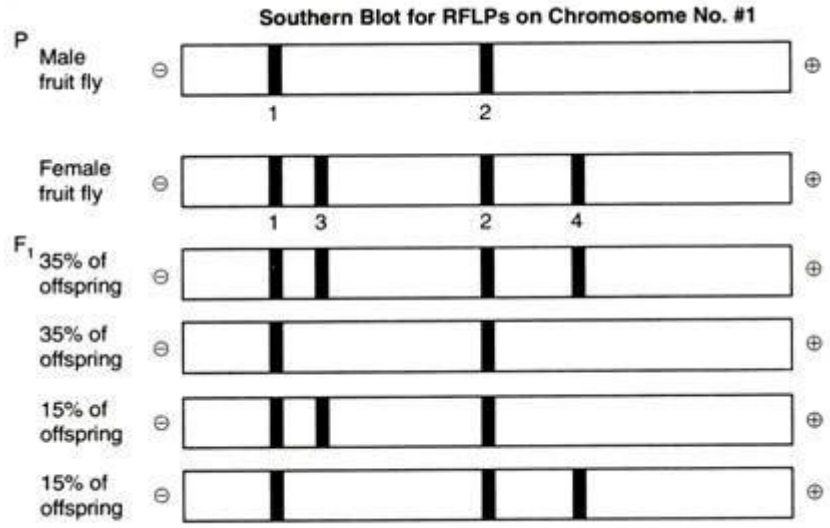


Fig. 21.9. Southern blot for RFLP's on chromosome no. 1.

In this example, 70% of the progeny were produce from parental genotype eggs and 30% were produced by recombinant genotype eggs. Therefore, these two RFLP loci are 30 centiMorgans apart from each other.

iii. PCR-RFLP:

Isolation of sufficient DNA for RFLP analysis is time consuming and labor intensive. However, PCR can be used to amplify very small amounts of DNA, usually in 2-3 hours, to the levels required for RFLP analysis. Therefore, more samples can be analyzed in a shorter time. An alternative name for the technique is Cleaved Amplified Polymorphic Sequence (CAPS) assay.

iv. Limitations:

RFLP is a multistep procedure involving restriction enzymatic cleavage, electrophoresis, southern blotting and detection of specific sequences. It is a time consuming process.

Random Amplified Polymorphic DNA (RAPD):

This technique can be used to determine taxonomic identity, assess kinship relationships, detect inter-specific gene flow, analyze hybrid speciation, and create specific probes. Advantages of RAPDs include suitability for work on anonymous genomes, applicability to work where limited DNA is available, efficiency and low expense. It is also useful in distinguishing individuals, cultivars or accessions. RAPDs also have applications in the identification of asexually reproduced plant varieties for forensic or agricultural purposes, as well as ecological ones.

In RAPD by using different primers, molecular characters can be generated that are diagnostic at different taxonomic levels. This is really a stripped-down version of PCR but uses a single sequence in the design of the primer (i.e., two primers are still needed for PCR: the same primer is used at either end).

The primer may be designed specifically, but could be chosen randomly and is used to amplify a series of samples which will include both the material of interest as well as other control samples with which the experimental material needs to be compared. Choice of primer length will be critical to the determination of band complexity in the resulting amplification pattern. Eventually a particular probe will be found that is able to distinguish between the sample of interest and those that are different.

i. Procedure:

Unlike traditional PCR analysis, RAPD (pronounced 'rapid') does not require any specific knowledge of the DNA sequence of the target organism: the identical 10-mer primers will or will not amplify a segment of DNA, depending on positions that are complementary to the primers' sequence. For example, no fragment is produced if primers annealed too far apart or 3′ ends of the primers are not facing each other.

Therefore, if a mutation has occurred in the template DNA at the site that was previously complementary to the primer, a PCR product will not be produced, resulting in a different pattern of amplified DNA segments on the gel (Fig. 21.10). RAPD is an inexpensive yet powerful typing method for many bacterial species

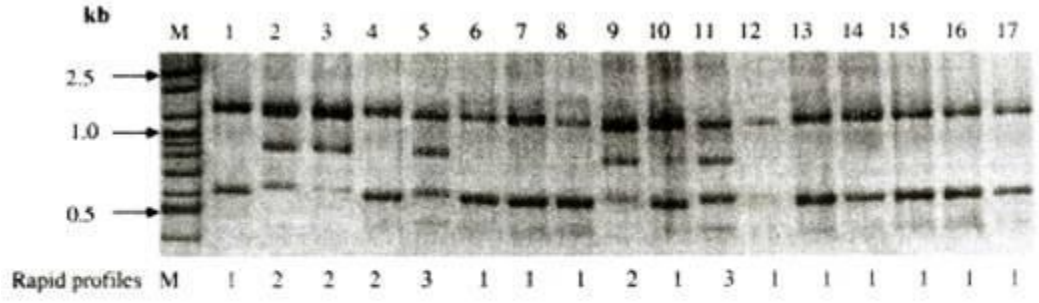


Fig. 21.10. Silver-stained polyacrylamide gel showing three distinct RAPD profiles generated by primer OPE15 for *Haemophilus ducreyi* isolates from different countries. Selecting the right sequence for the primer is very important because different sequences will produce different band patterns and possibly allow for a more specific recognition of individual strains.

RAPD amplification products can be either variable (polymorphic) or constant (non-polymorphic). In a RAPD analysis of several individuals within a species, and species within a genus, constant fragments diagnostic for a genus may be identified, as well as fragments which are polymorphic between species of the genus. RAPDs can be applied to analyze fusion of genotypes at different taxonomic levels. At the level of the individual, RAPD markers can be applied to parentage analysis, while at the population level, RAPD can detect hybrid populations, species or subspecies.

The detection of genotype hybrids relies on the identification of diagnostic RAPD markers for the parental genotypes under investigation. However RAPD markers tend to underestimate genetic distances between distantly related individuals, for example in inter-specific comparisons.

It is wise to be cautious when using RAPD for taxonomic studies above the species level. Conventional RFLP techniques are ill-suited for the analysis of paternity and estimation of reproductive success in species with large offspring clutches, because of the need to determine

paternity for each individual offspring. RAPD fingerprinting provides a ready alternative for such cases.

Synthetic offspring may be produced by mixing equal amounts of the DNA of the mother and the potential father. The amplification products from the synthetic offspring should ideally contain the full complement of bands that appear in any single offspring of these parents (Table 21.2).

ii. Limitations of RAPD:

1. Nearly all RAPD markers are dominant, i.e., it is not possible to distinguish whether a DNA segment is amplified from a locus that is heterozygous (1 copy) or homozygous (2 copies). Co-dominant RAPD markers, observed as different-sized DNA segments amplified from the same locus, are detected only rarely.
2. PCR is an enzymatic reaction, therefore the quality and concentration of template DNA, concentrations of PCR components, and the PCR cycling conditions may greatly influence the outcome. Thus, the RAPD technique is notoriously laboratory dependent and needs carefully developed laboratory protocols to be reproducible.
3. Mismatches between the primer and the template may result in the total absence of PCR product as well as in a merely decreased amount of the product. Thus, the RAPD results can be difficult to interpret.

Amplification Fragment Length Polymorphism (AFLP):

Amplified Fragment Length Polymorphism (AFLP) is a polymerase chain reaction (PCR) based genetic fingerprinting technique that was developed in the early 1990's by Keygene. AFLP can be used in the fingerprinting of genomic DNA of varying origins and complexities. The amplification reaction is rigorous, versatile and robust, and appears to be quantitative.

While AFLP is capable of producing very complex fingerprints (100 bands where RAPD produces 20), it is a technique that requires DNA of reasonable quality and is more experimentally demanding. AFLP uses restriction enzymes to cut genomic DNA, followed by ligation of complementary double stranded adaptors to the ends of the restriction fragments.

A subset of the restriction fragments are then amplified using 2 primers complementary to the adaptor and restriction site fragments. The fragments are visualized on denaturing polyacrylamide gels either through auto-radiographic or fluorescence methodologies.

i. Procedure:

AFLP-PCR is a highly sensitive method for detecting polymorphisms in DNA. The technique was originally described by Vos and Zabeau in 1993. The procedure of this technique is divided into three steps (Fig. 21.11):

1. Digestion of total cellular DNA with one or more restriction enzymes that cuts frequently (Msel, 4 bp recognition sequences) and one that cuts less frequently (EcoRI, 6 bp recognition sequence). The resulting fragments are ligated to end-specific adaptor molecules.
2. Selective amplification of some of these fragments with two PCR primers that have corresponding adaptor and restriction site specific sequences.
3. Electrophoretic separation of amplicons on a gel matrix, followed by visualization of the band pattern.

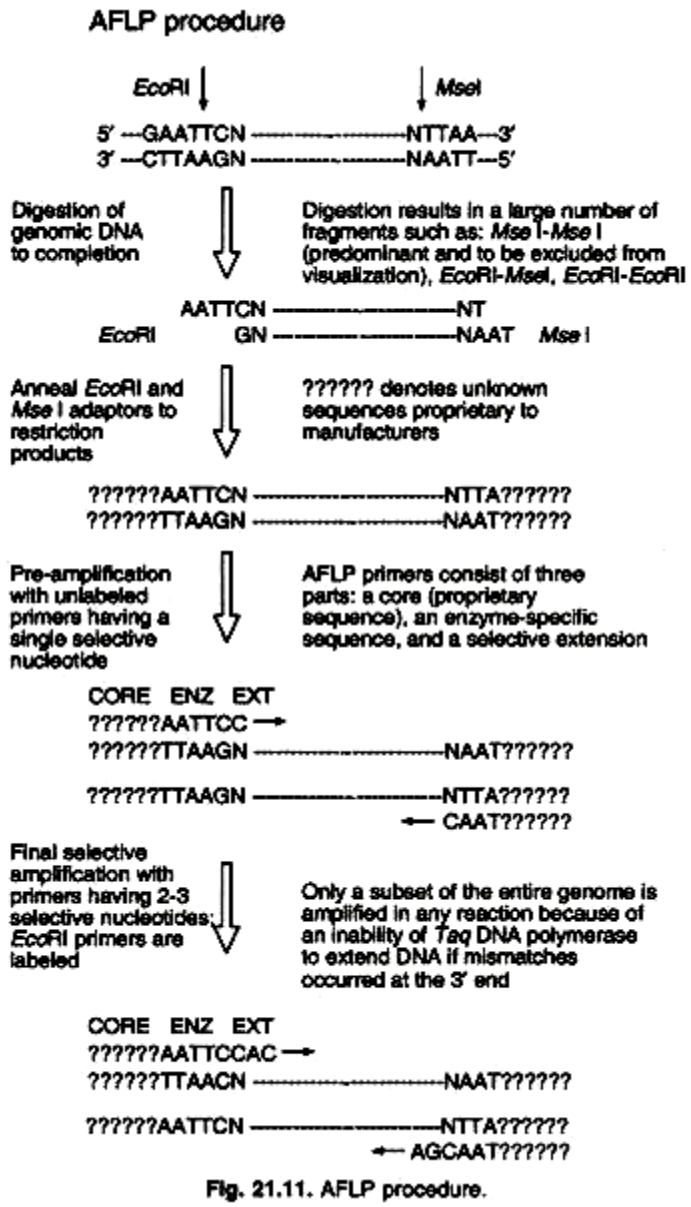


Fig. 21.11. AFLP procedure.

In a second, "selective", PCR, using the products of the first as template, primers containing two further additional bases, chosen by the user, are used. The EcoRI-adaptor specific primer used bears a label (fluorescent or radioactive). Gel electro-phoretic analysis reveals a pattern (fingerprint) of fragments representing about 1/4000th of the EcoRl-Msel fragments.

AFLP's, can be co-dominant markers, like RFLP's. Co-dominance results when the polymorphism is due to sequences within the amplified region. Yet, because of the number of bands seen at one time, additional evidence is needed to establish that a set of bands result from different alleles at the same locus.

If, however, the polymorphism is due to presence/absence of a priming site, the relationship is dominance. The non-priming allele will not be detected as a band. Compared to RAPD, fewer primers should be needed to screen all possible sites. AFLP can be used for mapping, fingerprinting and genetic distance calculation between genotypes. The advantage of AFLP is its high multiplexity and therefore the possibility of generating high marker densities.

One limitation of the AFLP technique is that fingerprints may share few common fragments when genome sequence homology is less than 90%. Therefore, AFLP cannot be used in comparative genomic analysis with hybridization-based probes or when comparing genomes that are evolving rapidly such as those of some microbes. Conversely, very homogeneous genomes may not be suitable for AFLP analysis.

A study on the genetic diversity of an endangered alpine plant (Eryngium alpinum L. (Apiaceae) demonstrated that AFLP markers enable a quick and reliable assessment of intraspecific genetic variability in conservation genetics. The study showed that although the endangered plant occurred in small isolated populations, these populations contained a high genetic diversity, a good indication that recovery of the species was possible.

ii. Limitations of AFLP:

1. Proprietary technology is needed to score heterozygotes and homozygotes. Otherwise, AFLP must be dominantly scored.
2. Developing locus-specific markers from individual fragments can be difficult.
3. Need to use different kits adapted to the size of the genome being analyzed.

Microsatellites:

Microsatellites can be used to determine genetic diversity within a species, as well as being able to distinguish varieties and even individuals, as well as parentage. The distribution of genetic variability is commonly used to verify species, subspecies or population division. Monitoring change in diversity may also be useful for predicting populations in peril as the persistence of a population partially depends on maintaining its evolutionary significance which requires genetic variation.

Microsatellites have been used to estimate demographic bottlenecks in some species. A bottleneck, when it severely and temporarily reduces population size, can also drastically reduce the genetic diversity of a population. A common theme in conservation genetics is the use of genetic variation to identify populations that have experienced bottlenecks, as numerous threatened or endangered species and populations have been found to have low levels of genetic variation.

Inter-Simple Sequence Repeats (ISSR):

ISSRs can be used to assess hybridization in natural populations of plants, as a study on *Penstemon (Scrophulariaceae)* did. Eight ISSR primers were used to examine patterns of hybridization and hybrid speciation in a hybrid complex involving four species, as well as allowing examination of pollen-mediated gene flow. Previous studies using allozymes, restriction-site variation of nuclear rDNA and chloroplast DNA failed to determine whether gene flow occurs among species other than *P. cenranthifolius*.

The previous studies also failed to provide support for hypotheses of diploid hybrid speciation. ISSR proved to be a much more successful technique in this study, allowing all species and all

DNA accessions to be differentiated. ISSR has also been used to detect varieties and diversity in rice, revealing much more data than RFLPs. The technique allowed for dissection below the subspecies level and this gives it a good level of applicability in the study of rare or endangered plants.

ISSRs have been used in conjunction with RAPD data to determine the colonization history of *Olea europaea* in Macronesia, along with lineages in the species complex. The two techniques have also been utilized in examining the historical biogeography of Sea rocket (*Cakile maritima*) and Sea Holly (*Eryngium maritimum*), comparing different and only distantly related taxa of broadly similar extant distribution. The trees generated by the different methods were largely similar topologically. Using the result, dispersal routes of the species along a linear coast line could be construed.

Joint use of RAPD and ISSR has also been used to examine clonal diversity in *Calamagrostis porteri ssp. insperata* (Poaceae), a rare grass that has little or no sexual reproduction, and spreads by vegetative reproduction. The relative advantages and disadvantages of various molecular markers in physical mapping are summarized in Table 21.3. This information suggests that RFLP, SSR and AFLP markers are most effective in detecting polymorphism.

However, given the large amount of DNA required for RFLP detection and the difficulties in automating RFLP analysis, AFLP and SSR are currently most popular markers.

Table 21.3. Comparison of most commonly used marker systems.

Features	RFLP	RAPD	AFLP	SSR	SNP
DNA required (µg)	10	0.02	0.5-1.0	0.05	0.05
DNA quality	High	High	Moderate	Moderate	High
PCR based	No	Yes	Yes	Yes	Yes
No. of polymorph loci analysed	1.0-3.0	1.5-5.0	20-100	1.0-3.0	1.0
Ease of use	Not easy	Easy	Easy	Easy	Easy
Amenable to automation	Low	Moderate	Moderate	High	High
Reproducibility	High	unreliable	High	High	High
Development cost	Low	Low	Moderate	High	High
Cost per analysis	High	Low	Moderate	Low	Low

The main uses of these markers include:

1. Assessment of genetic variability and characterization of germplasm.
2. Identification and fingerprinting of genotypes.
3. Estimation of genetic distances between population, inbreeds and breeding material.
4. Detection of monogenic and qualitative trait loci.
5. Marker assisted selection.
6. Identification of sequences of useful candidate genes.

Thing # 3. Restriction Mapping of DNA Fragments:

Restriction Mapping:

Genetic mapping using RFLPs as DNA markers can locate the positions of polymorphic restriction sites within a genome, but very few of the restriction sites in a genome are polymorphic, so many sites are not mapped by this technique.

We increase the marker density on a genome map by using an alternative method to locate the positions of some of the non polymorphic restriction sites. This is what restriction mapping achieves, although in practice the technique has limitations that means it is applicable only to relatively small DNA molecules.

Methodology for Restriction Mapping:

The simplest way to construct a restriction map is to compare the fragment sizes produced when a DNA molecule is digested with two different restriction enzymes that recognize different target sequences. An example using the restriction enzymes EcoRI and BamHI is shown in figure. 21.12. First, the DNA molecule is digested with just one of the enzymes and the sizes of the resulting fragments measured by agarose gel electrophoresis. Next, the molecule is digested with the second enzyme and the resulting fragments again sized in an agarose gel.

The results of subsequent use of two enzymes give clear picture about restriction sites creating a large number of fragments but this method do not allow their relative positions to be determined. Additional information is therefore obtained by cutting the DNA molecule with both enzymes together. In the example shown in Figure 21.12, the double restriction enables three of the sites to be mapped. However, a problem arises with the larger EcoRI fragment because this contains two BamHI sites and there are two alternative possibilities for the map location of the outer one of these.

The problem dissolved by going back to the original DNA molecule and treating it again with BamHI on its own, but this time preventing the digestion from going to completion by, for example, incubating the reaction for only a short time rousing a suboptimal incubation temperature. This is called a partial restriction and leads to a more complex set of products. The complete restriction products now being supplemented with partially restricted fragments that still contain one or more uncut BamHI sites.

In the example shown in Figure 21.12, the size of one of the partial restriction fragments is diagnostic and the correct map can be identified. A partial restriction usually gives the information needed to complete a map, but if there are many restriction sites then this type of analysis becomes bulky, simply because there are many different fragments to consider. An alternative strategy is simpler because it enables the majority of the fragments to be ignored. This is achieved by attaching a radioactive or other type of marker to each end of the starting DNA molecule before carrying out the partial digestion.

The result is that many of the partial restriction products become “invisible” because they do not contain an end-fragment and so do not show up when the agarose gel is screened for labeled products (Fig. 21.12). The sizes of the partial restriction products that are visible enable unmapped sites to be positioned relative to the ends of the starting molecule.

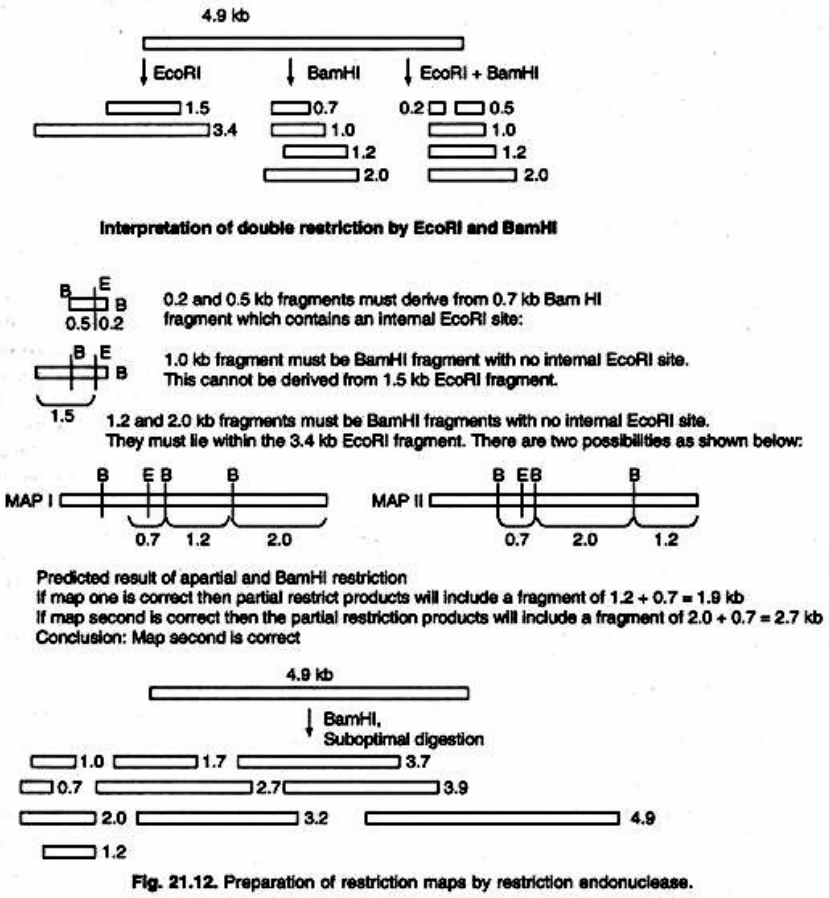


Fig. 21.12. Preparation of restriction maps by restriction endonuclease.

The scale of restriction mapping is limited by the sizes of the restriction fragments. Restriction maps are easy to generate if there are relatively few cut sites for the enzymes being used. However, as the number of cut sites increases, so also do the numbers of single, double and partial-restriction products whose sizes must be determined and compared in order for the map to be constructed. Computer analysis can be brought into play but problems still eventually arise.

A stage will be reached when a digest contains so many fragments that individual bands merge on the agarose gel, increasing the chances of one or more fragments being measured incorrectly or missed out entirely. If several fragments have similar sizes then even if they can all be identified, it may not be possible to assemble them into a clear map. Restriction mapping is therefore more applicable to small rather than large molecules, with the upper limit for the technique depending on the frequency of the restriction sites in the molecule being mapped.

In practice, if a DNA molecule is less than 50 kb in length it is usually possible to construct a clear restriction map for a selection of enzymes with six nucleotide recognition sequences. Restriction maps are equally useful after bacterial or eukaryotic genomic DNA has been cloned, if the cloned fragments are less than 50kb in length, because a detailed restriction map can then be built up as a preliminary to sequencing the cloned region. This is an important application of restriction mapping in projects sequencing large genomes.

Restriction mapping can be used for mapping of entire genomes larger than 50kb by slightly eliminating the limitations of restriction mapping by choosing enzymes expected to have infrequent cut sites in the target DNA molecule.

These "rare cutters" fall into two categories:

1. A few restriction enzymes cut at seven- or eight-nucleotide recognition sequences. Examples are SapI (5′-GCTCTTC-3′) and SgfI (5′-GCGATCGC-3′). The enzymes with seven-nucleotide recognition sequences would be expected, on average, to cut a DNA molecule with GC content of 50% once every 47 = 16,384 bp.

The enzymes with eight nucleotide recognition sequences should cut once every 48 = 65,536 bp. These figures compare with 46 = 4096 bp for enzymes with six-nucleotide recognition sequences, such as BamHI and EcoRI.

Cutters with seven-or eight-nucleotide recognition sequences are often used in restriction mapping of large molecules, but the approach is not as useful as it might by simply because not many of these enzymes are known.

2. Enzymes can be used whose recognition sequences contain motifs that are rare in the target DNA. Genomic DNA molecules do not have random sequences and some molecules are significantly deficient in certain motifs. For example, the sequence 5′-CG- 3′ is rare in the genomes of vertebrates because vertebrate cells possess an enzyme that adds a methyl group to carbon 5 of the C nucleotide in this sequence.

Domination of the resulting 5-methylcytosine gives thymine. The consequence is that during vertebrate evolution, many of the 5′-CG'3 sequences that were originally in these genomes have become converted to 5′-TG-3′.

Restriction enzymes that recognize a site containing 5′- CG-3′ therefore cut vertebrate DNA relatively infrequently. Examples are SmaI (5′-CCCGGG-3′), which cuts human DNA once every 78 kb on average, and BssHII (5′- GCGCGC'3′), which cuts once every 390 kb.

The potential of restriction mapping is therefore increased by using rare cutters. It is still not possible to construct restriction maps of the genomes of animals and plants, but it is feasible to use the technique with large cloned fragments, and with the smaller DNA molecules of prokaryotes and lower eukaryotes such as yeast and fungi.

If a rare cutter is used then it may be necessary to employ a special type of agarose gel electrophoresis to study the resulting restriction fragments. This is because the relationship between the length of DNA molecule and its migration rate in an electrophoresis gel is not linear, the resolution decreasing as the molecules get longer (Fig. 21.13A). This means that it is not possible to separate molecules more than about 50 kb in length, because all of these longer molecules run as a single, slowly migrating band in a standard agarose gel.

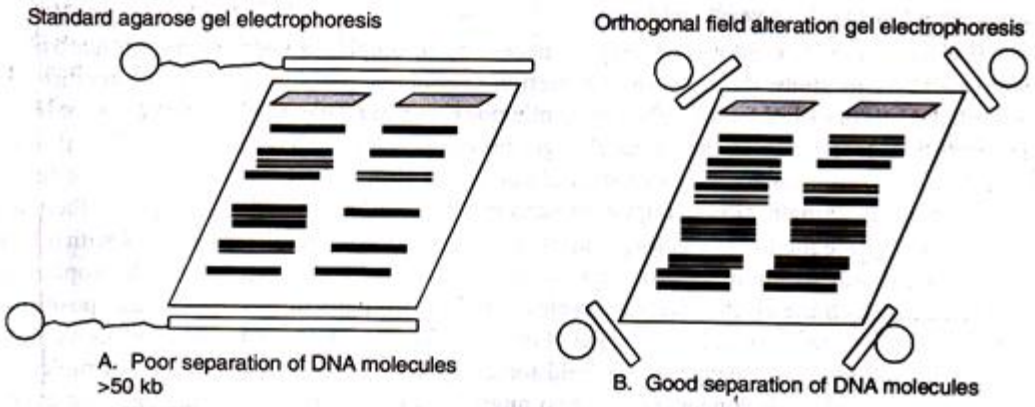


Fig. 21.13. A,B Conventional and non-conventional agarose gel electrophoresis.

To separate them it is necessary to replace the linear electric field used in conventional gel electrophoresis with a more complex field. An example is provided by orthogonal field alternation gel electrophoresis (OFAGE), in which the electric field alternates between two pairs of electrodes, each positioned at an angle of 45″ to the length of the gel (Fig. 21.13B).

The DNA molecules still move down through the gel, but each change in the field forces the molecules to realign. Shorter molecules realign more quickly than longer ones and so migrate more rapidly through the gel. The overall result is that molecules much longer than those separated by conventional gel electrophoresis can be resolved. Related techniques include CHEF (contour clamped homogeneous electric fields) and FIGE (field inversion gel electrophoresis).

Centromere mapping with ordered tetrads

In *Neurospora*, products of meiosis are present in a linear order, so that **ordered tetrads** are obtained. Such a situation is not found in *Chlamydomonas* although tetrad analysis can be conducted. Due to ordered tetrads in *Neurospora*, a cross between normal *(a^+)* and mutant *(a)* strain, will give rise to a linear arrangement of four normal spores *(a^+)* at one end followed by four mutant spores (*a*) at the other end. Such an arrangement would be disturbed if crossing over occurs between centromere and the gene, because crossing over occurs at four strand and not at two strand stage. If crossing over at two strand stage was possible, it would again lead to four normal spores followed by four mutant spores. Therefore, by analysing linear arrangement, a 4 : 4 arrangement will suggest absence of crossing over and a paired arrangement (24 : *2a* : 24 :2*a* or

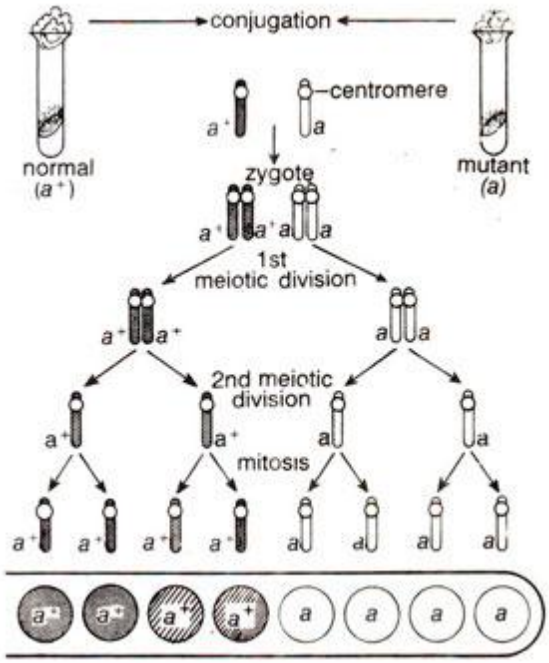


Fig. 11.6. Results showing first division

2*A* :4*a* :2*A or* 2*a* :4*A* :2*a)* will suggest that crossing over has taken place between the gene and corresponding centromere. When crossing over is absent leading to 4 : 4 arrangement, this is described as **First Division Segregation** (Fig. 11.6) and when crossing over takes place leading to paired arrangement, it is described as **Second Division Segregation.** (Fig. 11.7, 11.8).

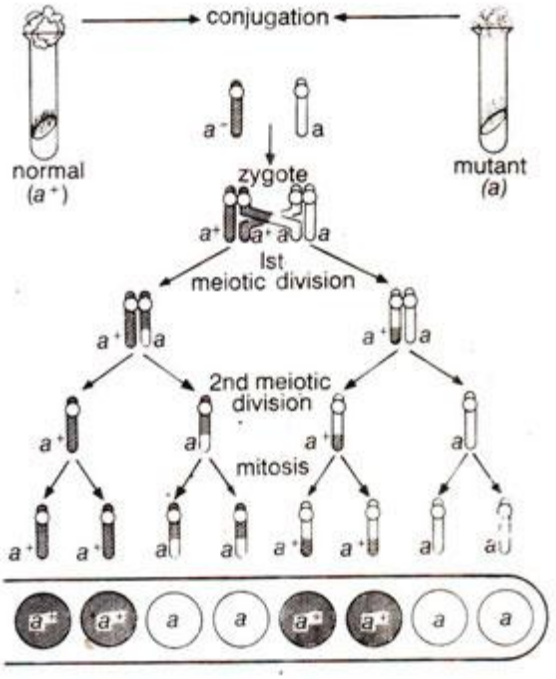


Fig. 11.7. Results of a linear order of ascospores (2:2:2:2) showing second division segregation in a cross between normal *(a*$^{+}$*)* and mutant *(a)* strains of *Neurospora*, in which crossing over takes place between the gene and the centromere.

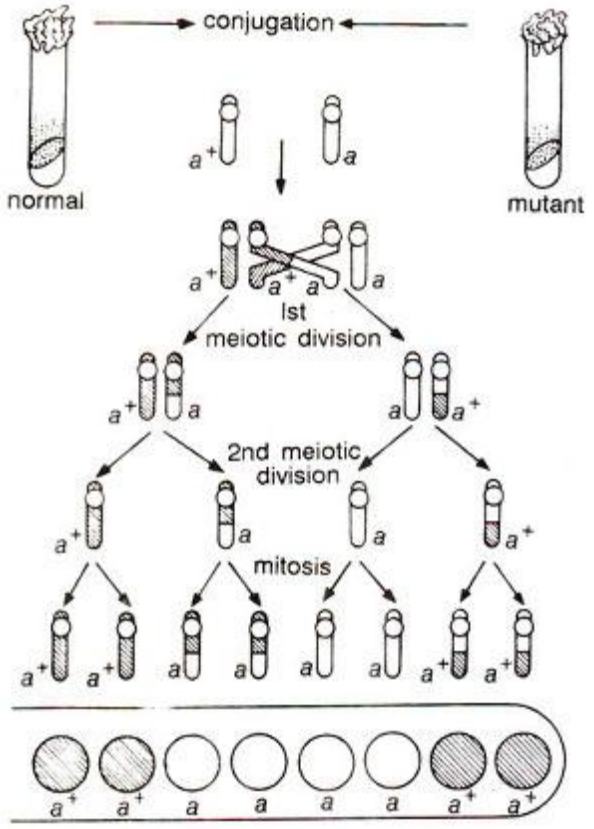


Fig. 11.8. Results of a linear order of ascospores (2:4:2) which is different than the one shown in Figure 11.7. but still showing second division segregation in a cross between normal *(a*$^{+}$*)* and mutant (a) strains of *Neurospora*, in which crossing over takes place between the gene and the centromere.

segregation in a cross between normal (*a*$^{+}$) and mutant *(a)* strains of-*Neurospora*, in which crossing over between the gene and centromere does not take place.

Frequency of recombination between centromere and gene. In *Neurospora,* since crossing over between a gene and centromere leads to second division segregation, the relative proportion of asci exhibiting second division segregation will give an estimate of crossing over. This will not be possible in *Chlamydomonas* due to absence of ordered tetrads. As an illustration, in an experiment in *Neurospora,* cross between albino (*al*) and wild type *(al^+)* gave first division segregation *(al al al^+ al^+)* in 129 asci and second division segregation *(al al^+ al al^+)* in 141 tetrads. Since only two of the four strands undergo crossing over in second division segregation, recombination between centromere and the gene is

$$\frac{1}{2} \times \frac{141}{(129+141)} = 0.26 \text{ or } 26 \text{ per cent.}$$

If a gene is located far away from centromere, crossing over may take place between the gene and centromere in each and every tetrad so that 100% tetrads should exhibit second division segregation. However, since in such cases, double crossovers, and multiple crossovers are also possible, some of the tetrads due to double crossover and other even number of crossover events, will give rise to first division segregation. Thus, second division segregation will never reach 100% level and in practice does not exceed 67% giving rise to only 33% recombination frequency *(*½ x 67, since only two of the four chromatids are involved in crossing over). This 33% recombination frequency will actually represent 50 units or more. Therefore, recombination frequencies should not be estimated over long distances, but should be estimated over small distances to avoid underestimation due to double and even multiple crossovers.

Detection of linkage between two genes. We have already discussed that tetrads can be of three types (i) **parental ditype (PD), (ii) non-parental ditype (NPD)** and **(iii) tetratype (T).** The linkage will be detected by comparing your results with the results one would expect on the basis of independent assortment. We have already discussed that if two parents *AB* and *ab* give rise to zygote *AaBb,* it will give PD (2*AB* : 2*ab),* NPD (2*Ab* : 2*aB*)and T (1*AB* : 1*Ab* : 1*aB* : 1*ab).* If two genes are located on separate chromosomes and independent assortment holds good, PD and NPD tetrads should be equal and T-tetrads should be relatively fewer (since they result due to crossing over in one of the two pairs of chromosomes). On the other hand, if linkage is there, then NPD will result due to four strand double crossovers and tetratype will result due to single crossover, so that the number of NPD tetrads will be much lower, not only in comparison to PD but also in comparison to T (Fig. 11.9). This will be illustrated with the help of some data. In *Neurospora,* a cross between an adenine requiring strain *(ad)* of mating type *A* and a wild type strain (+) of mating type *a* gave the results presented in Table 11.1.

The data presented in Table 11.1 can be used for comparing PD and NPD and also for comparing NPD and T. It will be seen that PD and NPD are not significantly different, a result expected only in case of independent assortment. Further, tetratype (T) tetrads are fewer than NPD, a result, which is also expected only from independent assortment, although their frequency relative to the PD or NPD during independent assortment will be governed by the distance between the genes and their corresponding centromeres. It can, therefore, be concluded from the data that loci for *ad* and *A* are not linked and are located close to centromere (T tetrads are few). It may be emphasized here that tetrads and not the ascospores or their pairs are considered as units. For instance, if tetrad ratio is 10 PD : 6 NPD, the difference may not be significant, but the ratio of meiotic products will be 40 PD : 24 NPD and the ratio of meiotic products will be 80 PD : 48 NPD, the differences in these cases being significant, thus leading us to erroneous conclusion of linkage.

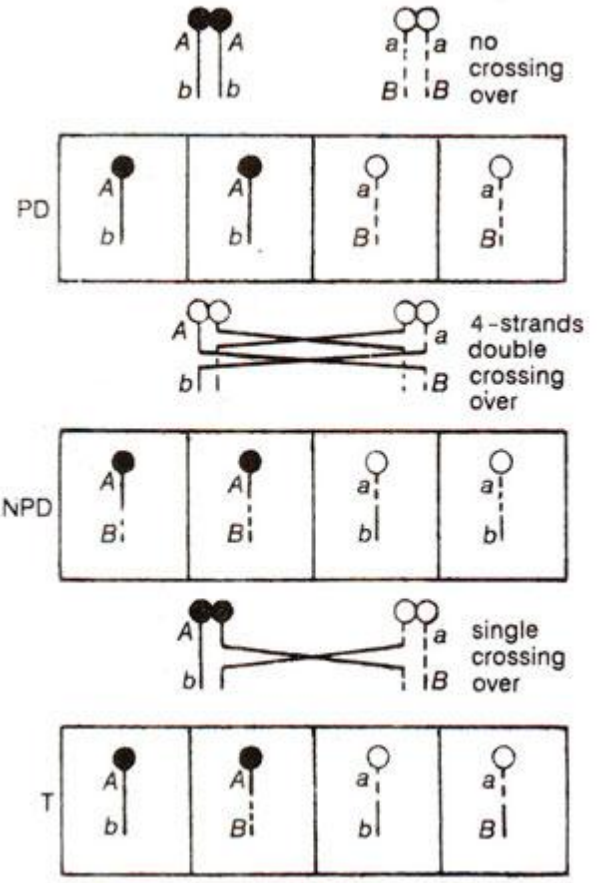


Fig. 11.9. Production of ordered tetrads classified as parental ditype (PD), non-parental ditype (NPD) and tetratype (T) in a dihybrid *(AaBb)*, with the genes *A* and *B* linked together (modified from Strickberger's "Genetics").

In another example of *Neurospora*, an albino strain *(al)* of mating type *A*, and a wild type strain (+) of mating type *a* produced the results shown in Table 11.2.

Since in this case, difference between PD (24) and NPD (3) is statistically significant and tetratypes exceed the number of NPD, the data suggest presence of linkage between the two loci.

Table 11.1. **Results of tetrad analysis in a cross *adA* × *+ a* in *Neurospora* (*ad* = adenine requiring, *a* = mating type).**

Parental ditype (PD)	Non-parental ditype (NPD)	Tetratype (T)
ad A *ad A* *+ a* *+ a*	*ad a* *ad a* *+ A* *+ A*	*ad a* *ad A* *+ a* *+ A*
frequency 10	9	1

Table 11.2. **Results of tetrad analysis from a cross *alA* × *+a* in *Neurospora* (*al* = albino; *a* = mating type).**

Parental ditype (PD)	Non-parental ditype (NPD)	Tetratype (T)		
		1	2	3
al A *al A* *+ a* *+ a*	*al a* *al a* *+ A* *+ A*	*al a* *+ a* *al A* *+ A*	*al a* *al A* *+ a* *+ A*	*al a* *+ A* *al A* *+ a*
24	3		27	

Preparation of linkage map. Data on tetrad analysis can be further utilized for determining linkage distances between genes. Distance between two genes can be sum of the distances of each irotn the centromere, if they are located on opposite arms or it will be difference of the distances of each gene from centromere if they are on the same arm. Once linkage is detected, one has to find out whether the two genes are on the same arm or on different arms. The relative frequencies of different orders in tetrads will help in deciding this fact. The different possible ordered tetrads obtained from the cross *AB* x *ab* are given in Table 11.3.

Table 11.3. **Seven possible tetrad types (with their frequencies), in a *Neurospora* strain heterozygous at two loci (*AaBb*), resulting from first and second division segregations (M_I, M_{II}) at either of the two loci.**

Parental ditype (PD)		Non-parental ditype (NPD)		Tetratype (T)		
1	2	3	4	5	6	7
A B A B a b a b	A B a b A B a b	A b A b a B a B	A b a B A b a B	A B A b a B a b	A B a B A b a b	A B a b A b a B
M_IM_I	$M_{II}M_{II}$	M_IM_I	$M_{II}M_{II}$	M_IM_{II}	$M_{II}M_I$	$M_{II}M_{II}$
808	90	1	1	90	5	5

In the example used in Table 11.3, the two genes are linked *(AB/ab).* In this case either of the two genes or both the genes may exhibit **first division segregation (M_I)** or **second division segregation (M_{II}).** If the order of loci is centromere-*A*-*B*, then the crossing over between A and B will produce M_I segregation for *A* and M_{II} segregation for *B*, as shown in Fig. 11.9 for tetratype (T) tetrads. On the other hand, crossing over between centromere and *A* will give M_{II} segregation for both and absence of crossing over in the region 'centromere-*A*-*B*' will give M_I segregation for both the genes.

In Table 11.3, seven possible classes are given, but the order with in the half ascus was ignored, because it only reflects results of random spindle attachment. For instance PD shown in class 2 in Table 11.3, will include all the following four classes, which are equivalent, and can be obtained from each other by inverting top two and/or bottom two genotypes, as expected from random spindle attachment.

		Bb	
		M_I	M_{II}
Aa	M_I	PD, NPD	T only
	M_{II}	T only	T, PD, NPD

Fig. 11.10. The results of frrst and second division segregations (Mi, Mn) for two genes *(Aa, Bb),* showing that for the production of tetratype (T) tetrads, second division segregation for atleast one of the two genes is necessary.

The PD, NPD and T types in Table 11.3 can be obtained by different combinations of M_I and M_{II} segregations of A and *B* genes (Fig. 11.10), It may be seen that a tetratype can occur only

where there is a crossover between atleast one locus and centromere. In other words for T type tetrads atleast one gene should exhibit M_{II}segregation.

AB	*ab*	*AB*	*ab*
ab	*AB*	*ab*	*AB*
AB	*AB*	*ab*	*ab*
ab	*ab*	*AB*	*AB*

The distances of two genes from centromeres can be calculated as follows :

$$\text{Centromere-}A = \frac{90+1+5+5}{1000} = \frac{101}{1000}$$
$$= 10.1\% = 5.05 \text{ m.u.}$$
$$\text{Centromere-}B = \frac{90+1+90+5}{1000} = \frac{186}{1000}$$
$$= 18.6\% = 9.3 \text{ m.u.}$$

With the above information, following three possibilities exist : (i) *A* and *B* are on two different chromosomes, (ii) *A* and *B* are on two different arms of the same chromosome and (iii) A and *B* are on the same arm of a chromosome, A lying between the centromere and *B.* The first possibility should lead to independent assortment and can be ruled out, because PD and NPD are not equal and. T exceeds NPD. Since M_{II} segregation for both' genes exceeds the frequency of a class showing M_{II} segregation for either A or *B* alone and since single crossovers will be more frequent than double crossovers, possibility (ii) above is less probable than the possibility (iii), which will give M_{II} segregation for both genes due to single crossover between the centromere and A.

Once it is ascertained that both genes are on the same arm, recombination frequency between *A* and *B* can be calculated. This distance can not be calculated by simple subtraction 9.30 - 5.05, because in calculating 9.30 we ignored double crossovers as in class 6 showing M_I segregation for *B* (Table 11.3).

Since all the four products in NPD are recombinants and T has 50% recombinant products, the distance *A-B* can be calculated as follows and a map can be drawn (Fig. 11.11).

$$RF = \frac{NPD + \frac{1}{2}T}{\text{total asci}} \times 100$$
$$= \frac{2 + \frac{1}{2}(100)}{1000} \times 100 = 5.2 \text{ m.u.}$$

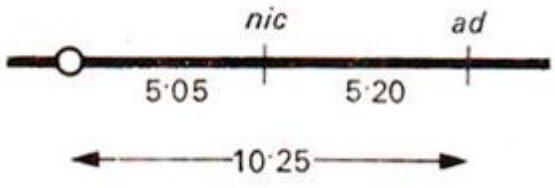


Fig. 11.11. A linkage map of genes *A* an *B* based on data presented in Table 11.3.

Cytogenetic mapping with deletions and duplications in *Drosophila*

Chromosomes in many plant and animal species can be stained with dyes such as Giemsa to reveal a pattern of dark and light staining bands. In Drosophila polytene chromosomes the

number of bands which can be seen is almost as large as the number of genes with a detectable phenotype. Each band covers about 25 kb. The Giemsa banding pattern in effect provides a high resolution (+ 100 kb) physical map of the genome. In the early days of Drosophila genetics this was largely used to map chromosomal re-arrangements such as deletions, duplications and inversions. The ability to fine map each of these types of re-arrangements has proved extremely useful. For example, deletions can be used to locate mutations on the physical map

Deletion Mapping and in situ hybridization

A form of complementation analysis can be used to determine whether a given mutation maps within a given deletion or outside it. Essentially the deletion stock and the mutant stock are crossed to generate flies carrying one chromosome of a pair of homologues carrying the deletion and the other carrying the mutation (mutation/deletion). If flies of this genotype have the mutant phenotype (the deletion fails to complement the mutation) the mutation must fall within the deleted region so the fly has no wild type copy of the gene. Conversely if flies of this genotype show a wild type phenotype(the deletion complements the mutation) the mutation must map to somewhere outside this deletion.

Collections of overlapping deletions for many regions of the Drosophila genome have been isolated and mapped. If a mutation fails to complement a number of overlapping deletions the gene must map to the region where they all overlap - the region which is deleted in all of them.

Polytene mapping can be used to locate cloned sequences as well as mutations. Cloned sequences can be used to generate labeled probes and hybridized to polytene chromosome preparations to identify the band(s) containing the cloned sequence. Cloned sequences that map to the same band as a particular mutations are good candidated for the gene disrupted in the mutation.

Duplications in P-elements and transgenesis

Unlike mice and C.elegans, Drosophila are refractory to being made transgenic by simple microinjection of transgene DNA. Integration of transgene DNA into the genome can be greatly facilitated by cloning the transgene into a vector where it will be flanked by the 31 bp terminal repeat sequences recognized by the P element transposase. It is useful to also include sequences coding for an eye colour marker gene such as rosy in these vectors. The transgene vector is then co-injected with DNA that codes for transposase in another vector that doesn't contain the repeats (helper plasmid). In the egg transposase from the helper plasmid catalyses integration of the transgene and marker gene flanked by the 31 bp repeats into the fly's genome.

Flies that hatch from the injected eggs will usually contain a small proportion of transgenic germ cells (sperm or eggs). Usually they don't show the marker gene phenotype. To make transgenic lines these flies have to bred with flies from the base stock used to derive the eggs for injection. Offspring derived from a transgenic sperm or egg will be completely transgenic and can be identified using the marker gene phenotype.

Transgenic Drosophila made in this way have some advantages over mouse and C.elegans transgenics:-

The transgene is almost always present in a single copy - mouse genes present in multiple copies are prone to having their expression shut down and the copies can interfere with each other's expression.

Because the transgenes are integrated in endogenous chromosomes their inheritance is very stable - unlike that of extrachromosomal arrays formed in C. elegans.

Inserts can be jumped to new locations by crossing transgenic strains with transposase expressing strains.

As in other organisms transgenesis in Drosophila can be used to:-

- identify regulatory elements
- misexpression studies of gene function
- generate and clone insertional mutations

Regulatory elements in a Drosophila Pax6 homologue

The Drosophila homologue of Pax6 (see lecture 3) turned out to be encoded by a locus named eyeless (ey), in which mutations had previously been isolated with a reduced eye phenotype. Ey was found to be expressed in the embryonic nervous system and in the early eye-antennal imaginal disc. Ey mutations were due to insertions in the intron 2 of the gene and affected its expression in the eye-antennal imaginal disc but not in the embryo. Transgenes in which intron 2 sequences were fused to a lacZ reporter were expressed exclusively in the eye-antennal disc. Drosophila More recently these Drosophila intron 2 sequence transgenes were also shown to confer early eye-specific expression in transgenic mice!

Misexpression of Pax6 in transgenic flies

Transgenes in which Ey coding sequences are fused to a promoter which is expressed in all imaginable discs (not just the eye-antennal disc) grow ectopic eyes on all disc derived body parts including mouthparts, legs and genitalia! This shows that Pax6 expression is both necessary and sufficient for eye development.

Insertional mutants and cloning Sine oculis

P-element based transgenes can insert in coding or regulatory sequences, which are essential for the function of an endogenous Drosophila gene. Like plant mutations caused by T plasmid or Ac-Ds insertions such mutations can be cloned relatively easily by plasmid rescue or inverse PCR (transposon tagging).

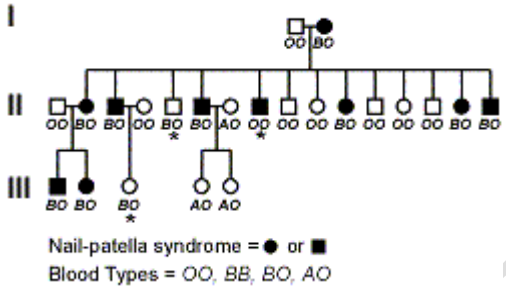
Sine oculis was originally identified as a mutation that produced an eyeless phenotype. Genetic tests showed that the original allele was hypomorphic (it did not completely knockout the function of the gene) and subsequently stronger alleles with a lethal phenotype were identified. So was mapped using overlapping deletions to the chromosome band 43C on the right arm of chromosome 2.

In 1994 a group using transgenics to look at the functions of the photoreceptor protein NinaE identified one transgenic line in which the insertion caused a recessive mutation. Flies homozygous for the insertion appeared to have relatively normal eyes with some roughness in peripheral regions. Internally they show defects in the optic lobes. In situ hybridization, using the NinaE P-element transgene to make a probe, mapped the insertion in these mutants to 43C suggesting that this insertional mutant might represent a new type of hypomorphic allele of so. The P- element insertion was used as to clone the so gene. It proved to have inserted in the 5 untranslated region of a large transcription unit coding for a homeobox protein. This protein was expressed in the embryo(in the developing optic lobes) and in larvae and adults in the eye. The original so allele only affects larval/adult expression, while the insertional mutant only affects

embryonic expression. The so homoeoprotein is homologous to the Six family of homeobox genes in mammals. Six3 knockout mice lack eyes and SIX5 mutations are associated with congenital cataracts in humans, suggesting that the role of these proteins in eye development is highly conserved.

Detection of linked loci by pedigree analysis in humans

The detection of gene action in human is difficult because of the lack of controlled crosses and the low number of progeny. The same difficulties plague the determination of linkage in humans. But some linkages can be detected in humans using pedigree analysis. Let's look at the following pedigree and see if we can determine the linkage relationships.



First, what can we determine about the inheritance of nail-patella syndrome? Clearly the disease is a dominant acting gene because affected individuals all had at least one parent with the disease. Secondly, can we determine the genotypes of the parents in generation I. The non-affected parent must be homozygous recessive for the disease. What about the affected parent in generation I. That parent must be heterozygous because offspring of this mating are homozygous recessive.

Next look at the blood type data. It is easier to genotype individuals for this phenotype because blood type is a codominant trait. Again the non-affected male is homozygous recessive whereas the affected female is heterozygous.

When we consider these two genes, the parental mating in generation I is reminiscent of what type of cross? Well the male is homozygous recessive at both genes, and the female is heterozygous at both genes. This is equivalent to a testcross. If you remember, the test cross was used to determine linkage in all of the examples that were described above. So the next step is to determine if we are seeing independent assortment of the two genes or if the two genes appear to move as a block into the gametes.

By looking at the pedigree, we can see that in almost all cases, individuals with nail-patella syndrome also posses the *B* allele. This strongly suggests that the nail-patella and blood type genes are linked, and that the dominant allele responsible for the disease is in coupling with the *B*allele at the blood type locus.

As stated above, most but not all of the offspring show this linkage. Which of the offspring are the recombinants? If we look at generation II, the second male offspring of the marriage (II-5) does not have the disease but does contain the *B* allele. Also, in this generation, the fourth male offspring (II-8) has the disease, but does not have a *B*allele. In generation III, offspring (III-3) of the marriage of the first male in generation II has the *B* allele, but is not

affected by the disease. So we can conclude that recombination has occurred between these two genes.

Now we would like to determine the linkage distance between the two genes. The original mating in genration I and the first two matings in generation II are test cross. The third mating in generation II is not informative because it involves the *A* allele which we are not following. We have a total of 16 offspring that are informative. Of these we determined that three were recombinant. As with all test crosses, this gives a genetic distance of 18.8 cM [100*(3/16)].

Somatic cell hybridization for positioning genes on chromosomes

Somatic cell hybrid genetics

Overview of the classical approach

The ability to derive long-term cultures of mammalian cells was perfected during the 1950s. Cell cultures provided important experimental material for early biochemists and molecular biologists interested in molecules and processes that occur within mammalian cells, but they were of little use to geneticists since somatic cell genomes remain essentially unchanged during continual renewal through mitotic division. This situation changed dramatically during the early 1960s when investigators discovered and developed methods for the induction of cell fusion in culture.

Normal diploid cells from all species of mammals carry approximately the same amount of DNA in their nucleus (twice the haploid amount of 3,000 mb). Thus, after fusion between any two mammalian cells, the hybrid cell nucleus becomes, in effect, tetraploid, with a genome that is twice the normal size. The enlarged genomes of hybrid cells are inherently unstable. Presumably, the increased requirement for DNA replication acts to slow down the rate of cell division, and as a consequence, cells that lose chromosomes during mitotic segregation will divide more quickly and outgrow those cells that maintain a larger genome content. Eventually, after many events of this type, cells can reach a relatively stable genome size that is close to that normally found in diploid mammalian cells. For reasons that are not understood, hybrids formed between particular combinations of species will preferentially eliminate chromosomes from just one of the parental lines. In hybrids formed between mouse cells and either hamster or human cells, mouse chromosomes will be eliminated in a relatively random manner. This process has allowed the derivation and characterization of a number of somatic cell hybrid lines that stably maintain only one or a few mouse chromosomes.

The field of somatic cell genetics had its heyday in the 1970s and early 1980s when it provided the predominant methodology for mapping loci — albeit, often to the resolution of whole chromosomes. The major tools for gene detection in this era (before the recombinant DNA revolution was in full gear) were species-specific assays for various housekeeping enzymes. Somatic cell geneticists could type each member of a panel of hybrid cells for the presence of a particular enzyme and then use karyotypic analysis to demonstrate concordance with a particular chromosome. In a strictly formal sense, this type of analysis is analogous to classical two-locus linkage studies with one marker being the enzymatic activity and the other marker being the particular chromosome that contains the gene encoding the enzyme.

The somatic cell hybrid approach has always been more important to human geneticists than to mouse geneticists. This is because well-established somatic cell hybrid lines with one or a few

mouse chromosomes are relatively rare compared to the large number of well-characterized hybrid lines with individual human chromosomes. There are several reasons for this state of affairs. First, the power of mouse linkage mapping has always been so great that somatic cell hybrid lines were never considered to be essential tools. Second, most mouse/hamster hybrid lines are chromosomally unstable and must be recharacterized each time they are grown in culture. With the difficulty of performing karyotypic analysis on mouse chromosomes, most investigators have shied away from this approach in the past. However, with alternative PCR-based methods for characterizing the chromosomal content of hybrids, this problem may have been overcome so that the derivation of new hybrids for special situations may no longer be as formidable as it once was.

The use of somatic cell hybrid panels as a general approach to gene mapping has now been superseded by *in situ* hybridization — which resolves map positions to chromosome bands rather than whole chromosomes — and, of course, classical linkage analysis. However, there are two special cases where somatic cell hybrid lines can provide unique tools for mouse geneticists. First, their DNA can be used as a source of material for the rapid derivation of panels of DNA markers to saturate particular chromosomes or subchromosomal regions. Second, their DNA can also be used for the rapid screening of new clones obtained from other sources for their presence in a particular interval of interest. This can be accomplished with the use of duplicate blots containing just three lanes of restriction digested and fractionated DNA from: (1) the somatic cell hybrid line containing the chromosome of interest; (2) mouse tissue (a positive control); and (3) the host cell line without mouse chromosomes (a negative control). Each blot can be subjected to repeated probing with different potential markers. A negative result allows one to discard a particular probe immediately; a positive result can be followed-up by higher resolution linkage analysis.

Radiation hybrid analysis

In 1990, Cox, Meyers, and their colleagues described a novel technique for determining gene order and distance which is as highly resolving as traditional linkage analysis but does not depend upon breeding. The approach used has similarities to, as well as differences from, both recombinational mapping and physical mapping. Radiation hybrid mapping was originally developed for use with the human genome, but with appropriate starting material and a sufficient number of chromosome-specific DNA markers, it can be used in the analysis of any species.

The starting material is a somatic cell hybrid line that contains only the chromosome of interest within a host background derived from another species. As indicated above, a common host species used for mouse chromosomes is the hamster. A well-established, stable hamster cell hybrid line containing a single mouse chromosome can be subjected to irradiation with X-rays that shatter each chromosome into multiple fragments. The irradiated cells are then placed together with pure hamster cells under conditions that promote fusion. Approximately 100 new hybrid clones are recovered that contain fragments of the mouse chromosome present in the original hybrid line. Finally, each of these lines are analyzed for the presence of various DNA markers that had been mapped previously into the chromosomal region of interest.

The order and distance of loci from each other can be determined according to the premise that X-rays will break the chromosome at random locations. Thus, the closer two loci are together,

the less likely it is that a break will occur between them. If two loci are side by side, they will either both be present or both be absent from all 100 cells with 100% concordance. If two loci are at opposite ends of the chromosome, there will still be cells that have neither or both, but there will also be a large number that have only one or the other. (A cell can carry both loci even if the frequency of breakage between the two is 100% since it is possible for a hybrid cell to pick-up more than one chromosomal fragment.) As the probability of chromosome breakage varies between 0% and 100% for various pairs of loci under analysis, the fraction of hybrid cells that carry both loci will vary from 100% down to a control value obtained for unlinked loci. Thus, by typing each of the "radiation hybrid cells" in the set of 100 for a series of DNA markers, it becomes possible to construct a linkage map that is highly analogous to traditional recombinational maps.

It is possible to obtain linkage maps at different levels of resolution through the use of different intensities of radiation to break chromosomes. For example, with high levels of radiation that break chromosomes once every 100 kb, on average, one could map loci from 10 kb to 500 kb; with lower levels of radiation, mapping could be performed over a window from 500 kb to 5 mb.

The analogy to classical recombination mapping is striking in that a determination of linkage distance in both cases is based on the probability with which chromosomes will break followed either by recombination (in the classical case) or by segregation upon cell fusion (in the radiation hybrid case). In both cases, linkage distances are determined by counting the ratio of offspring (pups or cells) that do or do not carry particular sets of DNA markers (alleles or genes). However, linkage distances obtained through radiation hybrid analysis are much more likely to be indicative of actual physical distances.

Although radiation hybrid analysis has provided a crucial tool for genetic analysis in humans, once again, it has not been as widely used by the mouse community because classical linkage analysis is so much more powerful. Nevertheless, the resolution of this protocol has been validated in a study of the region of mouse chromosome 2 surrounding the *agouti locus*. In this study, the radiation hybrid map that was obtained corresponded exactly with that predicted from linkage analysis, with a level of resolution that was approximately 40-fold higher. Thus, radiation hybrid mapping could serve to fill in the gap between linkage maps and physical maps, especially in "cold" regions between hotspots where distantly spaced markers cannot be separated by recombination

Notes 2:

The branch of genetics that studies cultivated somatic cells of mammals (including man), amphib ians, fish, insects, and higherplants.

The study of the genetics of somatic cells developed as an independent field in the mid-1960's, largely as a result of theapplication of many of the methods used in microorganism genet ics; these methods include the production of offspring from asingle cell in culture, the selection o f cells with a specific hereditary structure by means of selective nutrient media, thehybridization of cells, and the analysis of hybrids. The high resolving power of genetic experiments using som atic cells isdetermined by the rapid reproduction of somatic cells in a culture (the number of culti vated mammalian cells may double inonly 12–

14 hours) and the possibility of recording rare (with frequencies of 10^{-7}–10^{-8}) genetic events, including mutationsand the formation of hybrids. Such a high resolving power cannot be achieved in studies done at the level of the integralorganism.

Today, the areas most intensively studied in the genetics of somatic cells are the characteristics of the mutation process, themapping of genes in chromosomes and especially the mapping of human chromosomes, and the action and regulation ofgenes.

Natural and artificial mutagenesis in somatic-cell cultures was first studied in the early 1960's, and the ability of variousexternal factors to induce cell mutations was demonstrated in 1968. These developments contributed to the clarification ofthe molecular basis of mutagenesis, to the study of the relationship between the mutagenicity and carcinogenicity of varioussubstances and viruses, and to the evaluation of the danger posed by chemical and physical agents to human heredity.

One of the achievements of the genetics of mammalian somatic cells has been the production of hybrid cells by the fusion oftwo or more heterogeneous cells. These hybrid cells may combine the genomes of taxonomically distant species, forexample, of man and various rodent species, mice and chickens, and even man and mosquitoes. Special techniques toobtain hybrid cells include the treatment of cells with inactivated Sendai virus to increase the likelihood of their fusion and the use of selective mediums on which parent cells die and only hybrids survive and form colonies (Figure 1).

Two species of somatic hybrids obtained from human and mouse cells and from human and Chinese-hamster cells are usedto localize human genes on chromosomes. The principle underlying the mapping is the fact that human chromosomesdisappear in the course of the reproduction of hybrid cells. This disappearance is random, and therefore different humanchromosomes remain in each hybrid clone. In hybrids, comparison of the peculiarities of human cells with the survivinghuman chromosomes makes it possible to establish which chromosome contains the gene determining the particularcharacteristic. In the 1960's and 1970's enough genes were localized to permit the marking of almost all humanchromosomes.

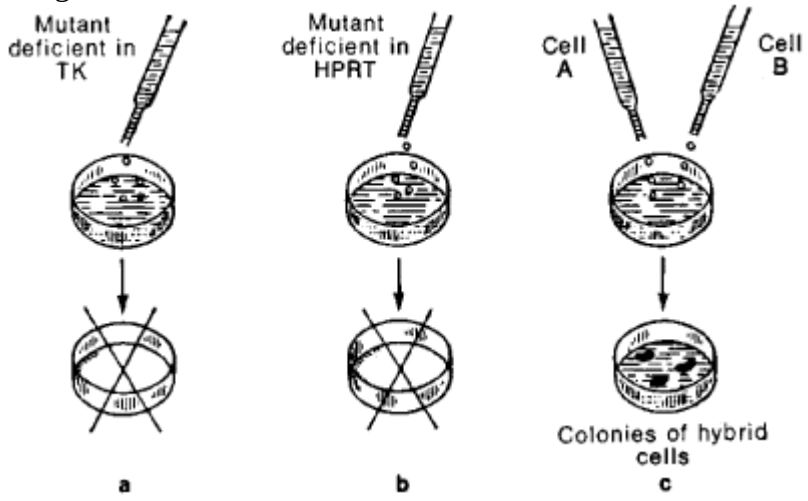


Figure 1. A method of isolating somatic hybrids using a selective medium on whichthe only cell s that can be cultivated are those capable of producing the enzymesthymidine kinase (TK) and hy poxan-thine-
guamne phosphoribosyl transferase(HPRT). If as a result of mutation the cells of one parent do n ot produce oneenzyme and the cells of the other parent do not produce the other enzyme, bothspe cies of parental cells will die and only the hybrid cells that produce both enzymeswill survive an d reproduce.

Mammalian cells are also hybridized to study the action of genes. Specifically, it has been shown that mammalian genomescontain regulatory genes that control the action of structural genes. Thus, the study of the genetics of mammalian somaticcells has proved to be a fruitful approach both in the analysis of theoretical questions and in the resolution of many practicalmatters, including the diagnosis of hereditary diseases, the evaluation of the harmful effect of individual environmental factorson genes, and the determination of the causes of malignant cellular degeneration.Plant cells are commonly used in genetic research because a large number of cells having the same genotype may beobtained from a single plant. The development of methods use d to isolate protoplasts, that is, the elimination of a plant cell'ssolid membranes, has facilitated the hybridization of somatic cells, genetic transformation, and the study of other genetic processes. The selective breeding of plants has also been initiated. In some cases a plant's ability to regenerate permits thegrowing of whole plants from a single cell. Today, the genetics of the so matic cells of plants and mammals offers manychallenges from both the theoretical and practical standpoints.

Pedigree conventions

Pedigree Analysis

A pedigree chart displays a family tree, and shows the members of the family who are affected by a genetic trait.

This chart shows four generations of a family with four individuals who are affected by a form of colorblindness.

• Circles represent females and squares represent males.

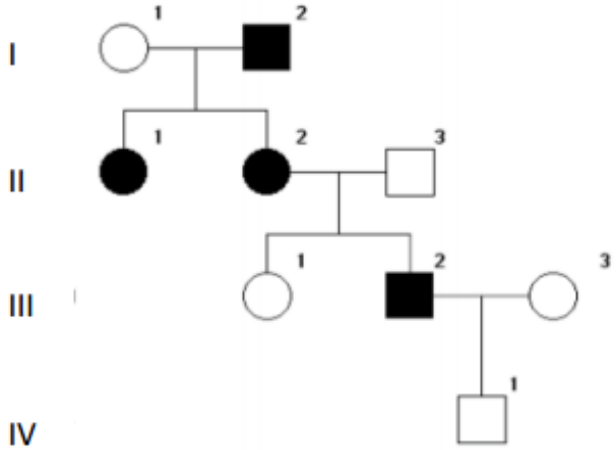
• Each individual is represented by:

– a Roman Numeral, which stands for the generation in the family,

– a Digit, which stands for the individual within the generation. (For instance, The female at the upper left is individual I-1.)

• A darkened circle or square represents an individual affected by the trait.

• The "founding parents" in this family are the female I-1 and



the male I-2 in the first
generation at the top.
• A male and female directly
connected by a horizontal line
have mated and have children. These three pairs have mated in this tree:
I-1 & I-2, II-2 & II-3, III-2 & III-3
• Vertical lines connect parents to
their children. For instance the
females , II-1 and II-2 are daughters of I-1 and II-2
• The "founding family" consists of the two founding parents and their children, II-1 and II-2.
In this pedigree, the unaffected founding mother, I-1, and affected founding father, I-2, are parents to two affected daughters, II-1 and II-2.
The affected founding daughter II-2 and the unaffected male II-3 who "marries into the family" have two offspring, an unaffected daughter III-1 and affected son, III-2.
Finally, this affected male III-2 and the unaffected female III-3 who "marries in" have an unaffected son, IV-1.
Pedigrees are interesting because they can be used to do some detective work and are often used to study the genetics of inherited diseases. For example, pedigrees can be analyzed to determine the mode of transmission for a genetic disease:
(1) Dominance - whether the disease alleles are dominant or recessive;
(2) Linkage - whether the disease alleles are X-linked (on the X chromsome) or autosomal
Autosomal chromosomes - The 22 chromosome pairs other than the XX (female) or XY (male) sex chromosomes.
Hemizygous - Males are "hemizygous" for X-linked genes – males only have one X chromosome and one allele of
any X-linked gene.
Allele - A version of a gene. Humans have 2 alleles of all their autosomal genes; females have 2 alleles of Xlinked genes; males have one allele of X-linked genes (and one allele of Y-linked genes).
Pedigree analysis is an example of abductive reasoning. In pedigree analysis you need to look for any clues that will allow you to decide if the trait is dominant or recessive and whether it is linked to an autosomal chromosome, or to the X chromosome.

General Assumptions

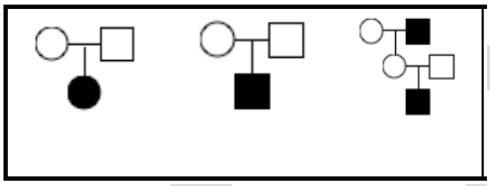
In the problems that follow, you'll be reasoning about the mode of transmission of genetic traits that are controlled by one gene, with two alleles, a dominant allele and a recessive allele.
We also make three simplifying assumptions:
1. Complete Penetrance. An individual in the pedigree will be affected (express the phenotype associated with a trait) when the individual carries at least one dominant allele of a dominant trait, or two recessive alleles of a recessive a trait.
2. Rare-in-Population. In each problem, the trait in question is rare in the general population. Assume for the
purposes of these problems that individuals who marry into the pedigree in the second and third

generations are not carriers. This does not apply to the founding parents – either or both of the individuals at
the top of the pedigree could be carriers.
3. Not-Y-Linked. The causative genes in these problems may be autosomal or X-Linked, but are not Y-linked.
5 Key Clues
There are five things to remember in reasoning about pedigrees.
(1) An unaffected individual cannot have any alleles of a dominant trait. (because a single allele of a dominant trait causes an individual to be affected).
(2) Individuals marrying into the family are assumed to have no disease alleles they will never be affected and can never be carriers of a recessive trait.(because the trait is rare in the population)
(3) An unaffected individual can be a carrier (have one allele) of a recessive trait.
(because two alleles of a recessive trait are required for an individual to be affected)
(4) When a trait is X-linked, a single recessive allele is sufficient for a male to be affected.
(because the male is hemizygous – he only has one allele of an X-linked trait)
(5) A father transmits his allele of X-linked genes to his daughters, but not his sons.
A mother transmits an allele of X-linked genes to both her daughters and her sons.

Characteristics of dominant and recessive inheritance

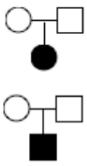
Key Patterns in Pedigree Analysis

Patterns that Indicate a RECESSIVE Trait



• The disease must be RECESSIVE if any affected individual has
2 unaffected parents.
Since this is a genetic disease at least one parent must have
an allele for the disease.
If neither parent is affected, the trait cannot be dominant.
(See Clue 1 above).

Patterns that Indicate a RECESSIVE Trait



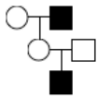
• AUTOSOMAL RECESSIVE: If any affected founding daughter has 2 unaffected parents the disease must be autosomal recessive.
An affected individual must inherit a recessive allele from both parents, so both parents must have an allele.
If the father had a recessive X-linked allele, he would have to be affected (since he only has

one X-linked allele).

• RECESSIVE: If an affected founding son has 2 unaffected parents, we cannot determine if the recessive disease is autosomal or x-linked.

If the trait is autosomal, both parents can be unaffected carriers of the disease.

If the trait is x-linked, the son must have inherited his allele from his mother only, and his father can be unaffected.

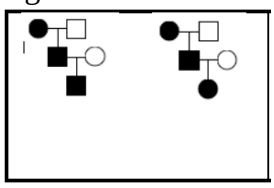


• X-LINKED RECESSIVE: When an affected non-founding son has 2 unaffected parents the disease must be X-linked recessive.

The father, who is marrying in, does not have any disease alleles, since he is marrying into the family;

so the affected son inherits an allele only from his unaffected mother.

A male cannot be affected by a single autosomal recessive allele, but can be affected by a single X-linked recessive allele.



Patterns that Indicate a DOMINANT Trait

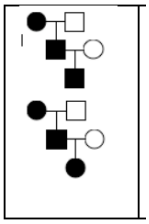
• The disease must be DOMINANT if every affected child of NON-FOUNDING parents has an affected parent.

The unaffected mother, who is marrying in, does not carry an allele for the disease;

so the affected child inherits an allele only from the affected father.

No child could be affected by a single autosomal recessive allele, or X-linked recessive allele, so the trait is dominant.

• When an affected son of non-founding parents has an affected father the disease must be



AUTOSOMAL DOMINANT.

A father does not transmit X-linked alleles to a son, so the disease cannot be X-linked dominant.

• When an affected daughter of non-founding parents has an affected father, we cannot determine whether the DOMINANT disease is autosomal or x-linked.

The affected father can transmit either an autosomal dominant allele, or an X-linked dominant allele to his daughter.

Applications of pedigree analysis

Pedigrees can be used in the clinical setting, such as genetic counseling sessions or genetic evaluations, or in genetic research. By analyzing how many family members have a genetic disorder, how these individuals are related, and the sex of the affected individuals, it is often possible to determine the inheritance pattern of the genetic disorder in the family. Together, the inheritance pattern and an accurate diagnosis help the genetic professional provide accurate risk information to the family. This includes risk information for future pregnancies or relatives who are currently unaffected, but who are at risk for developing the disorder based on family history information. Genetic testing options, if available, can then be offered to those at risk.

The pedigree is also a standard tool used by researchers. For example, in studies aimed at identifying genes that cause human genetic disorders, researchers must collect detailed information on relatives participating in the study, particularly those relatives who are affected with the disorder. Researchers compare the genes of affected individuals with the genes of those who did not inherit the disorder to identify the specific genes responsible. In other studies the disease-causing gene is known, and researchers study the gene mutation (s). A pedigree can help identify which family members should be included in mutation analysis, as only those family members who are affected or are at-risk could carry a mutation. Researchers can also pictorially show laboratory data, such as genotypes or haplotypes, on the pedigree.

CLASS: II B.Sc., BIOCHEMISTRY **COURSE NAME: CONCEPTS IN GENETICS**
COURSE CODE: 17BCU304B **UNIT:III-GENETICS OF EUKARYOTES AND HUMAN PEDIGREE ANALYSIS**
(BATCH-2017-2020)

POSSIBLE QUESTIONS

2 marks

1. What is linkage?
2. Explain crossing over.
3. What is pedigree and how it's employed in genetics?
4. What is centromere?
5. What are the methods involved in eukaryotic gene mapping?

8 Marks

1. What is linkage? Why is crossing over so important for biological diversity on Earth?
2. Comment on pedigree conventions with suitable examples.
3. Elaborate on centromere mapping.
4. Give detailed account on pedigree analysis.
5. What do you understand on eukaryotic genetic mapping? Explain in detail.
6. Comment on characteristics of dominant and recessive inheritance.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
II-B.Sc., BIOCHEMISTRY
CONCEPTS IN GENETICS (17BCU304B)
MULTIPLE CHOICE QUESTIONS
UNIT III

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	What was unique about the chromosomes studied by Creighton and McClintock?	They carried different alleles for two linked genes.	They were unable to cross over.	These chromosomes were unable to synapse during meiosis I.	Physical structures of the chromosomes allowed them to be distinguished from each other.	**Physical structures of the chromosomes allowed them to be distinguished from each other.**
2	Which of the following would be a reasonable way to demonstrate that twin spots are caused by mitotic recombination?	Collect a large number of flies with twin spots and verify that in all cases, both halves of the twin spot are the same size.	Collect a large number of petunias with twin spots and verify that they are able to pass this phenotype to their progeny.	Design cytologically distinct chromosomes that must recombine to cause the twin spot; analyze chromosomes from the spots with a microscope.	Try to clone new individuals from the cells in the twin spots and analyze their phenotype.	**Design cytologically distinct chromosomes that must recombine to cause the twin spot; analyze chromosomes from the spots with a microscope.**
3	You begin to study a novel plant species and discover that this diploid plant has 16 chromosomes. How many linkage groups would you expect to find?	16	8	4	It is impossible to tell from the information provided.	**8**
4	You calculate map distances between genes A, B, and C based on all pairwise dihybrid crosses. When you perform the trihybrid cross to verify your results, you discover that, despite analysis of a very large number of progeny, you have only about half as many double crossover progeny as you expect. Which explanation is the most reasonable?	Your dihybrid cross data are an underestimate of the distance between the more widely separated genes.	You are seeing an example of interference.	You are seeing an example of random sampling error.	None of these.	**You are seeing an example of interference.**
5	What is a convenient way to identify gene order in a trihybrid mapping cross?	Look for double crossover progeny and identify the gene that was flipped relative to parentals.	Look for the largest class of single crossovers and the two genes are the ones on the ends.	Look for the smallest class of single crossovers to identify the genes which are closest together.	A and B only.	**A and B only.**

6	Why is tetrad analysis useful for studying genetics?	Cells within a tetrad are haploid so genotype is directly reflected in the phenotype.	Cells within a tetrad reflect the events of meiosis.	Analysis of simple systems can elucidate general rules that apply more broadly. .	All of the above	**All of the above**
7	Tetrad analysis automatically includes the centromere as an additional point in mapping crosses, because	crossovers that are not between a gene and the centromere do not affect the relationship of that gene to the chromosome during division.	crossovers that are between a gene and the centromere cannot be identified in this system.	when studying crossovers between genes, the position of the centromere is irrelevant.	All of the above.	**crossovers that are not between a gene and the centromere do not affect the relationship of that gene to the chromosome during division.**
8	In an unordered tetrad with the parental ditype:	all spores have the same genotype.	all spores have the genotype of one of the parental strains.	all four spores have different genotypes.	all spores are diploiD. In a parental ditype, the spores are of both parental types in a 2:2 ratio.	**all spores have the genotype of one of the parental strains.**
9	You set up a mapping cross involving two genes in Aspergillus and determine the following percentages of tetrad types: parental ditype 51%, nonparental ditype 49%. What can you say about your genes?	They are closely linked.	They are unlinked.	They are lethal.	They are not expressed in the phenotype.	**They are unlinked.**
10	You set up a mapping cross involving two genes of interest in Saccharomyces and determine the following tetrad types: 8 nonparental ditype; 37 tetratype; 80 parental ditype. Calculate the most accurate map distance between these two genes.	18.7mu	32.3mu	64.6mu	None of these is accurate.	**32.3mu**
11	Formation of the following pattern in an ordered octad is due to what event during meiosis? Pattern: AAaaaaaAA	No crossing over and normal segregation in meiosis I.	Crossing over and normal segregation during meiosis I.	Crossing over and segregation during meiosis II.	No crossing over and segregation during meiosis II.	**Crossing over and segregation during meiosis II.**
12	You set up the following dihybrid mapping cross in fruit flies. P e+e+ roro cq cq x ee ro+ro+ cq+cq+. After backcrossing F1 males to ee roro cq cq females, you get the following results: gray body, rough eyes, claret eyes 576; ebony body, smooth, red eyes 564. How can you explain this result?	The genes are not linked.	Crossing over does not occur in male Drosophila.	These genes are X linked.	A and B only.	**Crossing over does not occur in male Drosophila.**
13	You set up a mapping cross involving your favorite gene (YFG) in Neurospora and determine the following octad types: 4:4 arrangement 83; 2:4:2 arrangement 7; 2:2:2:2 arrangement 10. What is the distance between your gene and the centromere?	7 cm	10 cm	17 cm	83 cm	**17 cm**

14	Crossing over has which of the following outcomes?	Formation of chiasmata in early prophase	Recombination of genes on the same side of the crossover point	multiple alleles for a given gene in a single chromatid	none of these	**none of these**
15	A heterozygous Neurospora cell is useful for mapping a chromosome's centromere because its tetrads are ordered and its alleles segregate in how many ways in meiosis?	only one way	two ways	four ways	none of these	**two ways**
16	Linkage is maize was reported by	Bateson	Morgan	Hutchinson	Sinnot	**Hutchinson**
17	Centromere is a part of	Chromosome	endoplasmic reticulum	ribosomes	mitochondria	**chromosome**
18	The inhibition of a crossover by another is known as:	**interference**	coefficient of coincidence	a marker	translocation	**interference**
19	Human pedigree analysis can provide which of the following pieces of information: 1. linkage between human genes 2. estimates of recombination frequencies 3. inheritance of chromosomal markers	1	2	3	all of these	**3**
20	The point, at which polytene chromosome appear to be attached together, is called	centromere	chromomere	chromocentre	centriole	**chromocentre**
21	Initiation codon in eukaryotes is	GAU	AGU	**AUG**	UAG.	**AUG**
22	Crossing Over occurs when the homologous chromosomes contain	One chromatid.	two chromatid	three chromatid	four chromatid	**two chromatid**
23	The eukaryotic genome differs from the prokaryotic genome because	the DNA is complexed with histone in prokaryotes	the DNA is circular and single stranded in prokaryotes	repetitive sequences are present in eukaryotes	genes in the former case are organized into operons.	**the DNA is circular and single stranded**

						in prokaryotes
24	Anticodon is an unpaired triplet of bases in an exposed position of	<-RNA.	w-RNA	r-RNA	both ‘b’ and ‘c’	**<-RNA.**
25	Coupling and repulsion are two aspects of same phenomenon called	Linkage	Crossing over	Evolution	Variation	**Linkage**
26	Linkage usually gets broken due to	Mutation	Cross over	Epistasis	Variation	**Cross over**
27	to linkage in maize, the number of colourless shrunken seeds produced are	4032	4035	149	152	**4035**
28	What pattern of inheritance is shown in the pedigree?	Autosomal dominant	Autosomal recessive	Sex linked dominant	Sex linked recessive	**Autosomal recessive**
29	Which of these is an example of cytogenetic mapping?	Hybridization studies in Drosophila determined the location of the yellow gene, which influences body color, to be near the tip of the X chromosome.	Karyotype analysis of humans with familial Down syndrome has identified a relatively small region of chromosome 21 containing genes that cause the Down phenotype.	Analysis of the chromosomes of a transgenic strain of mice indicates that the transgene has been inserted into the short arm of chromosome 15.	All of these.	**All of these.**
30	Cytologists can use which of the following to describe locations of a gene at a specific place on the chromosome?	DNA sequence of a chromosomal region.	Protein expression from a chromosomal region.	Banding pattern of stained chromosomes.	Comparison to markers located within a few thousand bp.	**Banding pattern of stained chromosomes.**

31	Cytogenetic mapping:	Requires that the gene have been cloned if in situ hybridization is to be used.	Is a high resolution method of gene localization.	Can be used to determine relative order of genes located very close to each other.	All of these.	**Requires that the gene have been cloned if in situ hybridization is to be used.**
32	Which of the following questions could be easily answered by FISH?	How often does recombination occur between the ebony and sepia loci in Drosophila?	In a particular family affected with Becker muscular dystrophy, is the mutation due to a deletion of the entire gene?	How large a piece of DNA would you need to clone in order to clone the group of human globin genes from chromosome 11?	All of these.	**How large a piece of DNA would you need to clone in order to clone the group of human globin genes from chromosome 11?**
33	Linkage mapping can determine the distance between which of the following pairs of DNA sequences?	AFLPs and RFLPs.	Two AFLPs.	Two known genes.	A known gene and any type of molecular marker.	**Two AFLPs.**
34	Somatic hybridization is achieved through	Grafting	Protoplast fusion	Conjugation	Recombinant DNA technology	**Protoplast fusion**
35	Alfred Sturtevent developed genetic maps known as	linkage maps	genetic localization	genetic expression	genetic path	**linkage maps**
36	Tendency of alleles that are located close together on chromosome to be inherited together during meiosis is	genetic linkage	genetic code	inheritance	gene expression	**genetic linkage**

37	Recombinant frequency of 1% is equivalent to	10 m.u	20 m.u	1 m.u	5 m.u	**1 m.u**
38	Linked genes:	are located near each other on the same chromosome.	violate the law of independent assortment.	segregate together during meiosis.	All of the above	**All of the above**
39	From the results in the previous question, calculate a map distance between these two genes.	0.15 map units	0.85 map units	1.5 map units	15 map units	**0.15 map units**
40	How did Morgan demonstrate crossing over?	A fly that received only y+w+m+ and ywm chromosomes from her parents was able to transmit different combinations of alleles to her offspring.	A fly with a gray body and red eyes (wild type) was able to produce offspring with gray bodies and white eyes.	True breeding yellow, white, miniature flies always had yellow, white, miniature offspring.	None of these examples demonstrate crossing over.	**A fly that received only y+w+m+ and ywm chromosomes from her parents was able to transmit different combinations of alleles to her offspring.**

UNIT 4
SYLLABUS

Developmental genetics, epigenetics and chromosomal aberrations

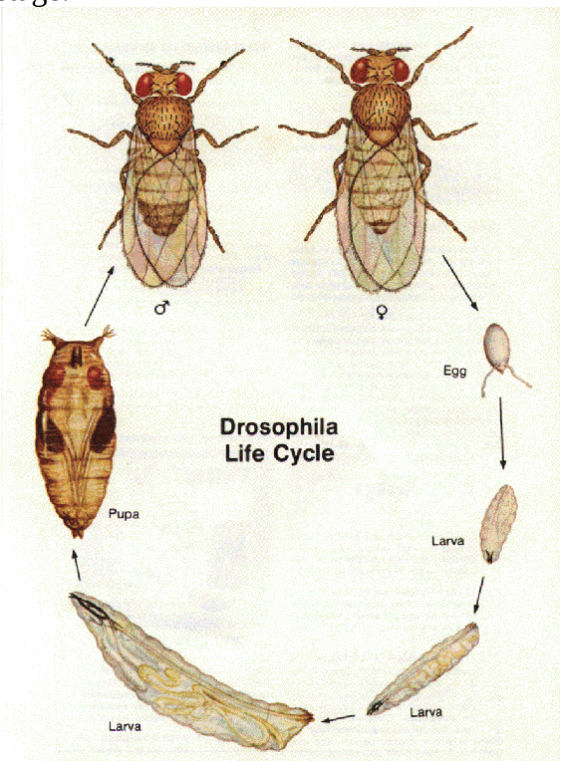
Model organism for genetic analysis, *Drosophila* development, maternal effect genes, morphogens and zygotic gene activity in development, sex chromosomes and sex determination, dosage compensation of X-linked genes. Extra nuclear inheritance, tests for organelle heredity and maternal effect, epigenetic mechanisms of transcriptional regulation & genomic imprinting. Variations in chromosome number- monosomy and trisomy of sex and autosomes. Variations in chromosome structure- inversions, deletions, duplications and translocations.

Model organism for genetic analysis

A model system is a simpler, idealized system that can be accessible and easily manipulated. Therefore, when selecting living organisms as models to work with, certain criteria are used depending upon the experimental purposes. As a result, there is a wide range of characteristics common to model organisms, including: 1) rapid development with short life cycles, 2) small adult size, 3) ready availability, and 4) tractability. Being small, growing rapidly and being readily available are crucial in terms of housing them, given the budget and space limitations of research and teaching laboratories. Tractability relates to the ease with which they can be manipulated.

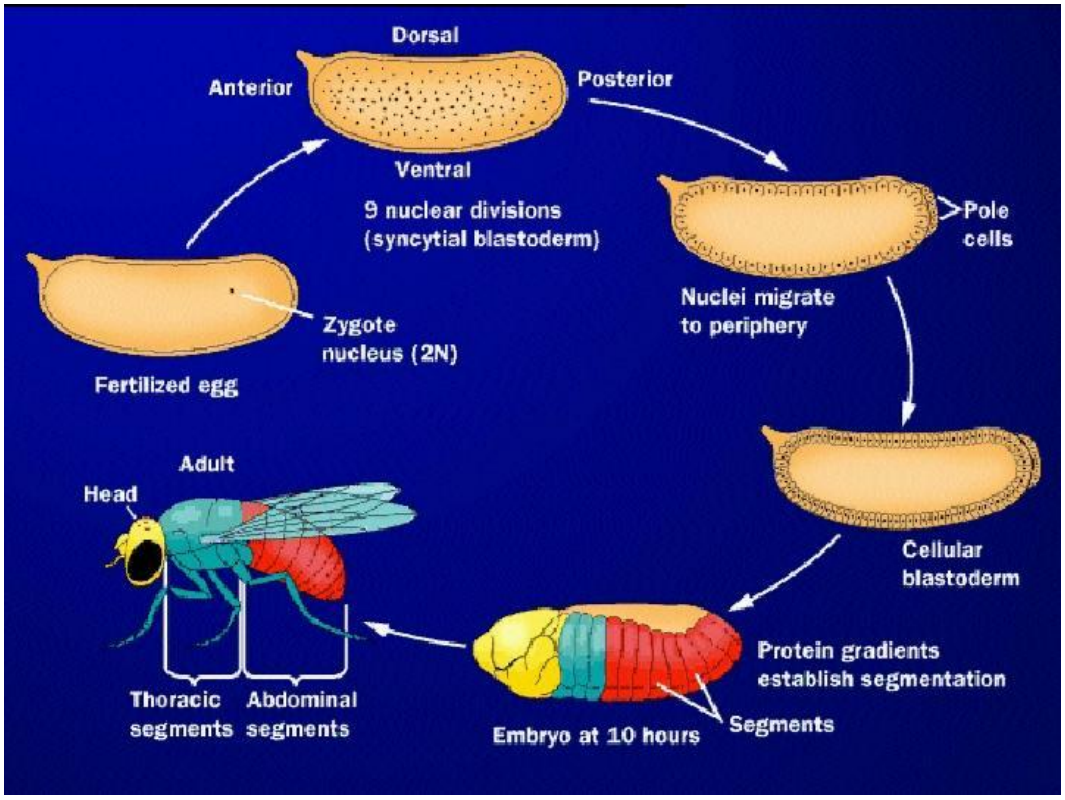
Drosophila development

The ***Drosophila* life cycle** consists of a number of stages: embryogenesis, three larval stages, a pupal stage, and (finally) the adult stage.



Embryogenesis in *Drosophila*

Following fertilization, mitosis (nuclear division) begins. However, cytokinesis (division of the cytoplasm) does not occur in the early *Drosophila* embryo, resulting in a multinucleate cell called a **syncytium**, or **syncytial blastoderm**. The common cytoplasm allows morphogen gradients to play a key role in pattern formation. At the tenth nuclear division, the nuclei migrate to the periphery of the embryo. At the thirteenth division, the 6000 or so nuclei are partitioned into separate cells. This stage is the **cellular blastoderm**. Although not yet evident, the major body axes and segment boundaries are determined. Subsequent development results in an embryo with morphologically distinct segments.

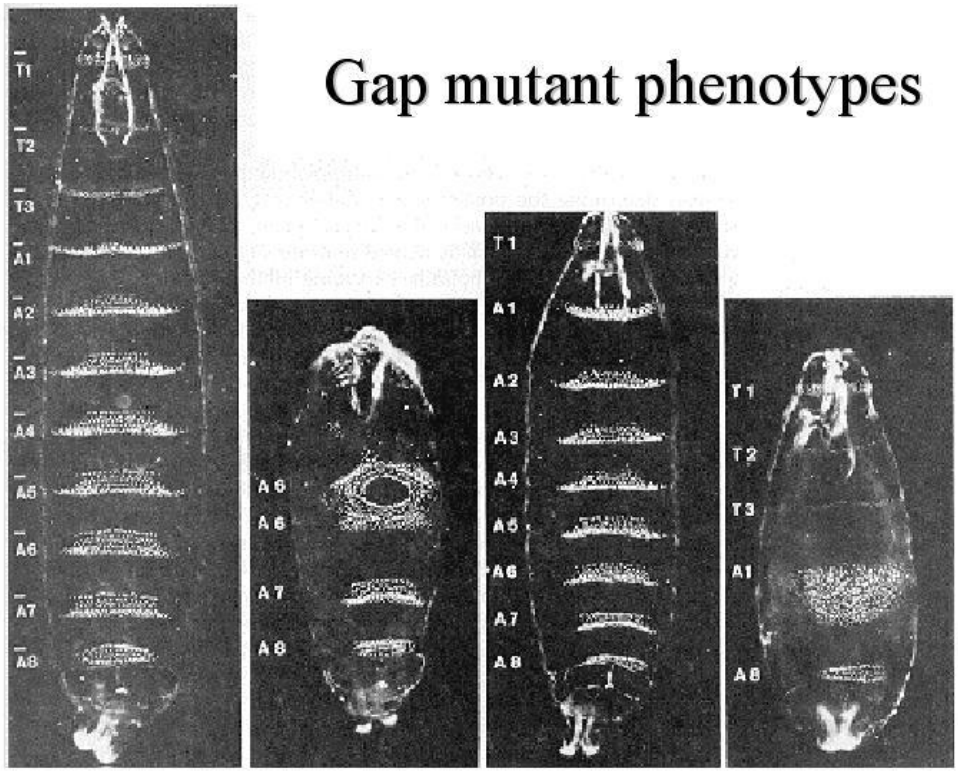


Genetic Analysis of *Drosophila* development

Much of what we understand about *Drosophila* development is based on the isolation and characterization of developmental mutants by three scientists, Ed Lewis, Christiane Nusslein-Volhard, and Eric Wieschaus, who were awarded the Nobel prize for their work in 1995. Lewis did pioneering research on late embryogenesis, while Nusslein-Volhard and Wieschaus concentrated their efforts on understanding early embryogenesis.

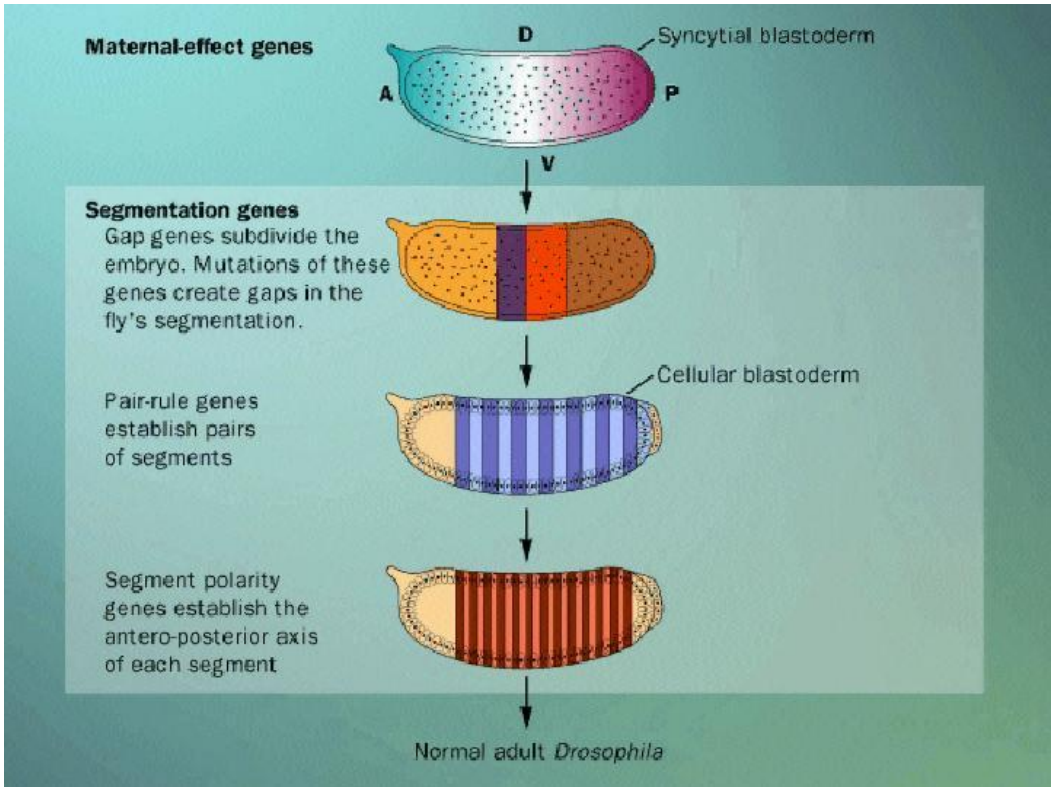
Nusslein-Volhard and Wieschaus set out to identify every gene required for early pattern formation in the Drosophila embryo. They looked for recessive embryonic lethal mutations, and classified them according to their phenotype before death. That is, they looked for and analyzed dead embryos.

Images of some of the mutants they identified are shown below. Notice the differences in segmentation patterns between the wildtype, shown on the left, and the mutant embryos.



A cascade of gene activation sets up the *Drosophila* body plan

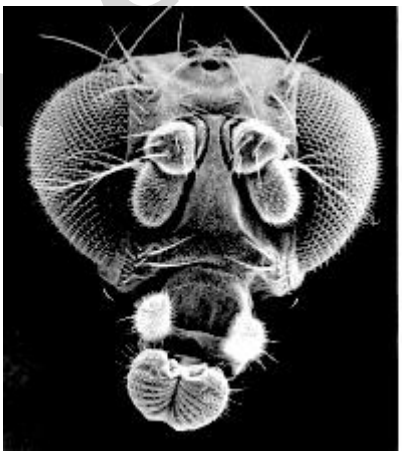
The maternal-effect genes, including bicoid and nanos, are required during oogenesis. The transcripts or protein products of these genes are found in the egg at fertilization, and form morphogen gradients. The maternal-effect genes encode transcription factors that regulate the expression of the gap genes. The gap genes roughly subdivide the embryo along the anterior/posterior axis. The gap genes encode transcription factors that regulate the expression of the pair-rule genes. The pair-rule genes divide the embryo into pairs of segments. The pair-rule genes encode transcription factors that regulate the expression of the segment polarity genes. The segment polarity genes set the anterior/posterior axis of each segment. The gap genes, pair-rule genes, and segment polarity genes are together called the segmentation genes, because they are involved in segment patterning.

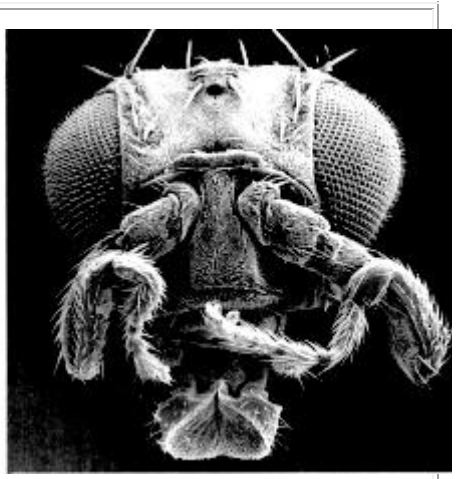


But how do these segments take of individual identities?

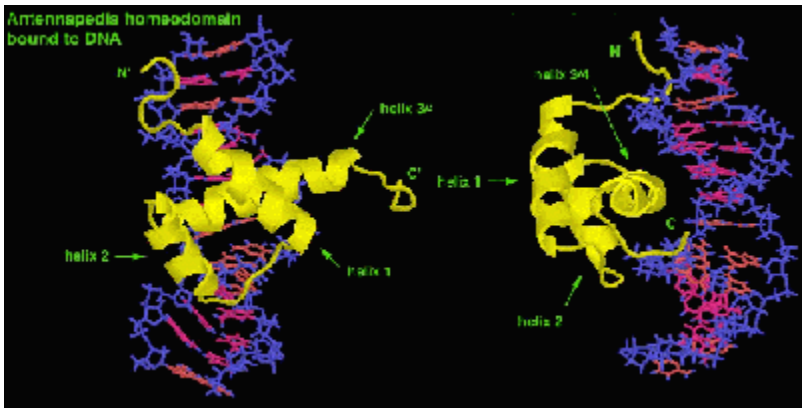
In normal flies, structures like legs, wings, and antennae develop on particular segments, and this process requires the action of **homeotic genes**. Enter Ed Lewis, who discovered homeotic mutants - mutant flies in which structures characteristic of one part of the embryo are found at some other location.

Homeotic mutations, such as *Antennapedia*, cause a misplacement of structures. These two scanning electron micrographs show fly heads. On the left is a wildtype fly. On the right is a fly with the dominant ***Antennapedia***mutation - and legs should be the antennae

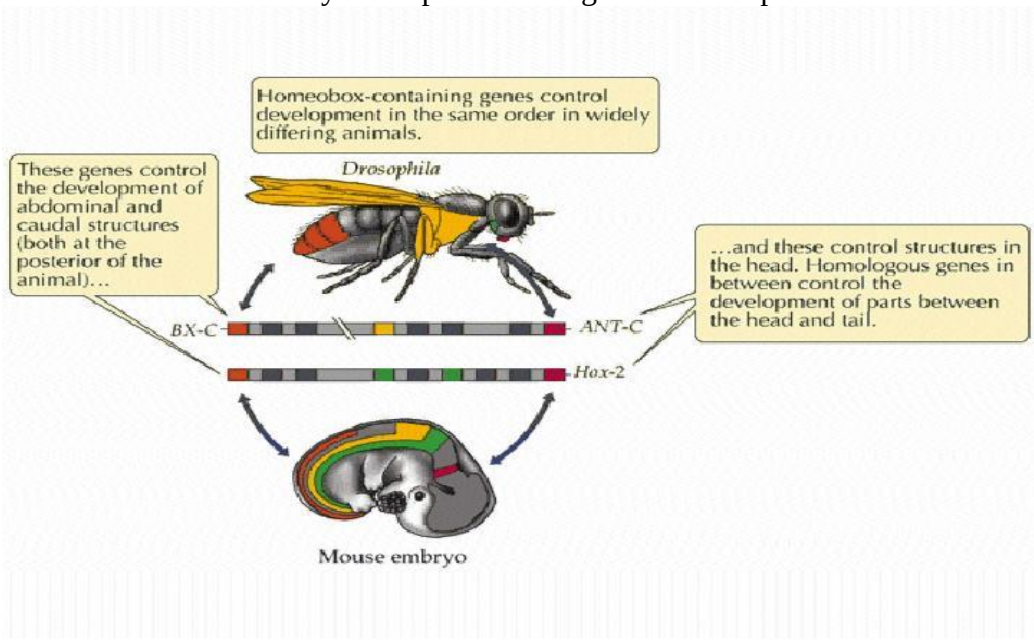




The homeotic genes encode transcription factors that control the expression of genes responsible for particular anatomical structures, such as wings, legs, and antennae. The homeotic genes include a 180 nucleotide sequence called the homeobox, which is translated into a 60 amino acid domain, called the homeodomain. The homeodomain is involved in DNA binding, as shown in the images below.

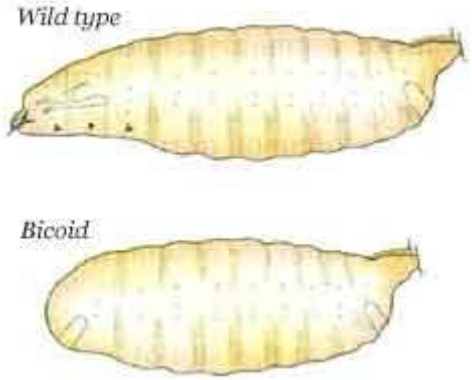


Homeobox-containing (or HOX) genes are found in many organisms, including worms, fish, frogs, birds, mammals, and plants. Interestingly, HOX genes are found in **clusters**, and the relative **gene order** within these clusters in conserved between organisms. That is, the order of related HOX genes in Drosophila and in mice is the same! In addition, the order of HOX-genes on the chromosome is related to where they are expressed along the anterior/posterior axis.



Maternal effect genes

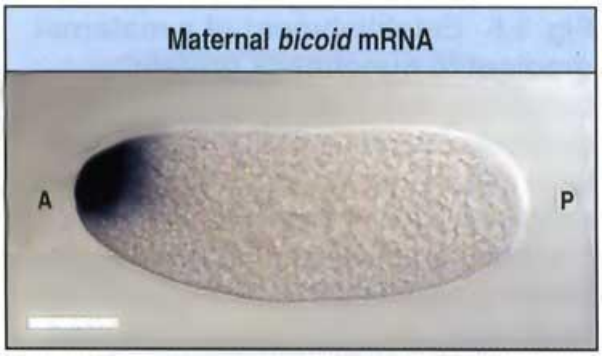
Maternal effect genes are a special class of genes that have their effect in the reproductive organs of the mutant; they are interesting because the mutant organism may appear phenotypically normal, and it is the *progeny* that express detectable differences, and they do so whether the progeny have inherited the mutant gene or not. That sounds a little confusing, but it really isn't that complex. I'll explain it using one canonical example of a maternal effect gene, *bicoid.*



Bicoid is a gene that is essential for normal axis formation in the fly, *Drosophila*. It is this gene product that basically tells the fly embryo which end is the front end—the cartoon to the right illustrates what mutant larvae look like. The top picture is a normal, or wild type, *Drosophila* larva. Students of the fly will recognize that this animal is facing to the left by the presence of the dark mouthparts; fly experts will be familiar with the stippling in the figure, which illustrates characteristic locations of bristles on the animal's cuticle.

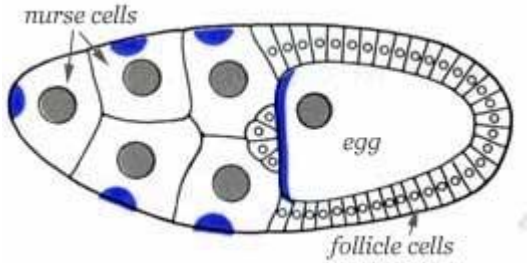
The picture below it is of an animal that lacks the *bicoid* gene product. It has no mouthparts, no head end at all; looking at the pattern of bristles, one can also see that the front end has the bristles found on the back end. (Yes, South Park fans, geneticists have created flies with two asses.)

The tricky part here is that that fly expressing the *bicoid* mutant phenotype may **not** carry the mutant gene. It could be genetically normal. What we know, though, from looking at it is that the poor two-assed fly's *mother* was a mutant. We know that because *Drosophila* embryos do not synthesize the *bicoid* gene product at all, not even the wildtype flies, and they all inherit it directly from Mom. The only way they could be lacking it is if their mother failed to pack it into the egg.



Here, for instance, is an *in situ* stain of a freshly laid wildtype *Drosophila* egg. An *in situ* stain is a way to dye specific RNA sequences, and in this case the egg is blue where ever *bicoid* RNA is present…and as can be seen, it is localized specifically to the front end of the egg. A photograph of a similarly stained egg from a *bicoid* mutant mother would look similar, except that there would be no blue spot at all—the egg would be a uniform gray. The key thing to understand is that that blue spot

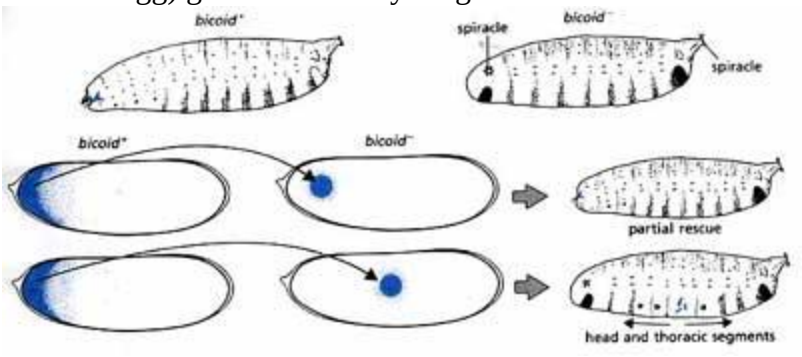
was not put there by the activity of the *egg's* genes, but exclusively by the action of the *mother's* genes. The diagram below illustrates how this pattern is set up during the formation of the egg in the mother fly's ovaries.



This is a follicle extracted from the ovary of a fly. It consists of several cell types, all of which will eventually be discarded except for one, the egg proper. The egg is going to grow to a relatively immense size, and it needs help to do that. An important contributor to that growth is a set of cells called nurse cells—the nurse cells are busy synthesizing essential proteins, such as yolk proteins, and stuffing them into the egg. They also make *bicoid* RNA (the blue stuff), which is similarly stuffed into the egg, along with other accessory proteins that make it sticky so that the *bicoid* RNA stays at that one end. There are also numerous cells called follicle cells that secrete the chorion, or shell that will surround the egg.

The diagram is only illustrating *bicoid*, but there are *many* RNAs and proteins that are being pumped into and secreted onto the surface of the developing egg. There are maternal genes that are necessary for the posterior end, and others that define dorsal and ventral sides, for instance.

Another important fact that isn't illustrated here is that the nurse and follicle cells are the *mother's*cells and have the mother's genotype. The egg, once it is fertilized and laid, is going to have a different genotype, and may actually acquire wildtype genes with wildtype *bicoid*…but it won't matter. If the maternal genes are defective, the damage is done *before* zygotic (from the fertilized egg) genes can do anything.



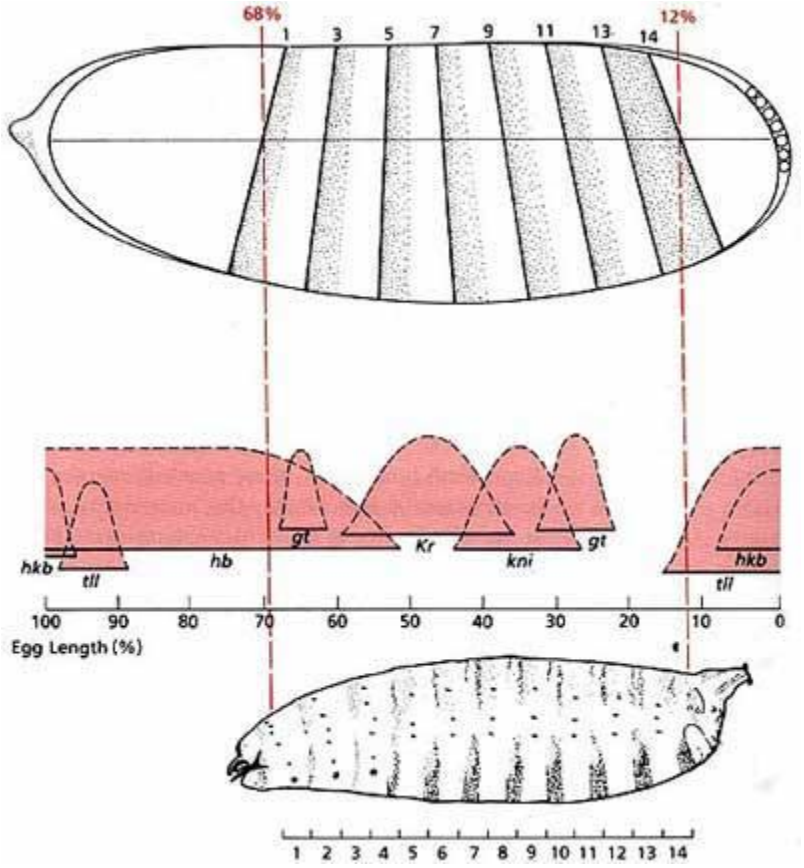
One clever experiment: the role of the bicoid gene product has been tested with what is called a rescue experiment, illustrated above. At left on the top is a *bicoid*$^{+}$, or wildtype larva, and on the right is a *bicoid*$^{-}$ larva that lacks any *bicoid* gene product. What if we injected it with *bicoid*? In the experiment, a little bit of *bicoid*$^{+}$ cytoplasm is sucked out of the anterior end of a normal egg, and injected into the anterior end of an egg deficient for *bicoid*. That's enough to do a partial rescue; it's hard to get a perfect rescue, because dosage and localization are impossible to get exactly as they are in the intact egg.

The experiment at the bottom illustrates another interesting result: if the *bicoid*$^{+}$ cytoplasm is injected into the middle of the egg, the embryo tries to form a head in the middle of it's body, as indicated by the difficult-to-see jumble of mouthparts that form there.

Maternal effect genes are common, and we know they are present in humans and other mammals—eggs contain many more informational macromolecules than just strands of DNA, and any organism above the level of a virus is going to pass information on to its progeny via the cytoplasm. However, maternal effect genes are most important in the very earliest stages of embryonic development, and defects in them are likely to be lethal. Maternal effect mutants in mammals aren't going to be seen as weird looking embryos, but as infertility problems, since embryos that fail in the first few days or weeks will simply be spontaneously aborted. There are few specifically identified maternal effect genes in mammals; one example is STELLA, identified in mice, in which homozygous carriers of the mutant allele look normal, but have severely reduced fertility.

Zygotic gene activity in development

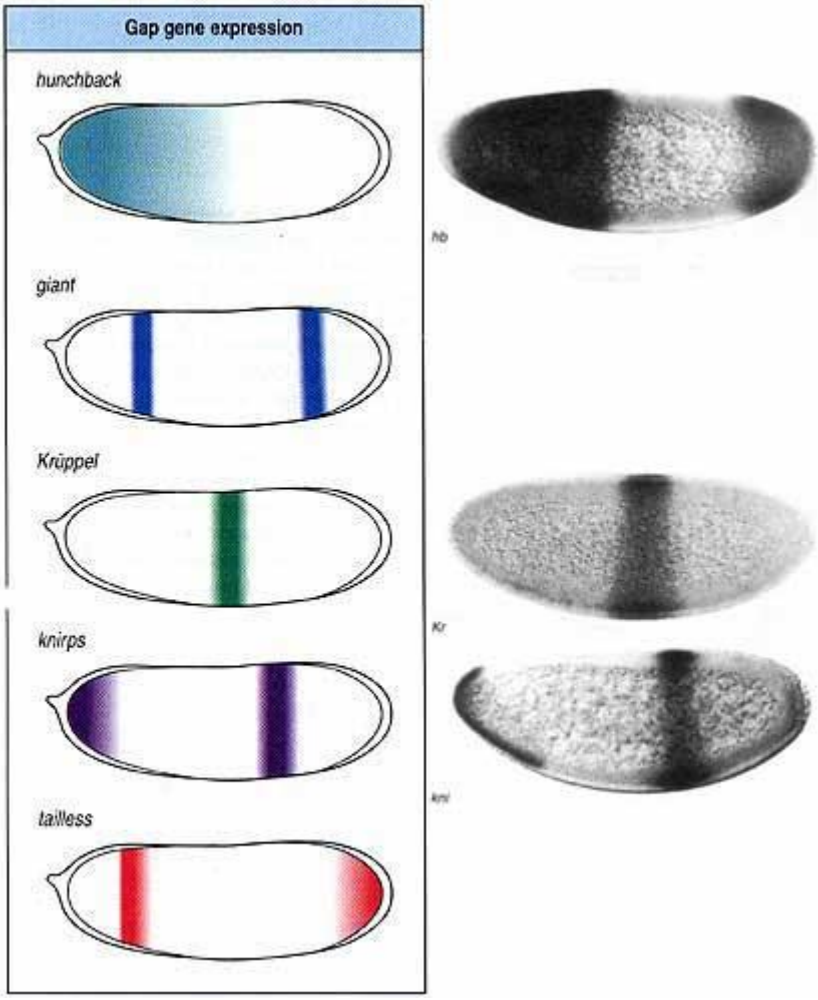
Bicoid is both a transcription factor and a morphogen. The gene product regulates the activity of other genes, controlling their pattern of expression in the embryo. Today I thought I'd get more specific about the downstream targets of *bicoid*, the gap genes.



Expression domains of the gap genes. The pink bars chart the strength of gene expression as a function of position along the lengths of the embryo for *hunchback (hb)*, *huckebein (hkb)*, *tailless*

(tll), giant (gt), Krüppel (Kr), and *knirps (kni).* These are also plotted against the egg/embryo where they are expressed, and the larva. (the larva is shorter than the egg because the anterior part of the embryo tucks itself inward to form the mouthparts.)

The gap genes are zygotic genes. Unlike maternal effect genes, which are transcribed from the mother's DNA, zygotic genes are activated in the fertilized embryo (the zygote) and are transcribed from the zygote's DNA. The gap genes in *Drosophila* get their name from the observation that mutations in these genes knock out, or cause a gap, in the body plan—lose the gap gene *krüppel,* for instance, and a chunk of the embryo's middle fails to develop. The genes themselves are expressed in restricted bands that correspond to the regions that are lost when they are mutated, as can be seen in this diagram and photographs of a few stained embryos.

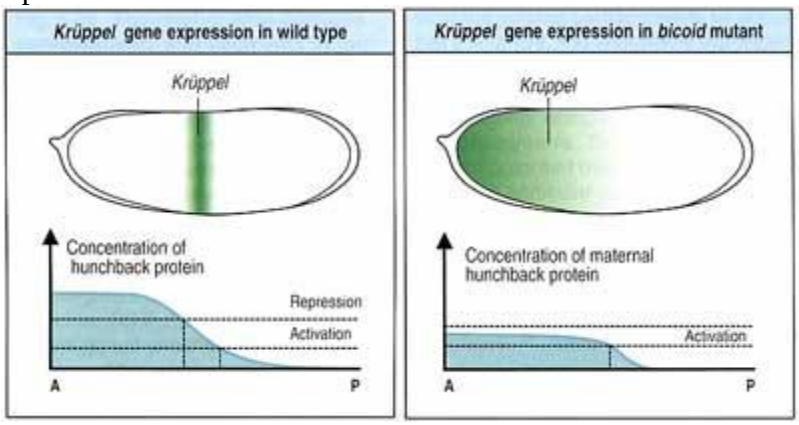


The interesting developmental question here is how a gradient of one gene product, *bicoid* is translated into a pattern of stripes of expression of other genes. The answer is that there is a complex pattern of interactions that have been teased apart, one by one.

For example, one gene, *hunchback,* is relatively simple. It is directly regulated by *bicoid,* so that wherever the *bicoid* concentration is above a certain level, *hunchback* is turned on. As you can see from the diagram, *hunchback* is therefore turned on in cells in the anterior half of the embryo. We can play with this, increasing the concentration of *bicoid* and seeing the expression

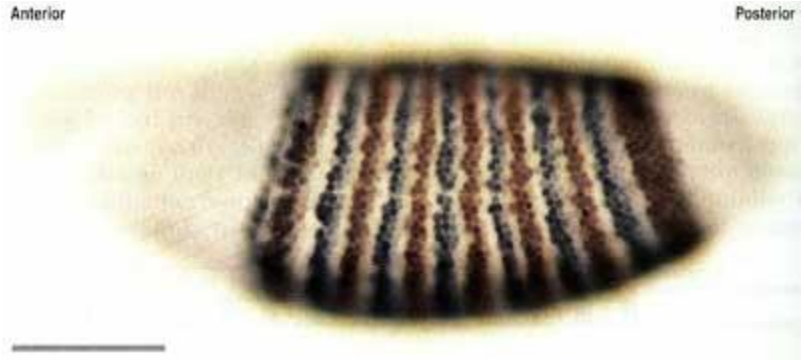
of *hunchback* turned on in more cells, or turning *bicoid* down and seeing *hunchback* similarly down regulated.

What about making stripes? Here, regulation gets more complicated. *Krüppel*, for instance, overlaps in its expression with *hunchback*, and is regulated by it. If there is no *hunchback*, there is no *Krüppel* expression. If there is a little bit of *hunchback*, *Krüppel* is turned on. If there is lots of *hunchback*, though, *Krüppel* is turned off again…so *Krüppel* is only turned on in a narrow band where the *hunchback* concentration is just right, where the fly has the perfect balance between repression and activation.



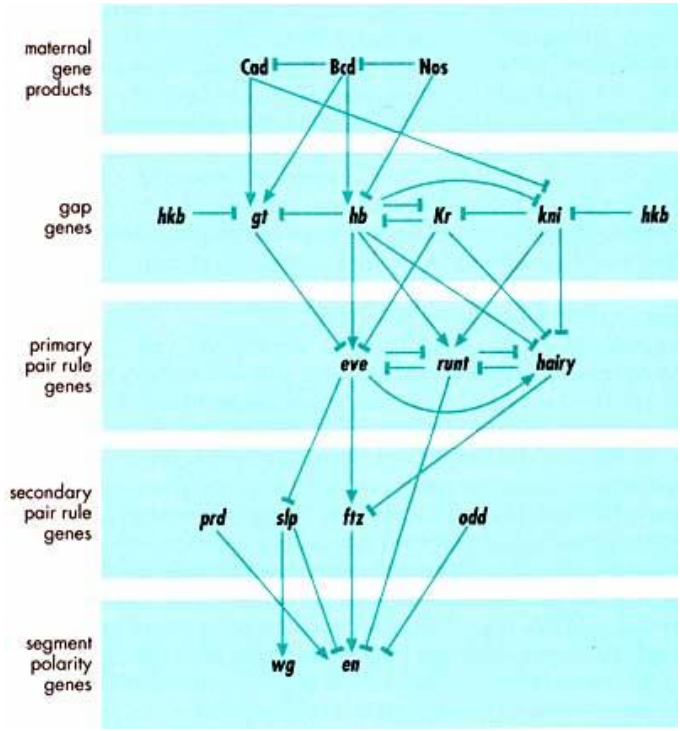
There is more to this story, of course. The gap genes regulate each other, so in addition to being controlled by *hunchback*, *Krüppel* is also inhibited by another gap gene, *knirps*. All of these genes are jostling one another, turning on some genes and turning off others, all triggered initially by the gradient of *bicoid* expression, to produce their final arrangement.

And it doesn't stop there! The gap genes are only the second step in the process. The gap genes in turn regulate another set of genes, the pair rule genes, which produce lovely alternating stripes of gene expression that correlate with every other segment, like this:

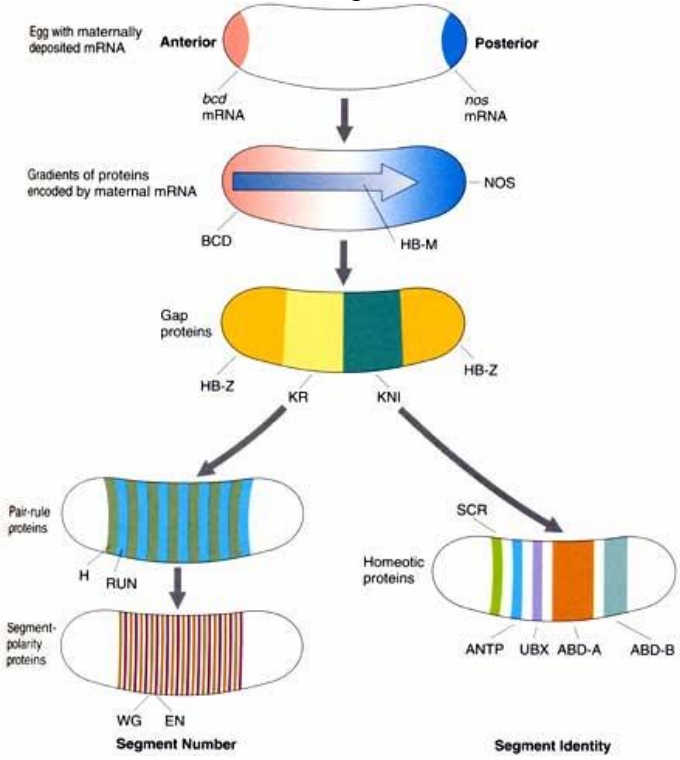


The pair rule genes are numerous, and also interact with one another, and will in turn regulate yet another level of the hierarchy, the segment polarity genes. The segment polarity genes are turned in every segment, within specific sub regions of the segment. Years of tracing these interactions now allows us to assemble diagrams of the regulatory cascade involved that look like this, where arrows indicate that a gene activates another, and bars indicate that it inhibits it:

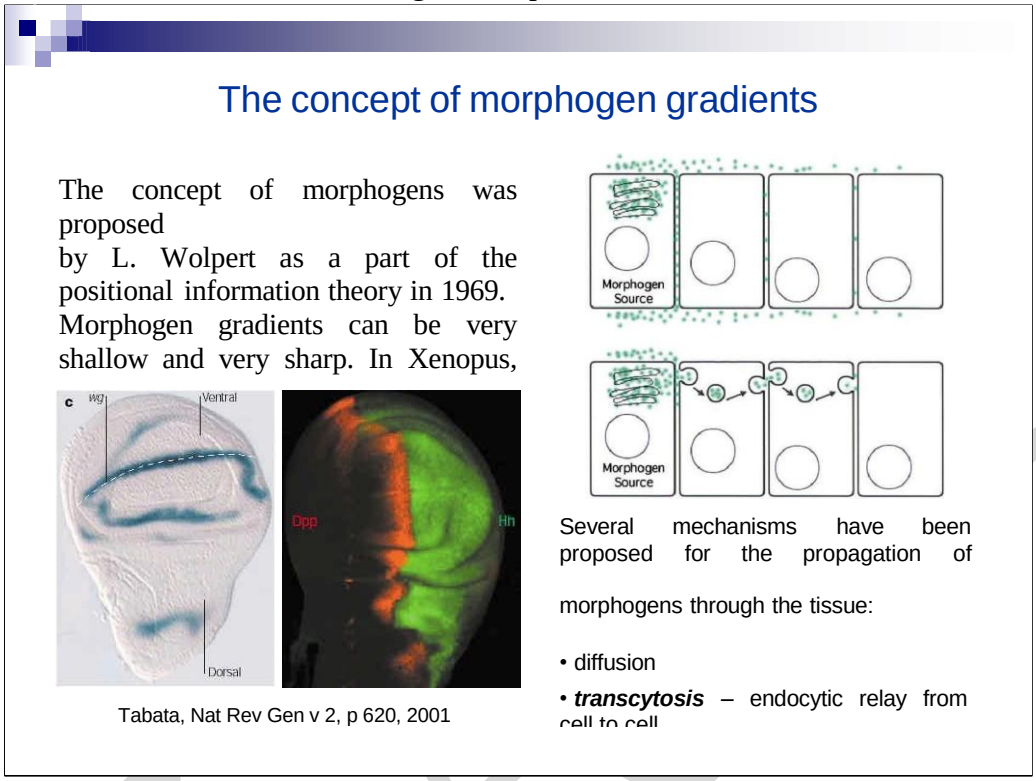
Note that this is not a complete map of all of the regulatory interactions, and that there is much more that has to be filled in below it—with the segment polarity genes *wingless (wg)* and *engrailed (en)*, we're still describing transcription factors that are going to go on to regulate yet another set of genes!



One last summary cartoon, since it illustrates these branching hierarchies of gene regulation simply and leaves me with a point that I'll want to follow up on later:



This shows how a localized deposition of maternal *bicoid* mRNA eventually leads to cells in different regions of the embryo expressing different genes. Note that I've only been discussing the effect of one gene, *bicoid*, and have said nothing so far about another interesting maternal gene product that is also localized in the egg, *nanos (nos)*. *Nanos* is also expressed in a gradient, like *bicoid*, but unlike *bicoid*, it is high at the posterior end and low at the anterior end.



Morphogens and zygotic gene activity in development

The main problem of morphogenesis can be formulated as one question. How do cells know what is their developmental fate? Early in the history of developmental biology it has become clear that for the cells to make a decision on choosing their future, they need to know their position in the developing tissue. This task to provide positional information to the cells was ascribed to the morphogens – diffusible substances able to form gradients in the tissue to enable cells to "read" both direction and the distance from the organizing centers. As opposed to Turing's idea, these morphogens do not have to form any complex patterns themselves, only a system of long and short gradients whose interpretation by individual cells will eventually result in gradual creation of a complex pattern through the process of iterative refinement. In this lecture we consider an example of a very well studied developmental system – Drosophila embryo – which clearly demonstrates how such patterning does occur in nature.

Despite earlier expectations, confirmed morphogens are almost all proteins and not low weight molecules. These are proteins from TGFb, hedgehog and wingless families. Interestingly most of morphogens require preprocessing, such as proteolytic cleavage, to become active. The pro-

protein and active forms may also have very different life times ability to diffuse.

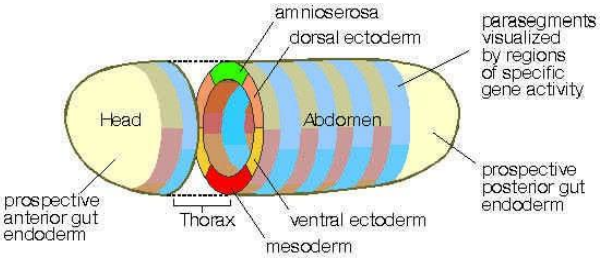
The development of Drosophila is peculiar in a sense that during some 13 first cell divisions the cell-cell boundaries are not formed and the nuclei divide in one giant cell called ***syncytium***. Only after the newly formed nuclei densely populate the near surface, cortex part of the syncytium the cellular membranes are formed to form one layer of cells that covers the original oocyte as a shell. The formation of pattern defining future body plan starts long before this ***cellularization*** process begins. Therefore, in case of early Drosophila development, the morphogen gradients are technically intracellular.

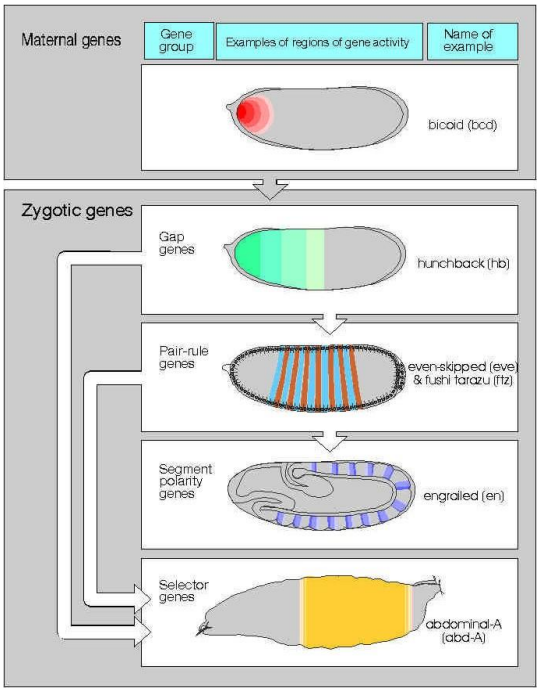
The formation of the body plan is governed by five groups of genes which act in a strict sequence and order. The emergence of pattern is mediated by their complex logic of mutual activation and inhibition.

The master plan of the Drosophila embryo

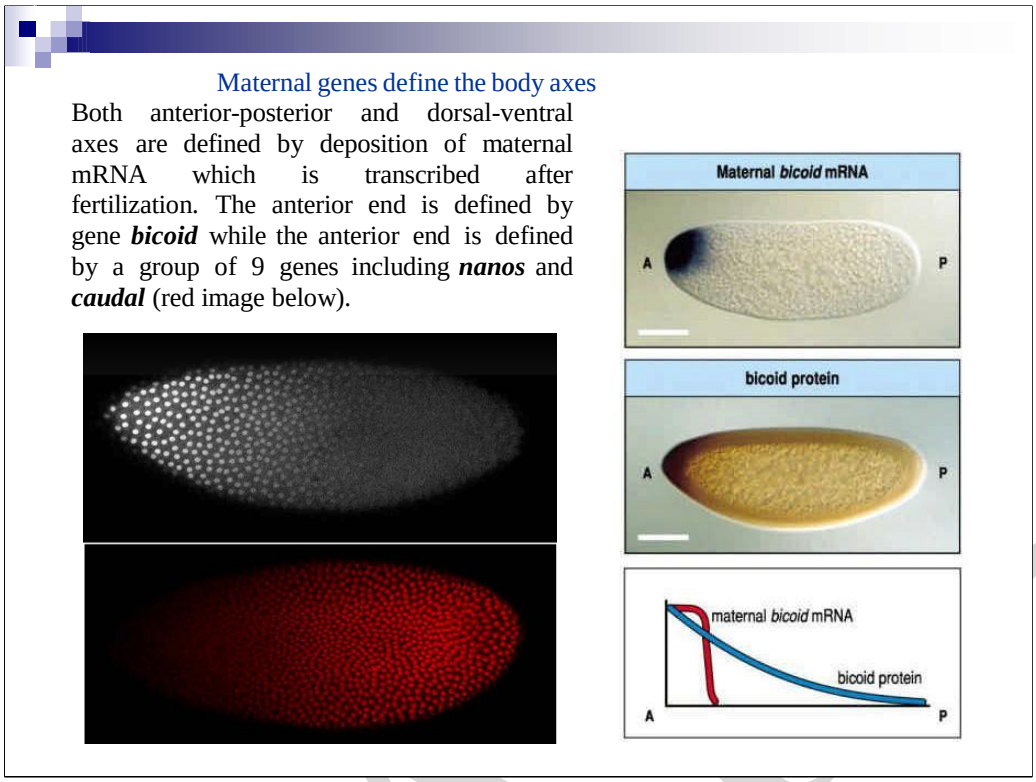
The creation of the body plan of Drosophila by the gradient forming proteins is perhaps the best understood morphogenetic process.

Formation of morphogenetic fields starts on the acellular level of ***syncytium***.





Unless specified otherwise, the illustrations are from

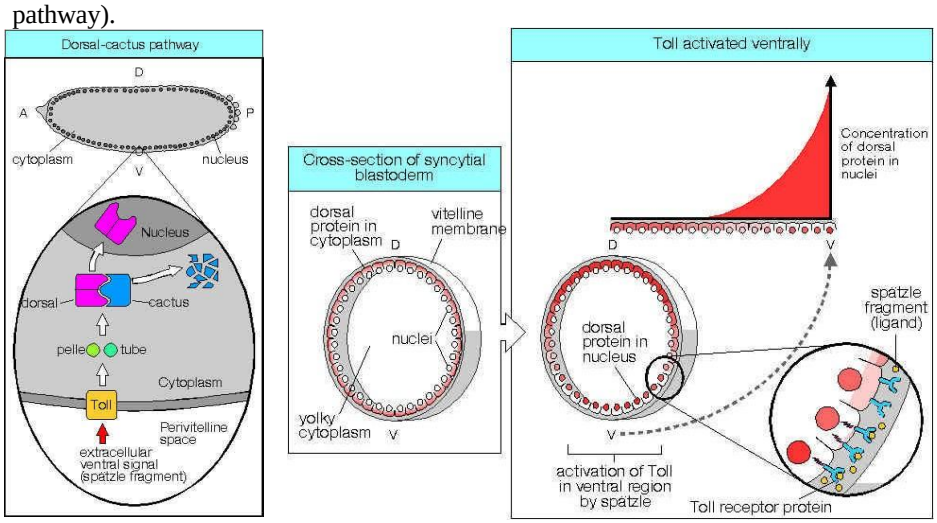


For establishment of the gross axis alignment, Drosophila embryo relies on maternal genes. They are deposited as highly compact and highly localized stores of RNA which begin to be translated into protein shortly after fertilization and long before zygotic transcription begins. This maternal control is exerted by at least 50 different genes while from the theoretical viewpoint it would suffice only two genes to define the two axes. This proves highly redundant and thus robust character of the development.

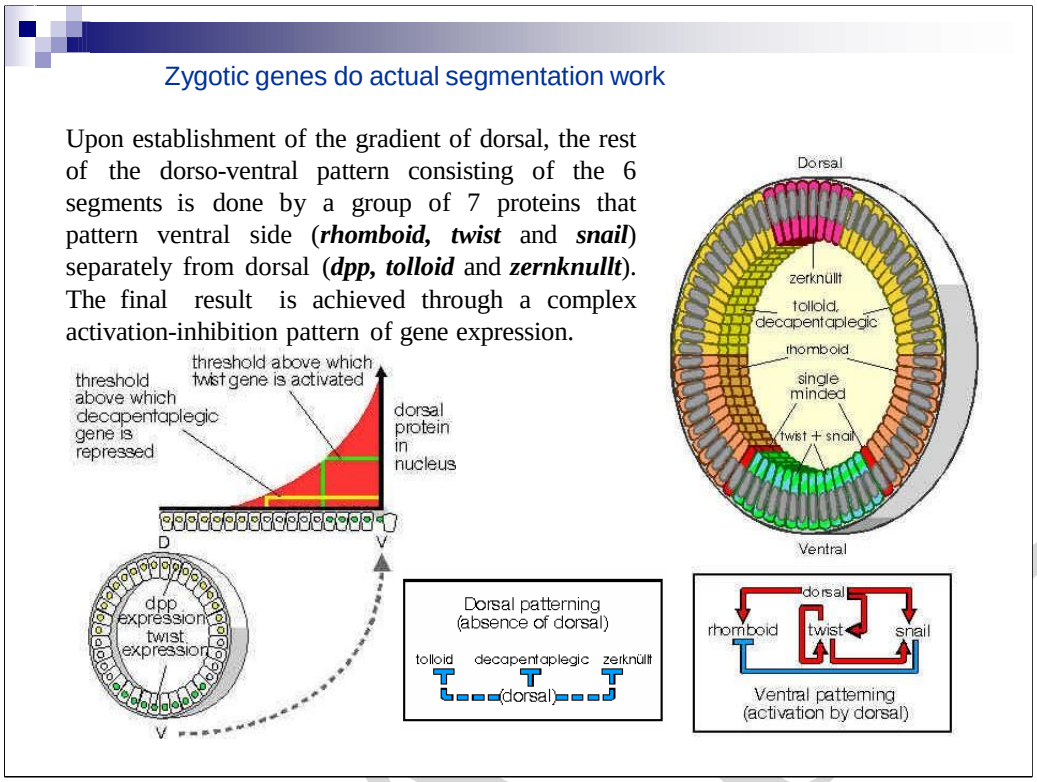
The anterior end of the embryo is defined by gene bicoid while posterior is defined by a group of genes with main genes being caudal and nanos. Bicoid as a transcription factor and acts directly as a morphogen by regulating downstream gap genes. Nanos, on the other hand exists to establish a gradient of another important gene – hunchback, as it suppresses its translation in the posterior end.

Dorso-ventral axis uses different mechanism

Specification of dorso-ventral axis occurs through activation of maternally deposited gene ***dorsal*** by means of extracellular ligand ***spatzle*** deposited on the ventral part of the egg shell through activation of ***Toll*** signaling pathway (analogue of the NF-kB pathway).



Very different from the previous slide mechanism is used to specify dorso- ventral axis. In this case the mRNA for the gene dorsal (NF-kB) is deposited homogeneously throughout the syncytium cortex. However, the inhomogeneity is achieved through spatially heterogeneous action of a signaling pathway which is a homolog of vertebrate NF-kB pathway. The ligand, spatzle, is deposited outside of the egg on the internal surface of the so-called ***viteline membrane*** that lines the inner surface of the egg shell. The ligand is deposited on the ventral surface and therefore the pathway is activated only there.

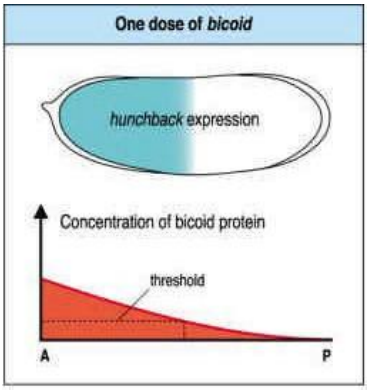


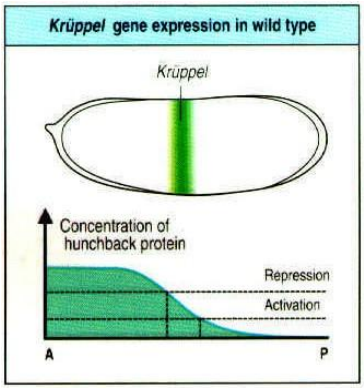
Once the initial symmetry breaking gradients are established throughout the syncytium, the more detailed patterning begins. This process is however already defined by zygotically transcribed genes. The dorso-ventral pattern of six segments is a typical example. In the ventral part, the established gradient of dorsal activates twist, rhomboid and snail. The three genes are related by complex pattern of activation and inhibition which results in spatio-temporal pattern of gene expression which is eventually responsible for definition of cell fates. On the dorsal side, the show is run by the gene product of ***decapentaplegic*** or simply "dpp". This is a homolog of BMP4, one of the most important morphogens in the TGFb family. Its gradient is created by complex interaction with other proteins, for example ***sog*** (short gastrulation) which we will encounter in the next lecture again.

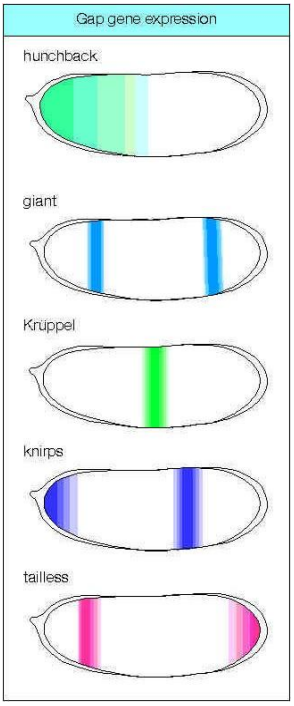
Note that this part of the embryo patterning occurs already on the cellular phase.

Gap genes pattern antero-posterior axis

After establishment of the antero-posterior gradient of maternal genes, the zygotic gap genes switch on. ***Hunchback*** is directly induced by bicoid while ***giant, kruppel*** and ***knirps*** are induced downstream of hunchback. All gap genes are transcription factors necessary for the following tissue patterning. The shown spatial patterns of expression result from complex pattern of mutual activation and inhibition.



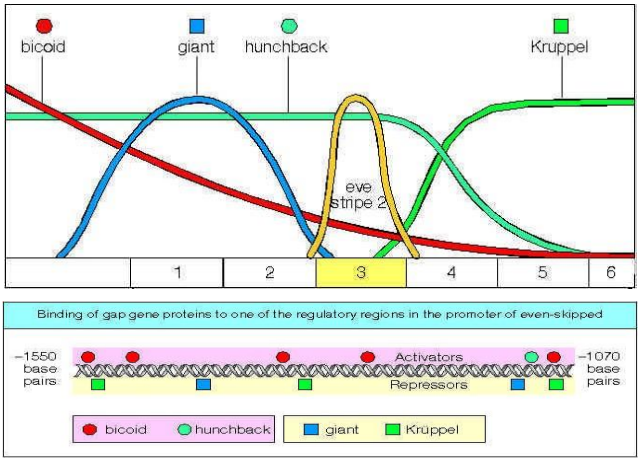


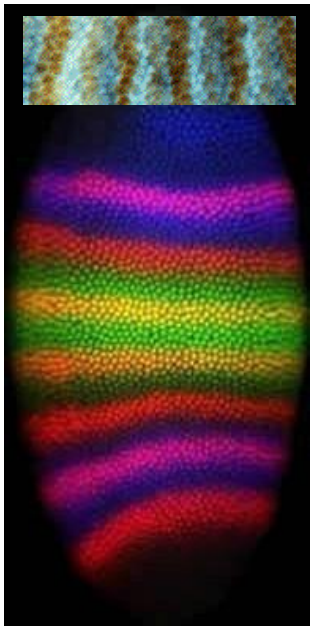


While on the stage of syncytium, the gap genes are induced by the maternal genes whose gradients have already been established. Bicoid directly induces hunchback. As shown on the slide there is low critical concentration of bicoid below which hunchback expression is not activated. The location of this threshold in the embryo defines the position of sharp hunchback boundary. Multistriped patterns of giant, kruppel and knirps are results of similar transformation of smooth gradients into binary outcome by use of thresholds. As shown on the slide, the stripe of kruppel forms on the decaying gradient of hunchback under the existence of two thresholds – for activation and inhibition.

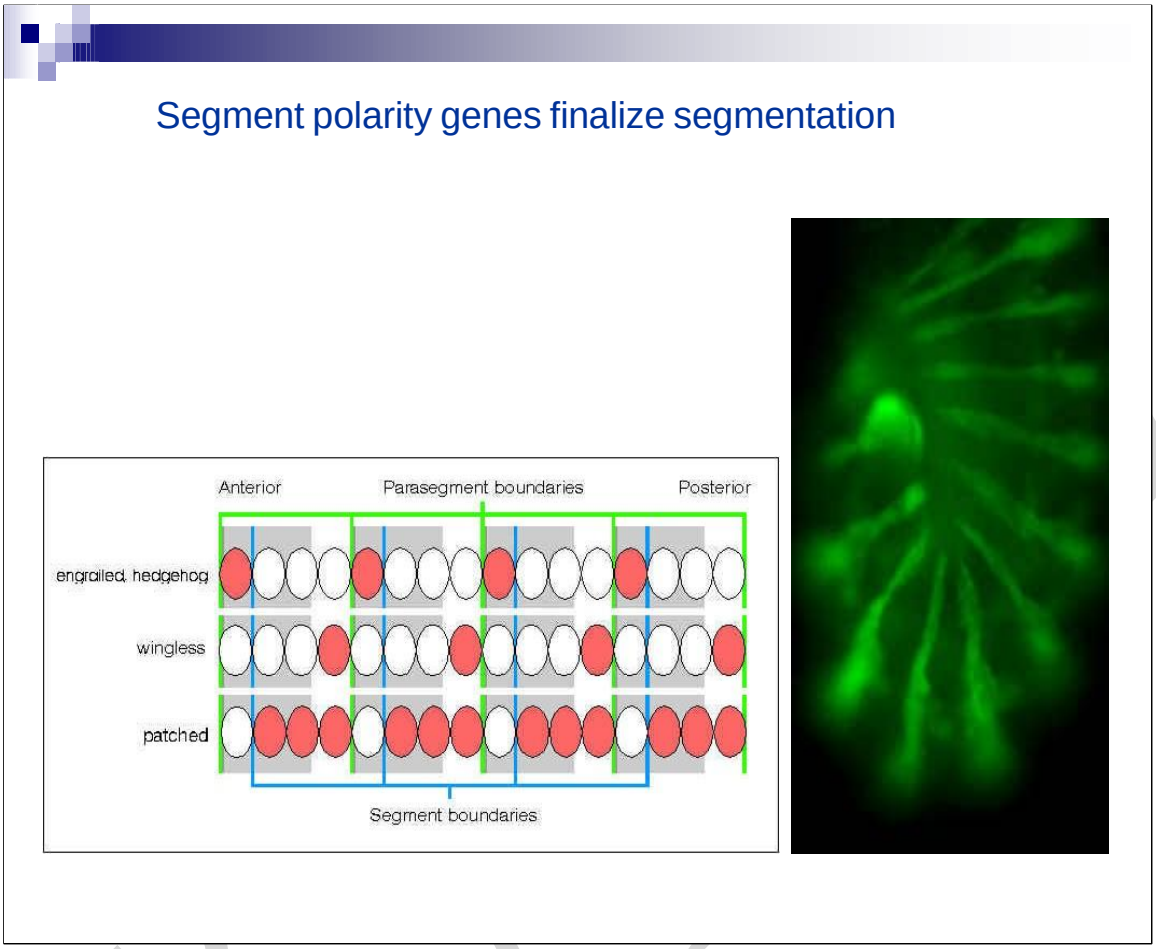
Segmentation and pair-rule genes

Shortly after the gap genes, pair-rule genes ***eve*** and ***fushi tarazu*** create 14 stripes of alternating expression which will become after some modification the segments of the larval body. Each strip has its own genetic control and is **not** a result of a periodic process!

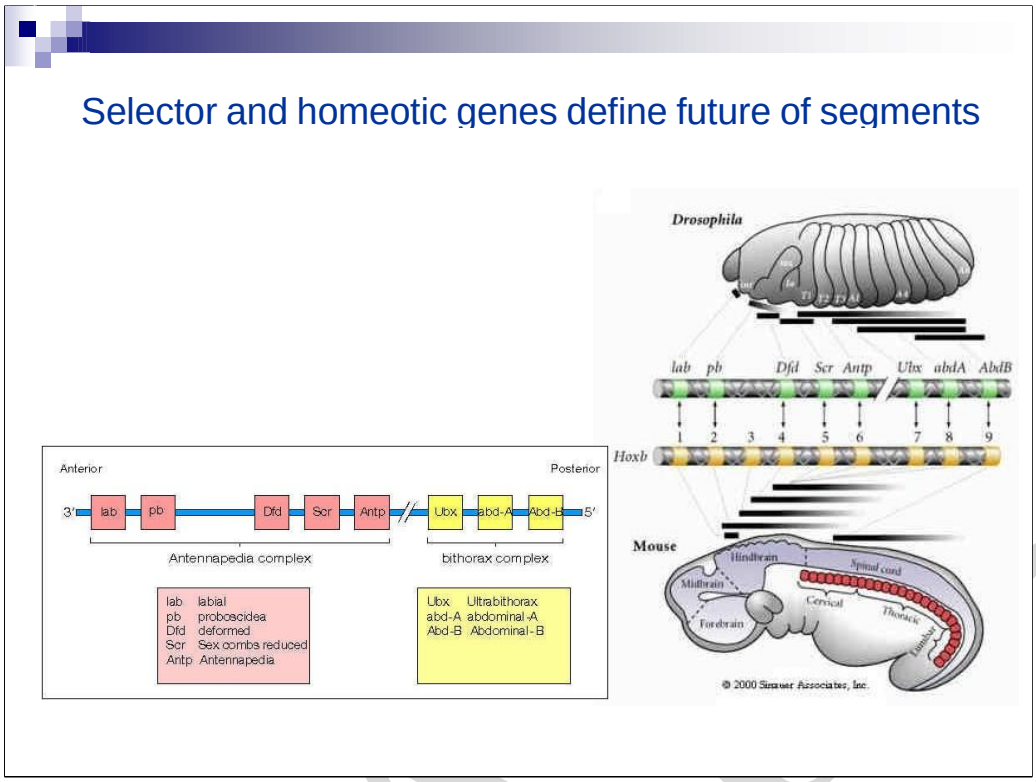




Emergence of periodic stripes of expression of genes even skipped (eve) and fushi tarazu is a magnificent pattern formation phenomenon. Various hypothesis and models were proposed to explain periodic stripes. To the great surprise of both experimentalists and theorists, each stripe is individually controlled by a specific combination of activation and inhibition by the gap genes. Not surprisingly, the regulation regions of eve gene are the most explored and the best understood gene control units up to date. Shown here as an example is the anatomy of the second stripe of eve with the corresponding details of the regulatory element on the eve gene. Overall, the expression of eve (in this stripe) is activated by bicoid and hunchback while giant and kruppel inhibit it. The group of pair-rule genes includes already known to you gene ***hairy***.



Segment polarity genes finally mark the boundaries of the body segments and provides them with antero-posterior direction. The genes of this group code for such important morphogens like ***wingless, engrailed*** and ***hedgehog*** and they are connected by complex interaction network.



Finally, the future of the Drosophila segments is defined by homeotic and selector genes which are homologous to the ***Hox*** genes of vertebrates. These genes can change the whole organ into another organ. For example, mutation ***antennapedia*** results in substitution of antennae by legs. The relationship to simple gradients can hardly be followed on this level of development. The complexity on this level exceeds the scope of this lecture. Just remember that these genes define which segments will turn into which parts of a body.

Sex chromosomes

The sex chromosomes include the X and Y chromosomes, which have a crucial role in sex determination. Males are defined by XY chromosomes and females by XX.

X chromosome

- Has more than 153 million base pairs
- Represents about 5% of the total DNA in women's cells, 2.5% in men's
- Contains about 2000 genes compared to the Y chromosome containing 78 genes, out of the estimated 20,000 to 25,000 total genes in the human genome
- Genetic disorders that are due to mutations in genes on the X chromosome are described as X linked.
- X chromosome inactivation is when most genes on one of the two X chromosomes in females are inactivated and do not produce any product. In somatic cells in normal females (but not in normal males), one X chromosome is inactivated early in development, thus equalizing the expression of X-linked genes in the 2 sexes. In normal female cells, the choice of which X chromosome is to be inactivated is a random one. Thus females are mosaic with respect to X-linked gene expression. In patients with extra X chromosomes, any X chromosome in excess of one is inactivated, but not all genes on that chromosome are inactivated.

Y chromosome

- Much smaller than X
- Carries only a few genes of functional importance, e.g. the TESTIS-DETERMINING FACTOR (also SRY protein – Sex-determining Region Y): certain genes that cause the male sex organs to develop. It is located on the short arm of the Y chromosome close to the pseudoautosomal region (see below).

In male meiosis, the X and Y chromosomes normally pair by segments at the ends of their short arms and undergo recombination in that region. The pairing segment includes the PSEUDOAUTOSOMAL region of the X and Y chromosomes so called because the X and Y-linked copies of this region are homologue to one another like autosomes.

Sex determination

Most body cells contain chromosomes in matched pairs. The number of pairs of chromosomes varies between species. Human body cells have 23 pairs of chromosomes in the nucleus. One of these pairs controls the inheritance of gender - whether offspring are male or female,

- in males, the two sex chromosomes are different. They are XY
- in females, the two sex chromosomes are the same. They are XX.

A set of human chromosomes can be separated from its cell, spread out on a microscope slide and magnified many thousands of times. When stained and photographed, they look like this:

Chromosomes from a male

Chromosomes from a female

Dosage compensation of X-linked gene

Dosage compensation is the process by which organisms equalize the expression of genes between members of different biological sexes. Across species, different sexes are often characterized by different types and numbers of sex chromosomes. In order to account for varying numbers of sex chromosomes, different organisms have acquired unique methods to equalize gene expression amongst the sexes. Because sex chromosomes contain different numbers of genes, different species of organisms have developed different mechanisms to cope with this inequality. Replicating the actual *gene* is impossible; thus organisms instead equalize the *expression* from each gene. For example, in humans, females (XX) silence the transcription of one X chromosome of each pair, and transcribe all information from the other, expressed X chromosome. Thus, human females have the same number of expressed X-linked genes as do human males (XY), both sexes having essentially one X chromosome per cell, from which to transcribe and express genes.

There are three main mechanisms of achieving dosage compensation which are widely documented in the literature and which are common to most species. These include random inactivation of one female X chromosome (as observed in *Mus musculus*), a two-fold increase in the transcription of a single male X chromosome (as observed in *Drosophila melanogaster*), and decreased transcription by half in both of the X chromosomes of a hermaphroditic organism (as observed in *Caenorhabditis*

elegans). These mechanisms have been widely studied and manipulated in model organisms commonly used in the laboratory research setting. A summary of these forms of dosage compensation is illustrated below. However, there are also other less common forms of dosage compensation, which are not as widely researched and are sometimes specific to only one species (as observed in certain bird and monotreme species).

As human females have two X chromosomes, this means they have the ability to produce twice as much X-linked gene product, this Dosage Compensation Mechanism is responsible for keeping levels of the X-linked gene product similar in both males and females. Therefore, such a mechanism only exists in females. Dosage compensation requires RNA polymerase, Tsix transcript and Xist transcript as well as the X chromosomes of developing embryonic cells. Note that each transcript is transcribed on one chromosome each. Xist acts as to ensure inactivation of X chromosome whereas Tsix ensures X chromosome remains active. Such activity is carried out from the X inactivation centre ., also known as XIC. The XIC is located at the centromere and is where decativation of the X Chromosome begins. This mechanism follows the process of X-inactivation, also termed Lyonization or single active X principle, in which regardless of the amount of X chromosomes are present, all but one are inactivated. However, although one of the X chromosomes is said to be inactivated, there is still around 15% of the genes on the inactivated X chromosome which aren't inactivated

In other organisms

Dosage compensation in other sexually-reproducing organisms is achieved through other mechanisms. *Drosophila* transcribe X chromosome genes in males at twice the level of the female X chromosomes. This increase in transcription is caused by changes in chromatin over the male X chromosome. A dosage-compensation complex forms and brings about this up-regulation. It is associated with histone-modifiying enzymes and two non-coding RNAs transcribed from the male X chromosome. The result is equal transcription of the X chromosome genes in males and females.

In nematode worms, there are two sexes - male and hermaphrodite. Males possess one X chromosome, whereas hermaphrodites contain two X chromosomes. Dosage compensation is brought about through the two-fold decrease in transcription in each of the X chromosomes in the hermaphrodite. A dosage compensation complex forms on each X chromosome, dissimilar to that in *Drosophila*, and results in chromosome changes and decreased X chromosome transcription.

Extra nuclear inheritance, tests for organelle heredity and maternal effect

Extra-Nuclear Inheritance in Eukaryotes:

Various cases of extra-nuclear inheritance in different eukaryotic organisms have been studied by several scientists.

Few important examples of extra nuclear inheritance in eukaryotes are stated under some classified subheadings:

(i) Maternal Inheritance:

Maternal inheritance means the inheritance controlled by extra-chromosomal, i.e., cytoplasmic, factors that are transmitted to the succeeding generation through the egg of female organism.

They show the following features:

i. reciprocal differences in F_1;

ii. which in most cases disappears in F_2;

iii. a smaller variation in F_2 as compared to that in F_3.

Maternal inheritance may be, broadly speaking, of two kinds:

i. If some treatments (chemical poison, heat shock etc.) are applied to the female parent, it may affect the egg's cytoplasm. As a result subsequent offspring's are modified in some way. Effects of this kind are called Dauer-modifications or persisting modifications.

It is observed that when protozoa are treated experimentally with chemical poisons or heat shocks, the treatments induce several morphological abnormalities in them. Such abnormalities go on decreasing generation after generation and, eventually, disappear completely through cell division if the treatments are removed.

Further evidences also come from fruit flies subjected to heat treatment and from bacteria treated with chemicals.

ii. Other kinds of maternal inheritance are also known which do not depend upon the repeated application of an external stimulus to the cytoplasm. In this case, maternal inheritance is truly controlled by independent cytoplasmic genes.

Maternal effects reflect the influence of the mother's gene on developing tissues. Many important characteristics of both animal and plants show maternal effects of which some examples axe described next.

(ii) Coiling of Snail Shells (Limnaea Peregra):

One of the earliest and classical examples of a maternal effect is that of the direction of coiling in shells of the water snail Limnaea peregra. In this snail, the shell is spirally coiled. Usually the direction of coiling of the shell is clockwise if viewed from the top of the shell. This type of coiling is called dextral. However, in some snails the coiling of shell is anticlockwise. This type of coiling is sinistral.

The direction of shell coiling of both types of snail is governed by genotype of the female parent and not by their own genotype. Further investigation suggests that coiling depends upon the early clearage in the zygote.

If the mitotic spindle is tilted to left (Fig. 22.1) of the median line of zygote, the successive cleavages will produce a spiral to left (sinistral) and if the orientation of spindle is tilted to the right of the median line of zygote, the successive cleavages will produce a spiral to right (dextral). The spindle orientation is controlled by the genotype of oocyte from which the egg develops.

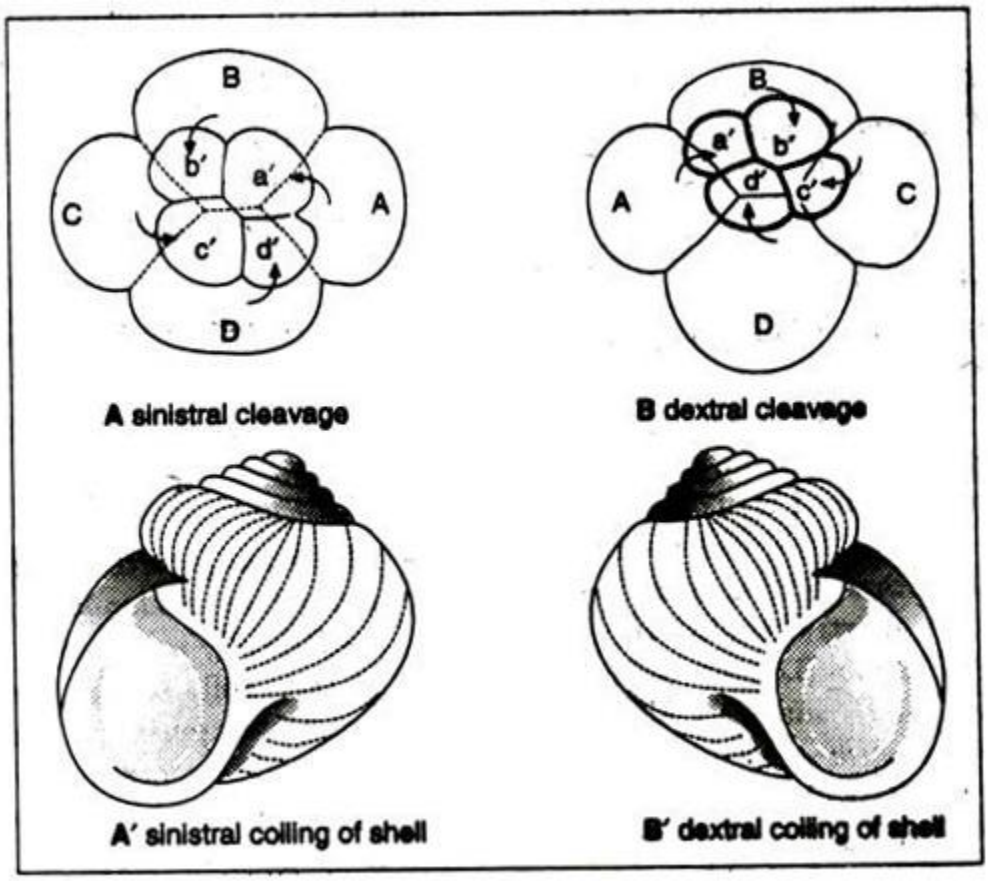


Fig. 22.1: A. Sinistral coiling of shell and B. Dextral coiling of shell.

In a cross (Fig. 22.2) between dextral 9 (female) snail and sinistral (male), (follow the left side coloum of the Fig. 22.2) all the Fi progeny have dextral coils like their mother and also indicates that dextral character (RR) is dominant over sinistral coiling (rr).

However, in the F_1 x F_1 cross (i.e., inbreeding or self fertilisation) all the F_2 snails are also dextral. The F_3 progeny from F_2 individuals with the genotype RR and R_r will show dextral coiling while those from rr F_2 individual will exhibit sinistral coiling of their shell; this produces the typical 3 : 1 ratio in F_3 generation.

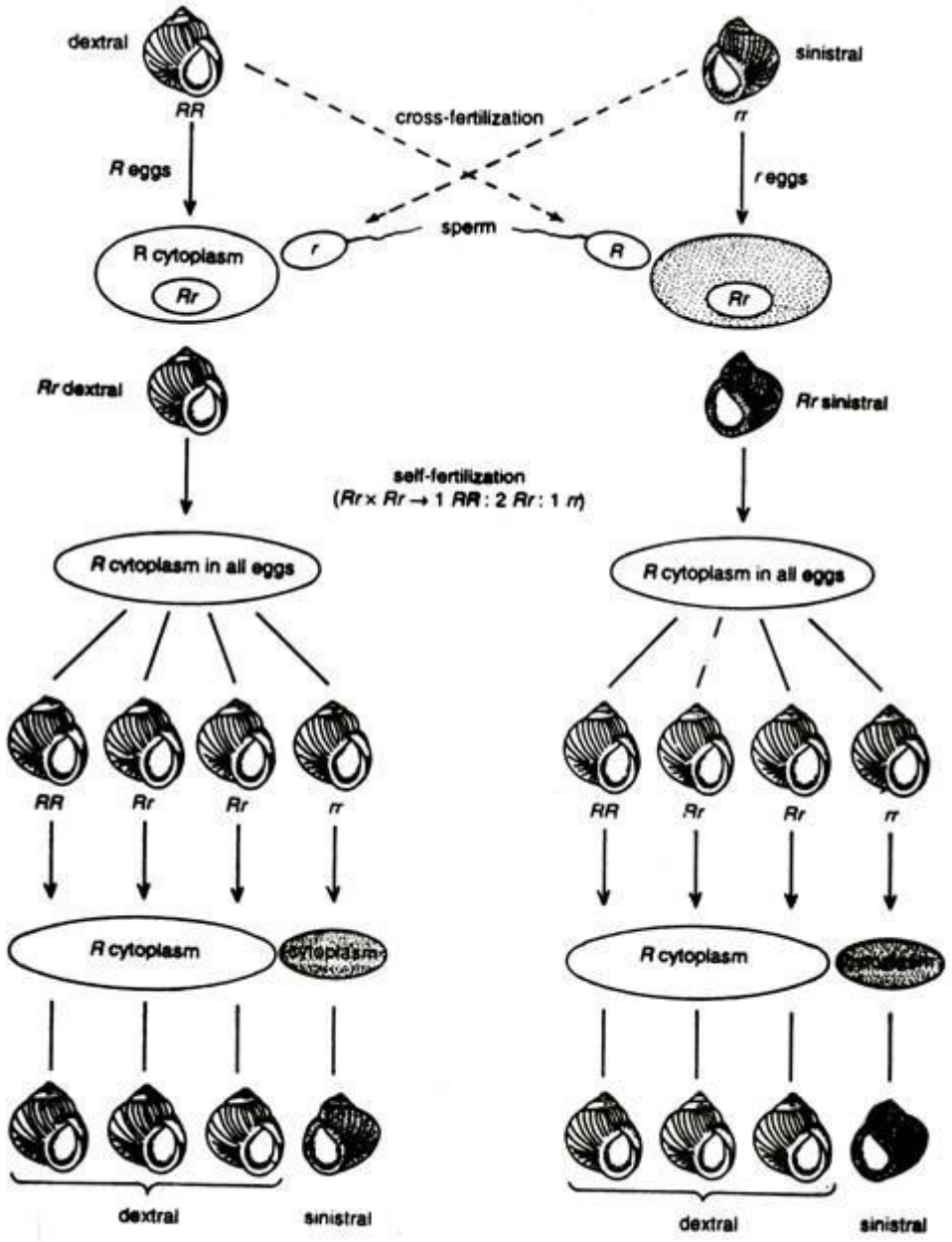


Fig. 22.2: Maternal effect in direction of coiling of the shell in *Limnaea peregra*.

In reciprocal cross (right side column of the Fig. 22.2) between dextral (male) and sinistral (female), all the F_x progeny have sinistral coiling (Rr) instead of dextral coiling. In this case, F_1 x F_1 cross, all the F_2 snails are, again, dextral. This F_3 progenies from F_2 also exhibit the typical monohybrid ratio of 3 : 1.

Thus the features of inheritance of coiling of Limnaea shells may be summarized as:

i. F_1s from reciprocal crosses show differences in coiling pattern.
ii. No segregation in F_2.
iii. Appearance of the typical 3: 1 ratio in F_3 in place of F_2.
The 3 : 1 ratio, although in F_3, clearly indicates that coiling of shell is governed by a single nuclear gene. But the segregation of this nuclear gene is apparently delayed by one generation and is visible in F_3instead of F_2 as in all other cases of Mendelian inheritance. This is because the direction of coiling in this snail is primarily determined by some substances already present in the cytoplasm or ooplasm of the egg cell.
Obviously these substances are produced by the female parent. As a result the offspring would produce the phenotype (in F_1) of the maternal parent since its nuclear gene product is possibly active after one generation later and shows delayed segregation in F_3.

(iii) Maternal Inheritance in Drosophila:

Several examples of maternal effects are known in drosophila:

i. Abnormal growth in the head region of *Drosophila melanogaster* was produced sporadically in a sample from a wild population collected at Acahuizotla, Mexico. Development of abnormal growth in the head region is called Tumorous head (Tu – h). Tu – h is governed by two major genes. But the frequency of tumour development in progeny is markedly influenced by the maternal effect.
When a cross is made between a normal female fly and a male fly with head tumour, less than 1% of the progeny exhibit head tumour. In contrast, when a reciprocal cross is made between a female having head tumour and a normal male, about 30% of the progeny show tumour development.
ii. In Drosophila, fertility and survival are occasionally influenced by maternal effect. A recessive nuclear gene, grandchild-less, affects the fertility of progeny in Drosophila subobscura. In this fly a female homozygous for which is fertile but all her offspring's are sterile.
The reason of this effect arises from cytoplasmic dependent pathway for the development of many organs, the egg cytoplasm formed by a female fly is not uniform and various parts of the egg appear to be specifically assigned for the formation of different tissues. Thus the fate of Drosophila germ cells to produce either various or testes is determined early in development.
A daughterless (da) gene in Drosophila causes death of all XX zygotes derived from eggs of da da females. It is reported that the cytoplasm of eggs of da da females affects the two X-chromosomes and does not inhibit the development of the female phenotype.
It should be noted that the action of da is determined by the genotype of the female producing egg and not by the genotype of the egg cells themselves. It is also reported that interaction of da gene product is evident if the zygote is XX but there is no such action affecting the survival of XY zygote.
Some genes in Drosophila have lethal effects on male embryos. A recessive gene, sonless, leads to the absence of males in the progeny of females homozygous for this gene. Another gene, abnormal oocyte, reduces the frequency of males in the progeny.
These effects on the frequency of male progeny are produced due to effect of egg cytoplasm on the survival of XY zygote. Thus the genes for daughterless and sonless produce typical uniparental inheritance. These uniparental inheritance patterns depend on the nuclear genotype of the female parent.

(iv) Maternal Inheritance in Ambystoma:

Maternal inheritance has also been studied in Axololt of *Ambystoma maxicanum*. In this animal there is a lethal gene 'O'. It is a recessive gene. But its dominant gene is '+'. The heterozygous individuals with genotype '+O' are completely normal.

When two heterozygous individuals with genotype '+0′ are crossed, a homozygous offspring with genotype '00' may be produced. The homozygous individual ('OO') develop normally in the early stage but in later life show a slight retardation of growth.

Their regeneration capacity is reduced to great extent and the homozygous males (OO) are sterile and their testes are poorly developed and spermatogenesis does not produce beyond spermatogonial stages. On the other hand, the homozygous females with genotype 'OO' produce eggs.

The eggs are capable of fertilisation and normal clearage but at the onset of gastrulation, further development is retarded and embryos normally die. This abnormal course of development is affected by the maternal effect, i.e., the growth is resulted by the genotype of the mother irrespective to that of the offspring.

If the maternal genes are recessive, i.e., 'OO', the growth of the offspring even with genotypes '+O' is retarded. According to Biggs and Justus (1967), in the egg of homozygous female 'OO' a protein-like substance called corrective factor is absent.

This protein-like substance is necessary for normal growth and development of the offspring and is synthesized by normal genes '++' of homozygous or heterozygous normal females ('++' or '+O') during oogenesis.

(v) Maternal Inheritance of Eye Pigmentation in Water Flea and Flour Moth:

The maternal inheritance has also been observed in case of water flea (*Gammarus sp.)* and the flour moth (*Ephestia kuhniella*). The normal colour of both the invertebrates is dark due to the dominant gene (AA or KK) in which the dominant gene A or K directs the production of a hormone-like substance called Kynurenine which is involved in the pigment synthesis.

The recessive gene (i.e., a or k) fails to direct the synthesis of kynurenine. In absence of kynurenine, colour of eye becomes light.

Therefore the recessive mutants do not possess pigment in the eye and have the genotype aa or Kk. When a cross is made between a heterozygous male (Aa or Kk) and a double recessive female (aa or kk), only half of the larvae show dark pigment in the eye. Again, when a cross is made between a heterozygous female and a double recessive male, all larva are with dark eye. But on reaching the adult stage, half of the progenies having the genotype 'aa' or 'kk' become light-eyed.

This indicates that some kynurenine molecules diffuse from the 'Aa' mother into all young (larvae) enabling them to produce pigments regardless of their genotype. But the aa progeny is unable to synthesize kynurenine and, consequently, develops light eye as the kynurenine obtained from mother is used up. This example suggests an ephemeral type of maternal inheritance.

(vi) Maternal Inheritance in Mammals:

Maternal effects in mammals are very prominent because mammalian mothers may affect the development of their offspring's not only through ovum cytoplasm but also through the uterine environment. Effects of the maternal Rh blood group on the developing foetus in humans and that of maternal genotype for hair loss gene on the survival of young embryos in mice are the examples of maternal effect.

Other examples of maternal effects in human are: the embryonic defects caused by maternal diabetes and maternal phenylketonuria, a maternal influence on left-or right-handedness and a maternal effect on body weight.

(vii) Maternal Inheritance in Plants:

Cytoplasmic male sterility in cross-pollinating plants is a classic example of maternal inheritance. Plants which are unable to produce functional pollen but possess normal fertile female structure aired known as male sterile.

Thus male sterility is characterized by nonfunctional pollen grains. This occurs in many flowering plants. In maize, wheat, sugar-beets, onions and some other crop plants, fertility is controlled, at least in part—by cytoplasmic factors and results in male sterility.

This type of male sterility is referred to as cytoplasmic male sterility. Nuclear gene did not control this type of sterility, rather it is transmitted from generation to generation through egg cytoplasm. In other plants, however, male sterility is controlled entirely by nuclear genes. But over present discussion will be focused on only cytoplasmic male sterility in relation to maternal inheritance in plants.

Cytoplasmic male sterility is determined by cytoplasmic factors. Since the bulk of cytoplasm of zygote is contributed by the egg cell and the pollen tube containing male gametes contributes negligible or no cytoplasm, sterility factors present in the cytoplasm of egg cell will be transmitted to the offspring which would always be male sterile.

A case of cytoplasmic male sterility was discovered and carefully analyzed by M. M. Rhoades (1933) in maize. He observed a male sterile plant in maize. In this plant pollens are aborted in the anther. The male sterile plant is produced when an egg cell containing cytoplasmic male sterility factor is fertilized by pollen from normal male fertile plants.

It indicates that male sterility is contributed by the cytoplasm of female parent. It is also- observed that when a male sterile female plant is crossed with wide range of fertile males, all progenies axe male sterile in the subsequent generations.

In maize, three distinct male sterile source (Cms) cytoplasm's axe known which are designated as Cms-T, Cms-C and Cms-S. The normal male fertile cytoplasm is known as N- cytoplasm. Each of the three Cms cytoplasm's shows strict maternal inheritance—even when all chromosomes axe replaced from male sterile plants by a male fertile source through repeated backcrossing.

Even then male sterility characteristics could not be avoided and the characteristics till persists. It indicates that if the character is present on chromosome as nuclear gene, then male sterility could be eliminated by repeated backcrossing.

Therefore, it is confirmed that male sterility is not controlled by nuclear gene, i.e., nuclear gene has no influence on cytoplasmic male sterility. In rare cases, male sterile plants produce a few fertile pollen grains.

When reciprocal crosses axe made between male sterile parent (with fertile pollen) and normal male fertile (female), the progeny is found to be male fertile. Such cases confirm maternal inheritance of male sterility.

Recent studies have shown that cytoplasmic male sterility is controlled by either some unique polypeptide produced by mitochondria or by some plasmid-like elements which axe not found in the mitochondria of normal cytoplasm.

Cytoplasmic male sterility due to S-cytoplasm is different from the cytoplasmic male sterility due to T-cytoplasm in several ways. The mt DNA of S-cytoplasm contains two unique plasmid-like DNA fragments called SF (Mol. wt. 3.45×10^6) and S-S (mol.wt.4.10×10^6).

These plasmid-like DNAs are not found in the isolated DNA of chloroplast or nuclei. Therefore, these plasmid-like DNAs are the characteristic of mitochondrial DNA of S-cytoplasm.

These plasmid-like DNAs are also absent in the mitochondrial DNA of N- cytoplasm (Normal-fertile), T-cytoplasm (male sterile) as well as C-cytoplasm (male sterile). Hence it has been suggested that such plasmid- like DNAs in mitochondria are responsible for causing male sterility.

T and C-cytoplasm of male sterile plant is comparatively stable and irreversible. It means that they never give rise to fertile cytoplasm even by applying mutagens.

On the other hand, S-cytoplasm is stable. It is found to change into fertile condition in some cases due to one of the two kinds of change:

i. Cytoplasmic mutation from male sterile to male fertile;

ii. Nuclear mutation giving rise to a new repressor gene.

It is also shown that both these changes are involved to make it fertile.

When a fertile line derived from S-cytoplasm is crossed as male () with a Cm-S tester female plant (), in some cases the offspring is male sterile. It indicates that the restorer gene is, possibly, absent. In other cases the offspring is semi-fertile. It suggests that the restorer gene is possibly, present. This restorer gene is different from the normal nuclear restorer gene – Rf_3 meant for S-cytoplasm.

These new restorer genes are likely to be located on different chromosomes where they are, possibly, attached like episome at different times and bring the change from sterile to fertile condition. It has also been suggested that male fertility genes could be originally located on organelle DNA and were later transposed to a nuclear site giving rise to restorer gene.

The gene or DNA segment that has migrated from the organelle to the nucleus or to the other organelle is termed promiscuous DNA.

When this fertility gene is absent from both organelle and nucleus, this might have led to cytoplasmic male sterility. The restorer gene present in the nucleus as dominant gene generally nullifies the effect of cytoplasmic male sterility so that individuals having a restorer gene in homozygous or heterozygous state are fertile even in the presence of male sterile cytoplasm.

In case of Cms-T, plasmid-like event in the mitochondria is absent but some unique polypeptides are produced in the mitochondria which bring the male sterility. When restorer gene is present in the nucleus, it prevents the production of unique polypeptides in Cms-T, and the plant becomes fertile.

But when nuclear restorer gene is absent, the plant achieves male sterile cytoplasm. Cms-C has also two additional plasmid-like elements like Cms-S. These elements are associated with cytoplasmic male sterility.

The maternal inheritance mechanism that transmits male sterility in maize have also been demonstrated by Dhawan and Paliwal in 1964. In their experiment they used two strains of maize—Sikkim primitive-2 and another strain from Colorado—for reciprocal crosses.

When Sikkim primitive-2 was used as female parent in the cross, the offspring showed little vigour and poor yield, but when Colorado strain of maize was used as female parent, the hybrid were more vigorous and showed high yield potency. These differences in hybrids of reciprocal crosses suggests that yield and vigour are governed by female cytoplasm.

Extra-nuclear Inheritance by Cellular Organelles:

Extra-nuclear inheritance is also associated with certain cytoplasmic organelles (mitochondria, plastids) that contain naked circular DNA and protein synthesizing apparatus. These extra nuclear genetic materials present in the organelles axe autonomous and code only for limited number of enzymes and polypeptides. Certain enzymes required for cellular respiration are synthesized in the mitochondria.

Similarly, chlorophyll and other pigments axe synthesized in the plastid. Besides the involvement of such biosynthetic activities, these organelles DNAs axe directly associated with the inheritance of some phenotypes which are not controlled by the nuclear genes. The genetic material of chloroplasts and mitochondria are transmitted almost exclusively via the egg.

The inheritance pattern is well-illustrated by the following examples:

Plastid Inheritance in Mirabilis:

Plastid inheritance means the inheritance of plastid characteristics due to plasma genes located in plastids. Plastid inheritance was first described by C. Corens (1908) in the four o'clock plant, *Mirabilis jalapa*.

Leaves of *Mirabilis jalapa* may be green, white or variegated and some branches may have only green, only white or only variegated leaves. Variegation means the presence of white or yellow spots of variable size on the green background of leaves.

Thus it forms the mosaic pattern of coloration on a leaf. Due to certain inheritable defects chloroplast of all cells or some cells of leaf often are unable to synthesize the chlorophyll pigments. Such cells remain non- green and form white or yellow coloured leaf, or white or yellow patches, interspersed with areas containing normal green cells with healthy chloroplasts.

Variegation may be produced by:

(a) Some environmental factors,
(b) Some nuclear genes,
(c) Plasma-genes in some cases.

Since the first and second causes of leaf variegation do not concern cytoplasmic inheritance, the inheritance of variegation due to plasma-genes will be discussed in this article.

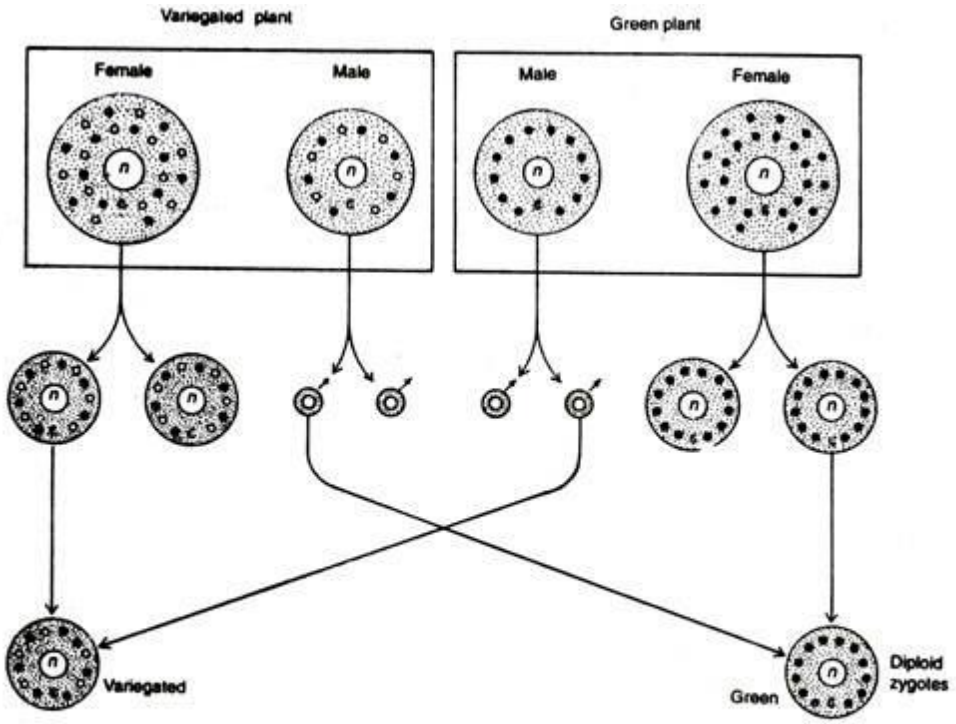


Fig. 22.3: Uniparental (*maternal*) inheritance in *Mirabilis jalapa* which has little or no cytoplasm in pollen gametes.

Correns made reciprocal crosses (Fig. 22.3) in all combinations among the flowers produced on these three types of branches.

These results are summarized in Table 22.1.:

Table 22.1: Plastid Inheritance in Variegated four O'clock Plant

Leaf phenotype of branch used as male plant	Leaf phenotype of branch used as female plant	Leaf phenotype of the progeny (F_1)
Green	Green	Green
	White or yellow	White or Colourless
	Variegated	Green, White or Colourless, Variegated
White or Colourless	Green	Green
	White or Colourless	White or Colourless
	Variegated	Green, White or Colourless, Variegated
Variegated	Green	Green
	White or Colourless	White or Colourless
	Variegated	Green, White or Colourless, Variegated

The results obtained from various crosses of leaf phenotypes of *Mirabilis jalapa*, as shown in Table 22.1, clearly indicates that leaf phenotype of the progeny is the same as that of the female parent (Fig. 22.3). The phenotype of male parent did not contribute anything to the progeny.

This phenomenon is referred to as uniparental transmission. Again, the results of the crosses of *Mirabilis jalapa* cannot be explained by sex-linkage.

The inheritance of different leaf colours in *Mirabilis jalapa* might be explained if the plastids are somehow autonomous and are never transmitted through male parent. For an organelle to be genetically autonomous, it must be provided with its own genetic determinants that are responsible for its phenotype.

Since the bulk amount of cytoplasm containing many plastids is contributed by the egg and the male gametes contribute negligible amount of cytoplasm, therefore plastids present in the cytoplasm of egg is responsible for the appearance of maternal colour in the offspring and the failure of male plant to transmit its colour to offspring is reasonable.

In the offspring from variegated female parents, green, white and variegated progeny are recovered in variable proportions. The variegated parent produces three kinds of egg- some with colourless plastids, some contains only green plastids, and some are with both chloroplasts and leucoplasts.

As a result, zygotes derived from these three types of egg cells will develop into green, white and variegated offspring's, respectively.

Inheritance of Lojap Trait in Maize:

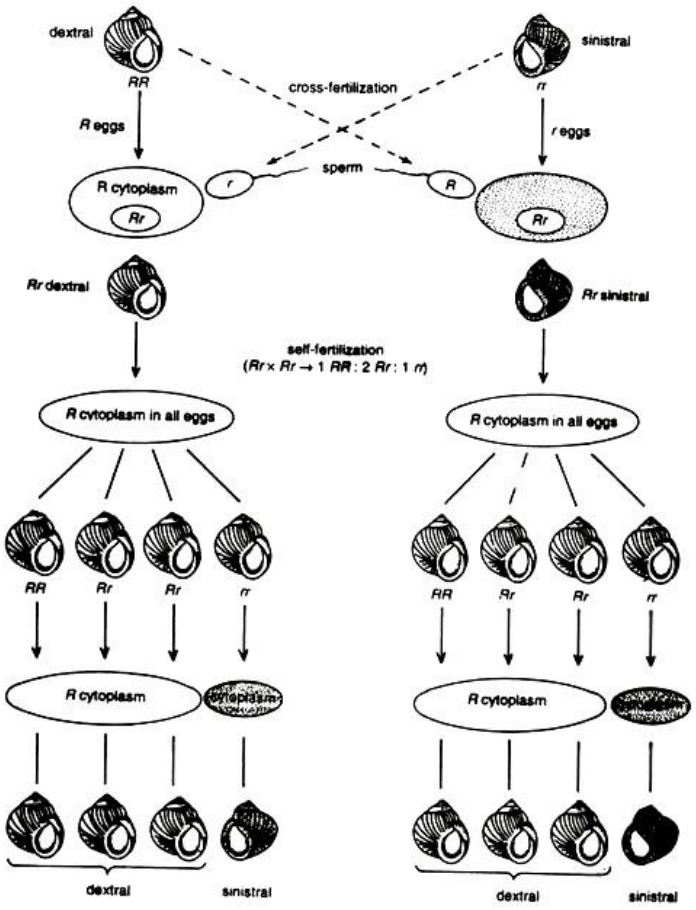
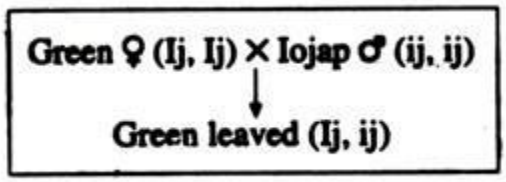


Fig. 22.2: Maternal effect in direction of coiling of the shell in *Limnaea peregra*.

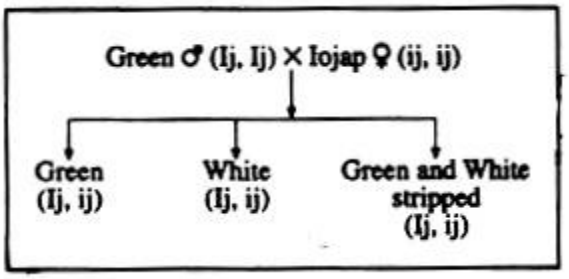
In maize plant, *iojap* is a trait which produces green and white stripped leaves. This trait is controlled by a recessive chromosomal gene (ij) when present in homozygous state. The name iojap was derived from 'Iowa' state (USA), the source of maize strain and japonica, the name of a stripped variety.

When a normal plant with green leaves used as a female parent is crossed (Fig. 22.4) with iojap parent, the offspring will be green leaved:



Again, when a reciprocal cross is made between a normal green plant (used as male) and iojap plant (used as female).

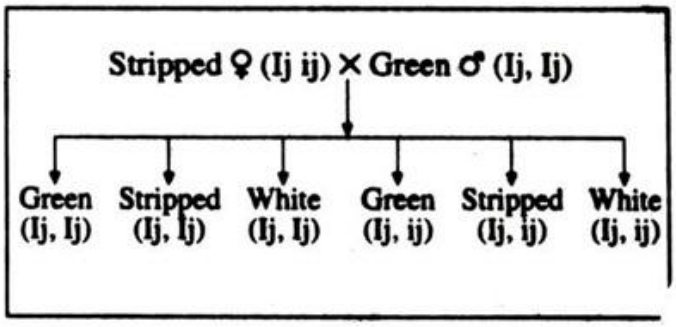
The offspring will be of three different types:



In iojap plants, green and white stripped trait of leaf is inherited from the female parent due to maternal inheritance. It seems that iojap plants contain two types of plastids— normal green, and abnormal iojap plastids.

During the formation of egg cells plastids are randomly distributed in the egg cells. If the egg cell receives normal green plastids it will produce green leaved plants irrespective of which plant acted as pollen parent. If the egg cell receives abnormal colourless plastids, it will give rise to white leaved plants. If the egg cell receives both green and abnormal plastids it will give rise to plants with green and white stripped leaves.

If stripped leaved F_x iojap (I_i,i_j) as female parent is crossed with normal green leaved (Ij Ij) as male parent the following types of offspring are obtained:



This backcross experiment shows that green males have no effect upon progeny. The appearance of iojap trait has been explained by two hypotheses. One hypothesis states that frequent mutation in the chloroplast genome produces the abnormal plastids.

Another hypothesis suggests that certain cytoplasmic elements other than chloroplast mutation bring about the bleaching of chloroplasts. It is also suggested that a nuclear gene controls the development of abnormal plastids in the cytoplasm. So this type of inheritance is a case of inaction between nuclear and cytoplasmic inheritance.

Extra-Nuclear Inheritance by Mitochondria of Yeast:

Yeast, *Saccharomyces cerevisiae*, are unicellular ascomycetes fungi. In this fungi, sexual reproduction takes place by the fusion of two somatic cells to form a diploid zygote nucleus. Next follows two successive nuclear divisions forming four haploid daughter nuclei, all of which take part in ascospore formation. Now the mother cell, i.e., zygote cell, is called ascus.

The diploid zygote can also be grown vegetatively as a diploid strain that will later sporulate. Respiration of yeast cell takes place both aerobically and anaerobically (fermentation). Certain mutant yeast cells are unable to utilise oxygen and are comparatively small- sized and slow growing producing small colonies on agar medium. These small colonies forming mutant strains of yeast are known as petites.

In petite strains, the necessary components (cytochrome b, c_1) and some enzymes (cytochrome oxidase a, a_3) for aerobic terminal respiration activity are absent. But these components are present in the cell of normal strain where they are associated with the inner membrane of mitochondria.

Petite strain can be maintained indefinitely in the vegetative state and can be mated with normal yeast cells. When such mating are carried out, three petite varieties can be classified:

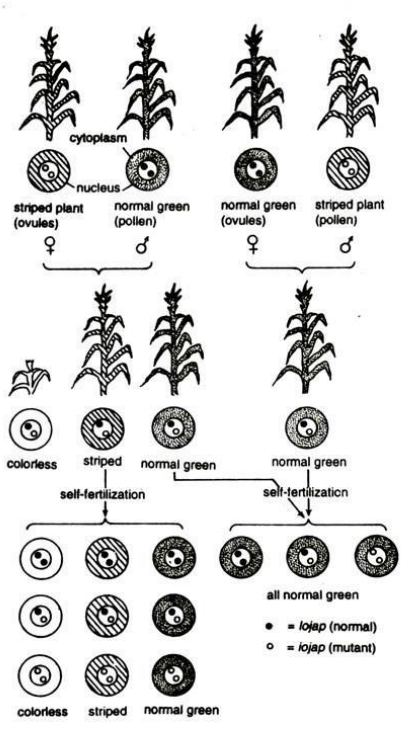


Fig. 22.4: Inheritance of Iojap trait in maize.

Nuclear (segregational) Petites:

When a normal haploid strain of yeast is crossed with a haploid petite strain, a normal diploid zygote is produced. The haploid ascospores produced from zygote by sporulation are segregated in the ratio 1 : 1 (petite : normal). Hence the result of such cross follows ordinary nuclear mendelian inheritance (Fig. 22.5).

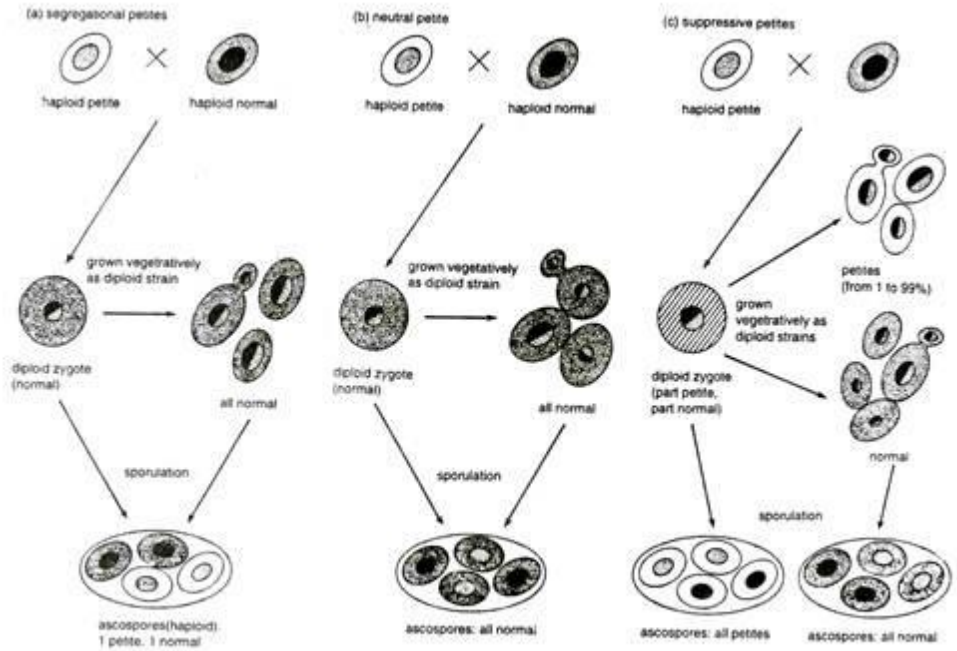


Fig. 22.5: Results of crosses between different kinds of *petites* and normal strain of yeast.

Neutral Petites:

In this type, only normal wild ascospores are produced from matting between petite and normal strain of yeast. The petite characteristics is absent in the product of segregation. So it shows the non-Mendelian inheritance. This non-Mendelian behaviour is very difficult to explain on the basis of nuclear genes and indicates that such petite characteristics are caused by extra-nuclear inheritance.

Suppressive Petites:

In this type, all ascospores produced from mating between normal and petite strain are petite type. Such petites seem to suppress normal respiratory behaviour and the suppressive petite factor acts as a dominant. Fig. 22.5 shows diagrammatic scheme for explaining some differences between neutral and suppressive petite in terms of DNA.

Therefore, there are two different genetic causes for respiratory deficiency in yeast. One is nuclear and the other is extra nuclear. On this basis a neutral petite having the nuclear gene for normally functioning mitochondria is crossed with a segregational petite (Fig. 22,6). The diploid zygote produced from such cross can use the normal nuclear genes from neutral petite and respire normally.

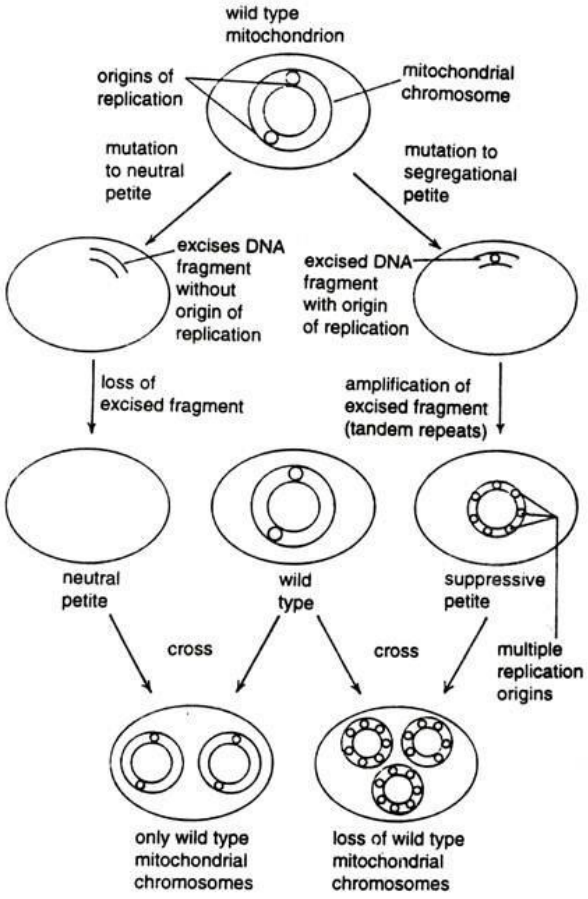


Fig. 22.6: Differences between neutral and suppressive petite in terms of DNA.

When such diploid zygotes are grown vegetatively as a diploid strain, they produce diploid colonies that are of normal size and respire normally. But when such diploid zygotes are allowed to sporulate, they undergo meiosis and produce four haploid ascospores of which two are petites and other two are normal.

It indicates that normal and petite characteristics segregate in the 1 : 1 ratios expected from mendelian segregation.

It is noted that the neutral petite contains the normal nuclear gene for the respiratory enzymes but the segregational petite does not contain respiratory enzymes, so it is obvious that the cytoplasmic factor of the neutral petite appears in the cytoplasm of diploid zygotes where the factor is possibly independent of nuclear control.

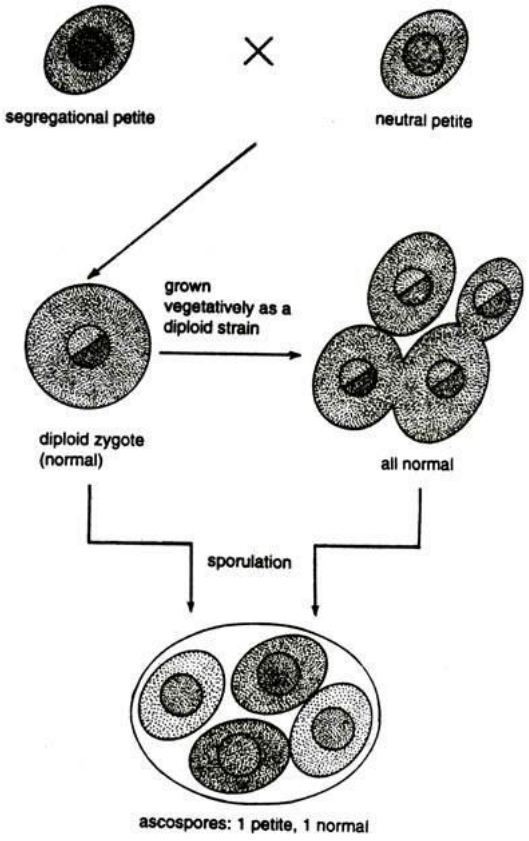


Fig. 22.7: Results of a cross between a segregational petite and a neutral peptite.

It is also noted that neutral petite strains are readily produced by subjecting normal strain in low doses of acriflavines dyes as well as ethidium bromide. But the treatment of such doses of dye does not induce any nuclear changes. Thus it strongly indicates the involvement of extra-nuclear change of gene controlling petite characteristics. Such changes ultimately shows the extra-nuclear inheritance of petite characteristics.

(viii) Extra-Nuclear Inheritance by Mitochondria of Porky Strain of *Neurospora*:

There are many examples for mitochondrial enzyme deficiency which are cases of extra chromosomal inheritance—petite yeast came from the studies of *Neurospora crassa*, a member of ascomycetes group of fungi. In this fungus, there is a slow-growing mutant strain called porky.

Such strain exhibits poorly differentiated mitochondria which are deficient in the membrane-bound cytochromes b, a_1 and a_3, essential proteins of the respiratory electron transport chain and also possess greatly reduced numbers of the small ribosomal units. As in yeast, this trait—inherited via the female parent In some strains in non-Mendelian fashion, indicate its extra chromosomal nature.

When porky as female parent (proto perithecial parent) is crossed with a normal strain as a male parent (mitochondrial parent), the progeny axe found to be porky. In reciprocal cross, the progeny are normal. This non-Mendelian uniparental inheritance suggests that the cytoplasm of female parent is important because the only difference between reciprocal crosses is the contribution of cytoplasm. Thus nuclear genotype has no effect on this particular phenotype.

(ix) Extra-Nuclear Inheritance by Symbionts:

There are many cases of cytoplasmic inheritance which are actually due to the presence of certain intra-cellular parasites like bacteria, virus particles etc. that make a symbiotic relationship with the host cell. They are able to reproduce within the host cell and look like the cytoplasmic inclusions.

These cytoplasmic symbionts provide some evidences regarding the cytoplasmic inheritance of the host cell. These symbionts are variously designated by Greek alphabets as a (sigma), k (kappa), etc. The various types of extra-nuclear inheritances due to parasites or symbionts are discussed next.

Kappa Particles in Paramoecium:

One of the most stricking and spectacular example of cytoplasmic inheritance due to symbiont bacteria is noted in the most common ciblate protozoam *Paramoecium aurelia*. In 1943, T. M. Sonneborn reported that some strains of P. aurelia contain kappa particles and are known as killer strain.

Kappa particles are the symbiont bacteria called *Caedobacter taeniospiralis*. The diameter of kappa particles are about 0.2µ. They are bounded by a membrane and contain a little bit of cytoplasm with DNA. The strain of Parameocium in which the kappa particles are absent are called sensitive strain. The sensitive strains are killed by the killer strain.

The destruction of sensitive strain occurs through secretion of a toxic substance called paramecin. This toxic substance is believed to breakdown the food vacuole membrane of the sensitive strain. Paramecin is diffusible in the liquid medium (Fig. 22.8).

When killers are allowed to remain in a medium for a time, they are not killed. It means that paramecin has no effect on killers. Paramecin is associated with a particular kind of kappa that occurs in about 20 percent of a kappa population.

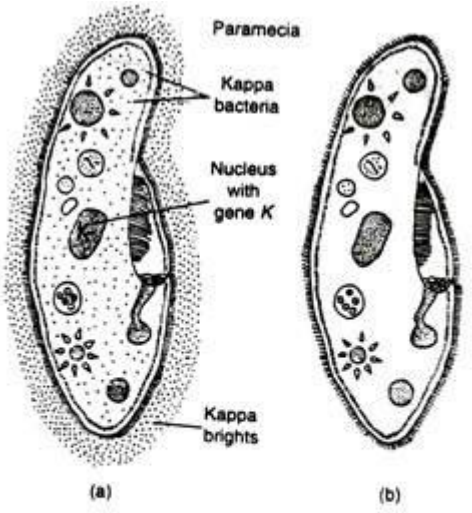


Fig. 22.8: (a) Killer strain Paramoecium with Kappa particles and nucleus with gene K; (b) Sensitive Paramoecium with no Kappa particles and nucleus with gene k.

These kappa bacteria possess a refractile protein containing 'R' body and are called brights because they are infected with a virus that controls the synthesis of a viral protein as well as R protein body in kappa bacterium.

The virus may act as the toxin in the killing response and R body facilitates the penetration of the toxin. The non-bright kappa bacteria may also contain virus but the virus may be in provirus state in them.

The killer character of Paramoecium has a nuclear as well as cytoplasmic basis. The existence of kappa particles is determined by presence of a nuclear dominant gene K. Kappa particles, like other bacteria, multiply through fission.

But their multiplication in the cytoplasm of Paramoecium depends on the presence of a dominant nuclear gene K which helps to make an environment necessary for the bacteria to reproduce.

When killer strain of *Paramoecium* conjugates with sensitive strain under appropriate condition for brief period and no cytoplasm exchange occurs, two kinds of clones result- one from the original killer cell which contains allele K (Kk) and kappa particles and the other from the original sensitive cell which carries the allele K (Kk) and lacks kappa particles.

It indicates that homozygous (either KK or kk) strains become heterozygous following an exchange of K and k genes without cytoplasmic exchange.

Following autogamy (a process of self- fertilization within one undivided cell resulting in homozygosity), half the progency (50%) are sensitive Paramecia. But all progenies of sensitives following autogamy will be sensitive's.

In this conjugation, following autogamy of killers, 50% progeny will receive Kk genotype with cytoplasmic kappa particles other 50% progeny will receive kk genotype with cytoplasmic kappa particles. But it will be sensitive, because kappa cannot reproduce in the cells unless a K allele is present in the nucleus and, as a consequence the kappa are eliminated.

On the other hand, in this conjugation the product of autogamy of sensitive strain obtained after conjugation are all sensitive. All through, 50% progeny of autogamy have KK genotype without cytoplasmic kappa particles because no cytoplasm has been transferred in this conjugation. Remaining 50% progeny of autogamy of sensitive's have kk genotype and no cytoplasmic kappa particles.

Under some conditions of conjugation persists much longer; a long connection is established between conjugants (killer and sensitive). In this conjugation, cytoplasm as well as nuclear genes are exchanged (Fig. 22.9). As a consequence both ex-conjugants will receive the genotype Kk and the cytoplasm with kappa particles.

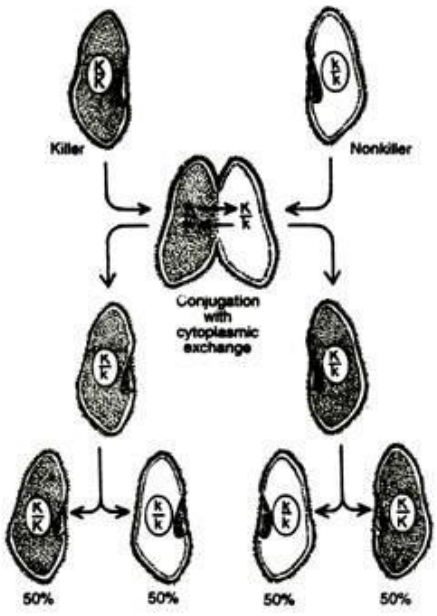


Fig. 22.9: Effect of conjugation for longer period with cytoplasmic exchange followed by autogamy.

Therefore, conjugation for longer period with cytoplasmic exchange will produce all killer strains. Autogamy of both ex-conjugants produces homozygotes KK (killer) and kk (sensitive) cell in the 1 : 1 ratios, respectively, as expected from Mendelian segregation.

Therefore, conjugation for shorter period without cytoplasmic exchange does not follow the Mendelian pattern of inheritance. Hence it confirms the cytoplasmic basis of inheritance of killer trait.

Mu Particles in Paramoecium:

There is another type of killer trait found in certain strain of *Paramoecium* due to presence of 'mu' particles in the cytoplasm. A *Paramoecium* with a 'mu' particle is called mate killer. On the other hand, a *Paramoecium* having no 'mu' particles is called mate sensitive.

It is so named because when a *Paramoecium* with 'mu' particle conjugates with a partner *Paramoecium* without 'mu' particle then the former kills the latter. The 'mu' particles exist only in those cells whose micronucleus contains at least one dominant of either of the two pairs of unlinked chromosomal genes such as M_1 and M_2. The 'mu' particles are symbionts which are made of DNA, RNA and other substances.

The maintenance of the 'mu' symbiont in a *Paramoecium* is dependent upon the genotype of the *Paramoecium*. In fact, the mate-killers of few genotypes maintain their normal number of particles for about seven generations. From the eighth generation, the particles suddenly and completely disappear from the a small fraction of the cell.

Gibson and Beale (1962) suggested that the maintenance of 'mu' particle in *Paramoecium* was due to the presence of another cytoplasmic particle called metagon. It is possibly a long-lived messenger RNA or informosome and may be a product of M_1 and M_2gene. One metagon may be necessary for the maintenance of hundred 'mu' particles.

(x) Sigma Virus in Drosophila:

Some strains of Drosophila melanogaster are sensitive to $C0_2$ as they die when briefly exposed to $C0_2$, while normal flies can be exposed for long periods to pure $C0_2$ without permanent damage. The high degree of $C0_2$– sensitivity is associated with the presence of a DNA virus called sigma factor found in the cytoplasm of $C0_2$ sensitive Drosophila.

Sigma factor is transmitted through the egg cytoplasm. When a cross is made between $C0_2$– sensitive female with normal male, all offspring's are $C0_2$ sensitive. Again, in reciprocal cross, i.e., a cross between normal female and $C0_2$ sensitive male, most of the offspring's are normal except for a small proportion of progeny which are $C0_2$sensitive. Therefore, the inheritance pattern of $C0_2$ sensitivity is non-Mendelian and confirms the cytoplasmic basis of inheritance.

(xi) Spirochaetes and Maternal Sex Ratios in Drosophila:

Spirochaetes sometimes enter into the female body cell of Drosophila and live there as endoparasites. When spirochaetes enter the egg cell and these infected egg cells are fertilised, the zygotes having XY sex. chromosome are killed early in embryonic development and XX zygotes survive.

Therefore, the presence of spirochaete in the female body gives rise to exclusively female progeny—this condition is known as maternal sex ratio. It is evident that XY embryos are killed by a toxic substance which may derive from a DNA virus present within spirochetes that live as endoparasite in the female body of Drosophila. Maternal sex ratios in Drosophila is also considered as an example of extra nuclear inheritance.

Milk Factor in Mice:

This is an interesting example of extra nuclear inheritance. It is found that certain types of mice are very susceptible to mammary cancer and this characteristic is found to be transmitted maternally.The results of reciprocal cross between susceptible mice and low-incidence mice depend on the trait of female parent.

When the young mice of low cancer incidence parent are allowed to feed milk by a susceptible foster mother, it produce, a high rate of cancer in them. Hence this is a case of infective agent transmitted in the milk. The milk factor responsible for causing cancer is possibly a virus. The presence of milk factor depends on nuclear gene.

(xii) RNA Viruses in Fungi:

Like *Paramoecium*, there are two strains of yeast (*Saccharomyces cerevisiae*). One strain is killer and other one is sensitive. The killer strain secretes a proteineceous toxic substance that kills the sensitive strain of yeast cell.

When a cross is made between killer and sensitive strain of yeast, only killer offspring's are produced—indicating uniparental inheritance. There are some other strains of yeast which are called neutral strains.

Neutral strains are neither killed by killer nor do they kill the sensitive strain. But the cytoplasm of both killer and neutral strains contain two types of double-stranded RNA in the form of isometric virus-like particles (about 39 nm in diameter).

The existence and maintenance of virus particles in the yeast cytoplasm are controlled by some dominant nuclear genes called MAK genes (maintenance of killer). Some other nuclear genes—e.g., KEXx (killer expression) and KEX_2—convert killers into neutrals.

A similar situation is noted is case of *Ustilago maydis*, a maize smut fungus. Here the cytoplasm of killer strain also contains maycovirus like particle containing double-stranded RNA. Killer strain secretes a toxin which kills sensitive strains but it has no lethal effect on resistant strains. Resistant

strains are particularly resistant to one of the killer strains designated as p_1, p_4 and p_6. Some nuclear genes denoted as P^r_1,P^r_4 and p^r_6 convert sensitive strain into resistant ones.

In all such cases mentioned above, the virus like particles are not the integral part of the normal cellular organization but their existence and transmission indirectly provides some evidences in favor of cytoplasmic inheritance.

(xiii) Uniparental Inheritance in Chlamydomonas Reinhardi:

R. Sager (1970) and N. Gilham (1968) have reported some cases of extra-chromosomal inheritance in green alga *Chlamydomonas reinhardi*. The alga reproduces by asexual as well as sexual means. The sexual reproduction takes place by fusion between two morphologically similar but physiologically dissimilar haploid gametes coming from different haploid parents designated as '+' and '-'.

The gametic fusion produces the zygote. The sex is determined by a single chromosomal gene. When meiosis occurs in the zygote, four haploid daughter protoplasts are formed which give rise to a new plant body. Although both the sexes contribute equally to the zygote, there is maternal transmission of certain cytoplasmic traits.

Chlamydomonas is a haploid unicellulate green alga. It has two mating types—'+' and '-'.The two mating types are governed by two alleles of a nuclear single gene. The alleles axe named as mt^+ and mt^-. The + mating type is considered as female, while the – mating type is regarded as male. During sexual reproduction one mt^+ and one mt^-cell pair and fuse together to form a zygote where there is mixture of cytoplasm coming from both mt+ and mt^- gametes.

The zygote undergoes meiosis to produce 4 haploid meiozospores of which two zoospores contain '+' alleles and other two contain '-' alleles, i.e., it shows typical 1 : 1 segregation for nuclear genes. But for their plasma genes all zoospores are identical and contain only mt^+type plasma genes by mt^+ plasma genes. The inactivation is not clear but it may involve an enzymatic process.

R. Sager isolated two strains of *Chlamydomonas:* one strain was resistant (S_r) to 500/xg of streptomycin per ml of culture solution and the other one is sensitive. The trait of streptomycin resistance is believed to be located in its cp-DNA (chloroplast DNA).

Mating between mt+ streptomycin resistant (S_r) and mt sensitive (S_s) cells produce only resistant progeny but the nuclear genes for mating type segregate as expected (Fig. 22.10). But the reciprocal cross between mt+ susceptible and mt^- resistant shows again the expected segregation for mating type but all progenies are sensitive type. Therefore, it clearly provides an example for extra-nuclear inheritance. It is also observed that in less than 0.1% of zygotes plasma genes from mt" parent are not inactivated and produce cytohets, i.e., heterozygotes for cytoplasmic genes.

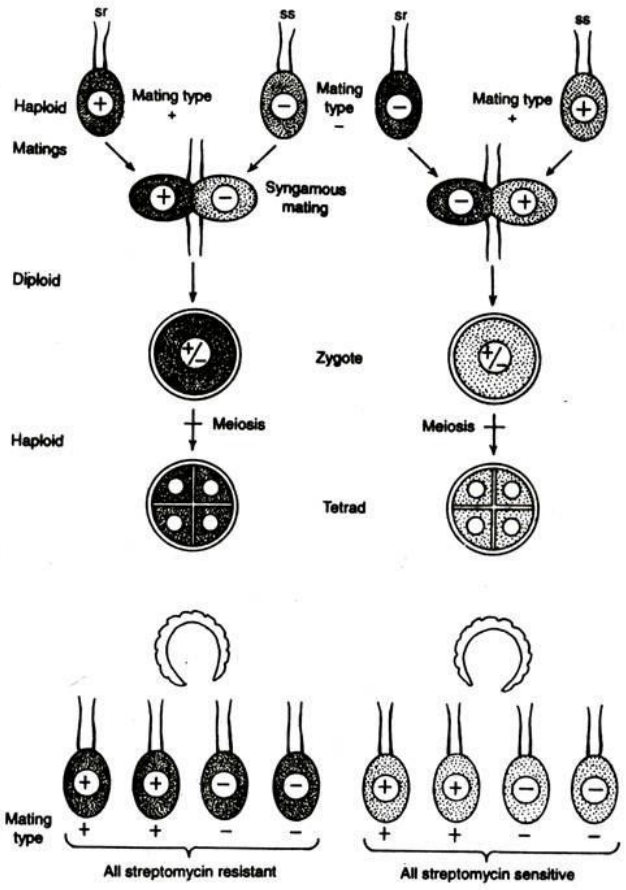


Fig. 22.10: Uniparental inheritance in *Chlamydomonas reinhardi*.

(xiv) Inheritance of Preformed Structures:

In some lower group of animals like ciliated protozoa there are some characteristics which show cytoplasmic inheritance. These traits are not controlled by any plasma gene. In *Paramoecium*, the cytoplasm is differentiated into cortex or ectoplasm and medulla or endoplasm.

The mouth and the contractile vacuole are the prominent preformed traits that are present in the cortical region of *Paramoecium*. These preformed structures can be transmitted independent of the transmission of nuclear genes and plasma genes.

For normal sexual reproduction two individuals of *Paramoecium* called conjugants come close together, exchange their nuclear material, and then separate as ex-conjugants with zygote nucleus. In some rare cases, conjugants do not separate and remain as doublet animal with two sets of mouth and contractile vacuole structures.

When this doublet animal reproduces asexually by binary fission, they also give rise to doublet ex-conjugants. When doublets are mated with normal singlet's, the progeny of doublets ex-conjugants are doublets and the progeny of singlet's are singlet's. This type of inheritance is also found when the animals reproduces through autogamy.

In these crosses it is noted that the nuclear genes of doublets and singlet's are inherited in normal Mendelian fashion. But the mode of duplicated mouth and contractile vacuole is independent of the mode of inheritance of nuclear genes as well as cytoplasmic hereditary factors.

Furthermore, the transplantation experiments reveal that such cortical structure (mouth, contractile vacuole) are autonomous and are maintained by cell division.

Some born has suggested that different parts of the cortex might serve as sites for the specific absorption and orientation of molecules derived from the milieu and genetic action. Besides this, preformed cortical structures could act by determining where , some gene products go in the cell, how these combine and orient and what they do.

(xv) Extra-Nuclear Inheritance in Bacteria:

In many cases, extra-nuclear inheritance has also been reported in bacteria. In addition to main chromosome, bacterial cell has got a unique extra-nuclear genetic system in their cytoplasm. Such extra-nuclear genetic system plays an important role for cytoplasmic basis of inheritance.

In bacteria, the cytoplasmic inheritance is performed by:

(i) plasmids, and

(ii) episome.

In addition to main chromosome, subsidiary DNA is also present in the bacterial cell in the form of plasmid. Plasmids are mini-circular DNA duplex and are capable of independent replication and transmission. By definition, a plasmid is a relicon that is stably inherited in an extra-chromosomal state.

Plasmids are not essential for the survival of bacterial cells, i.e., disposable-except under certain environmental conditions. Plasmids vary in size and contain three to several hundred genes. A bacterial cell may contain more than one plasmid.

There are several types of bacterial plasmids of which three general types have been studied extensively such as:

(a) F plasmid;

(b) R plasmids, and

(c) Col plasmids.

F plasmids carry genes for the development of F pili and are responsible for conjugation.

R plasmids carry genes which are responsible for resistance to antibiotics or other antibacterial drugs.

Col plasmids carry genes which code for colicins. Colicins are proteins that kill sensitive *E.coli* cells.

Plasmids may again be divided into two types on the basis of whether or not they mediate conjugative self-transfer.

They are:

(a) Conjugative and

(b) Non-conjugative.

Conjugative plasmids are also known as transmissible plasmids that mediate the transfer of DNA through conjugation. All F plasmids, R plasmids and some col plasmids are the examples of conjugative plasmids.

These plasmids spread rapidly among the bacterial cells of a population. Transmission of R plasmids is responsible for many pathogenic bacteria to become resistant to many of the widely used antibiotics such as penicillin, tetracycline, streptomycin, kanamycin, chloramphenicol etc.

The transmission of these plasmids are not only restricted among the population of the same species but are also known to transfer to others like Proteus, Salmonella, Hemophilus, Pasteurella, Shigella etc.

Non-conjugative or non-transmissible plasmids do not mediate DNA transfer through conjugation. Some R and Col plasmids are of this type.

Some plasmids are capable of becoming attached and integrated into the bacterial chromosome. Then they are named episomes.

Plasmid and episomes contain insertion sequences which are also present in bacterial chromosomes. Insertion sequences are transposable in that they can move about within and between chromosomes and mediate genetic recombination between otherwise non-homologous genetic elements within which they are located.

Insertion sequence are also responsible for the transfer of genetic controlling resistance to antibiotics from one genetic element to the other.

Considering the mode of transmission, location and the presence of genes controlling certain characteristics, it is clear that the phenomenon of extra-nuclear inheritance still exists even in most simple, prokaryotic unicellular organism like bacteria.

Considering the discussion of this article we can summarize the characteristic features of cytoplasmic inheritance as:

i. In case of cytoplasmic inheritance, reciprocal crosses show marked difference for characteristics governed by plasma genes.

ii. In most cases female parent contributes the plasma genes, i.e., uniparental inheritance or maternal inheritance.

iii. In general, F_2, F_3 and so on generations do not show segregation for cytoplasmically inherited characteristics. It is a non-Mendelian inheritance.

iv. In case of bi-parental inheritance, irregular segregation takes place in F_1.

v. Several plasma genes are associated with cp-DNA or mt DNA in higher eukaryotic organisms and with plasmids or eRisomes in prokaryotes.

vi. In many cases, a cytoplasmically inherited characteristic is associated with an endosymbiont or parasite or virus present in the cytoplasm of the organism.

In some specific cases the cytoplasmic inheritance of some preformed characteristics is not affected by exchange of cytoplasm and is not controlled by nuclear genes. They are autonomous and are maintained by cell division.

Epigenetic mechanisms of transcriptional regulation & genomic imprinting

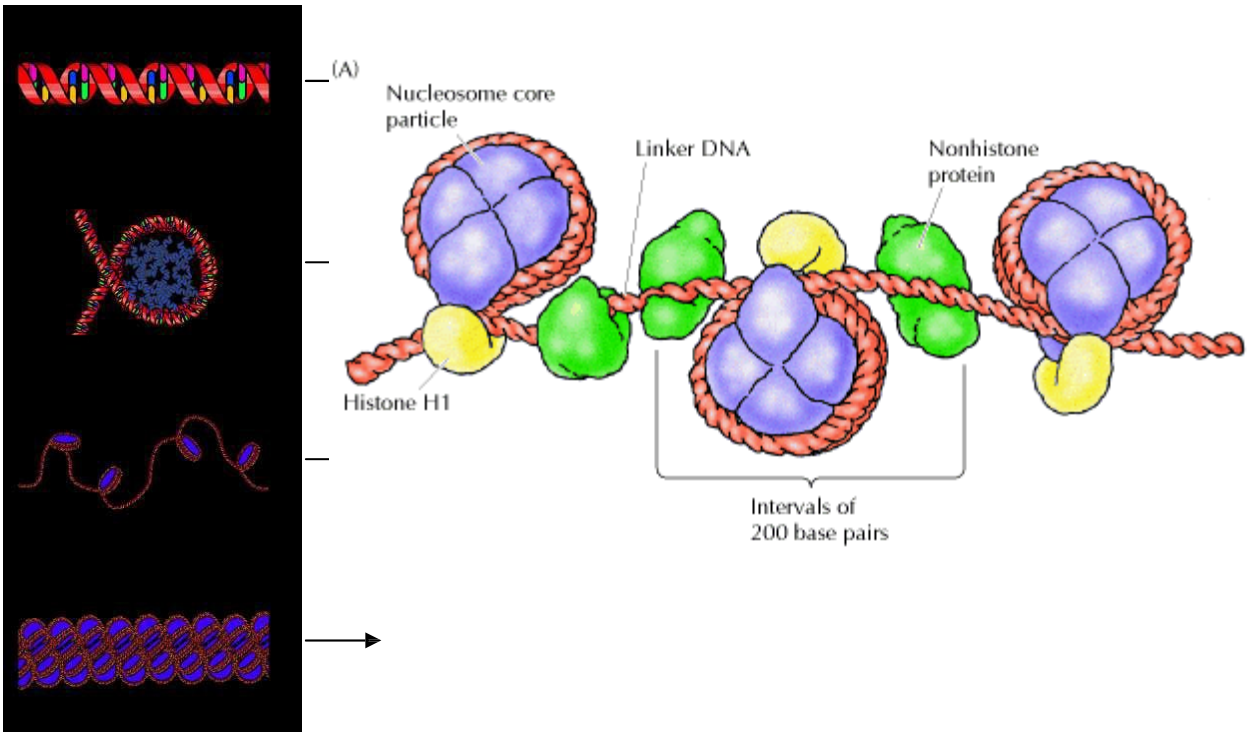
Epigenetics

Epigenetics is the study of mitotically heritable changes in gene expression that occur without changes in the DNA sequence (Wolffe & Matzke, 1999) The term 'epigenome' describes epigenetic modifications all over the Genome.

Length of extended human DNA is nearly 2 m, while the diameter of the nucleus is 5-10 µ µµ µm. DNA of eukaryotic cells is tightly bound to small basic proteins, the histones, that package DNA in an orderly manner to form chromatin.

Nucleosomes – Fundamental repeating units of eukaryotic chromatin. The nucleosome is composed of a short length of DNA (146 bp) wrapped around a core of histone

proteins in 2 turns. H2a, 2b, 3 and 4 constitute the core histones, while histone H1 helps in the packaging of the nucleosomes on each other.



Each core histone is associated with an amino-terminal tail, of 25-40 amino acid residues, that extends through the DNA into the space surrounding the nucleosome

♦ ♦♦ ♦ Thus, the coiling and supercoiling condenses the length of the DNA some 10,000-fold

♦ ♦♦ ♦ However, such a compact structure would block the transcriptional machinery from approaching its target and silence gene transcription

♦ ♦♦ ♦ To facilitate the transcription of certain genes and silence the expression of others, the DNA and the core histone tails undergo covalent modifications, which are collectively studied under the heading of 'epigenetic mechanisms of gene regulation'

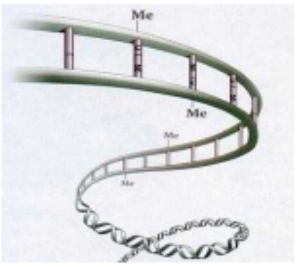
Mechanisms of epigenetic regulation,

♦ DNA methylation
♦ Histone modifications
∗ Histone acetylation
∗ Histone methylation
∗ Histone phosphorylation
∗ Histone ubiquitination
∗ Histone sumoylation
∗ Histone ribosylation

♦ Small RNA control
* siRNA
* miRNA

DNA methylation

♦ In eukaryotes, it refers to the process by which a methyl group is covalently added to the carbon (at position 5) of cytosine in the DNA strand



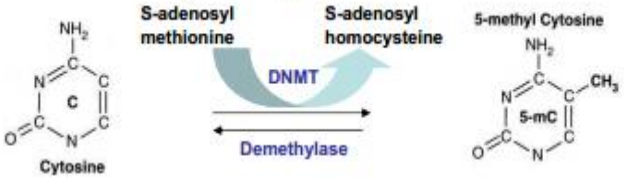
♦ Only those cytidine residues that are adjacent to guanidine i.e. the CpG sites (cytidine bound through a phosphate molecule to guanidine) in the DNA strand are targets for the methylation-inducing enzymes

♦ These CpG sites may occur in multiple repeats and are known as CpG islands

♦ The most important location for DNA methylation is in the promoter region of the gene, where extensive methylation (hypermethylation) of the CpG sites causes gene silencing

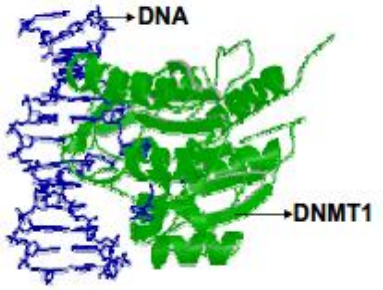
How is DNA methylated?

♦ DNA methyl transferases (DNMT)



♦ 4 DNA methyl transferases

* DNA methyl transferase 1 (DNMT 1)
* DNA methyl transferase 2 (DNMT 2)
* DNA methyl transferase 3a (DNMT 3a)
* DNA methyl transferase 3b (DNMT 3b)



Functions of the DNMTs

DNMT1: Maintenance methyltransferase, component of DNA replication complex

Functions – a. 'Copies' the methylation pattern from the template DNA strand to the newly synthesized strand after cell division

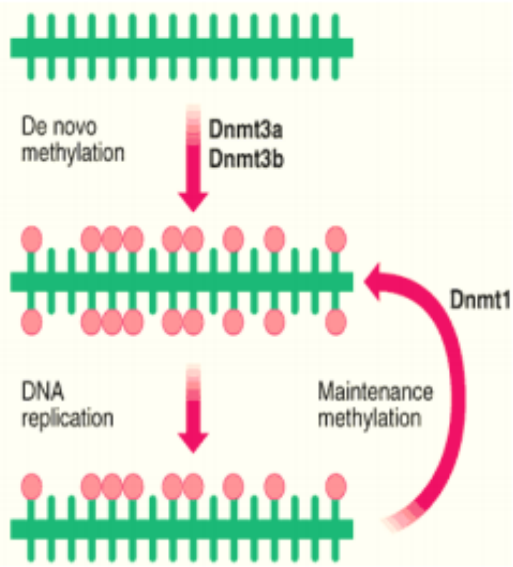
b. Supports long-term silencing of non coding DNA in addition to the epigenetic silencing of particular genes

DNMT3a & 3b:

Functions – a. Responsible for the de novo methylation of the non methylated DNA in response to environmental challenges and during development

b. Important role in epigenetic silencing of particular genes

DNMT2: Displays negligible evidence of transmethylase activity



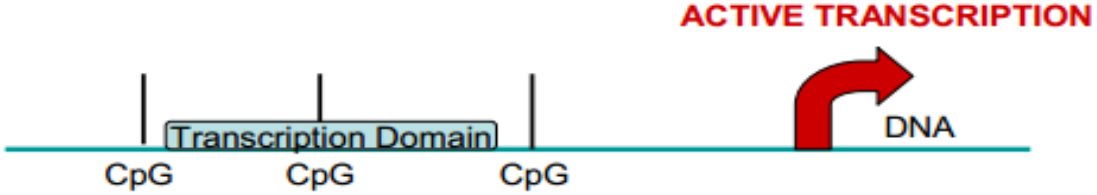
How does DNA methylation affect gene transcription?

How are genes transcribed?

- Actively transcribed genes are characterised by the presence of CpG sites in their promoter regions that are hypomethylated

Transcription Factors

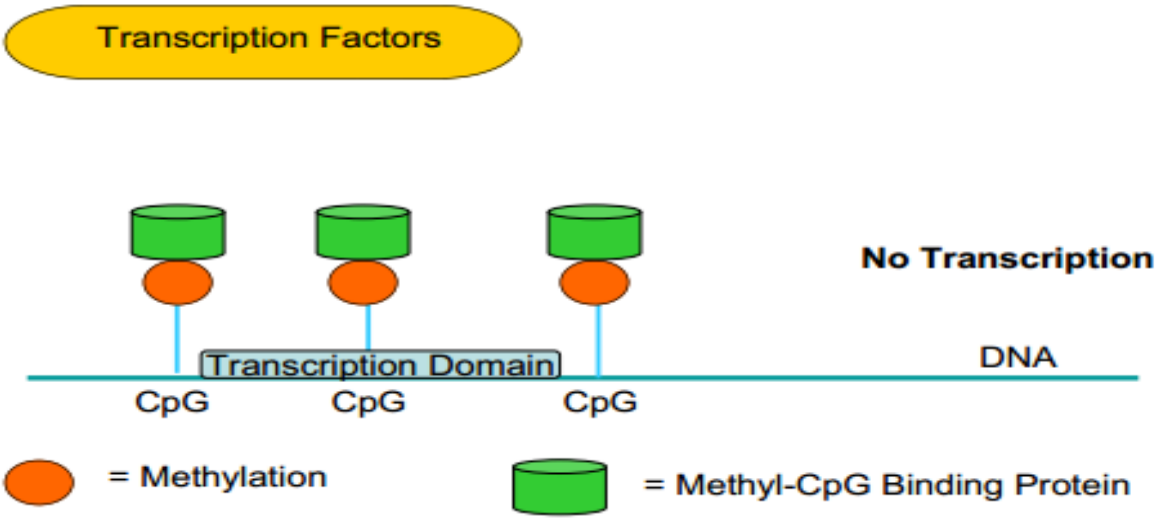
- As the CpG sites in the promoter region are hypomethylated, it is possible for appropriate transcription factors to bind to their recognition sequences in the promoter



"Simplified version of mechanism of active gene transcription"

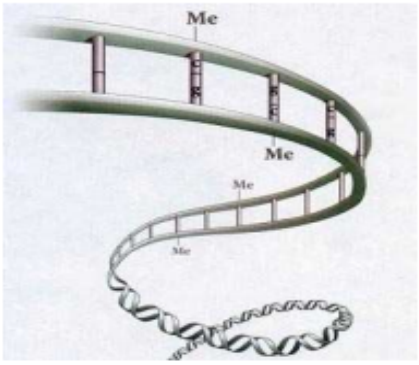
How does DNA methylation silence gene transcription?

- The nuclear matrix proteins, methyl-CpG-binding proteins 1 and 2, bind preferentially to the methylated cytosines
- Presence of the methyl-CpG-binding proteins, bound to the methylated cytosines in the promoter region, obstructs binding of appropriate transcription factors to their transcription domains



DNA demethylation

- DNA demethylation is an important component of the epigenetic control as methylated genes may need to be 'unsilenced' by demethylation in response to different environmental signals or during development



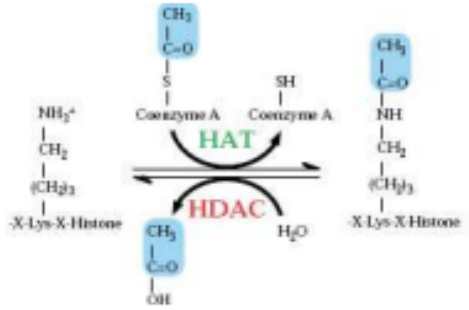
- Passive demethylation:
  - Occurs during DNA replication i.e. cell division and involves the inhibition of DNMT1 activity
- Active demethylation:
  - Glycosylase-dependent: Enzymes known to have 5-Methylcytosine-DNA-glycosylase activity cleave the bond between the DNA backbone and the methylated cytosine base (Zhu *et al.*, 2000, *Nucleic Acid Res*)
  - Direct removal of the methyl moiety from the methylated DNA (Ramchandani *et al.*, 1999, *PNAS*)

CLASS: II B.Sc., BIOCHEMISTRY
COURSE CODE: 17BCU304B

COURSE NAME: CONCEPTS IN GENETICS
UNIT:IV-DEVELOPMENTAL GENETICS, EPIGENETICS AND CHROMOSOMAL ABERRATIONS (BATCH-2017-2020)

Histone acetylation

- The enzymes histone acetyltransferases (HATs) catalyse the transfer of acetyl groups from acetyl coenzyme A to the amino groups of conserved lysine residues located in N-terminal tails of the core histones



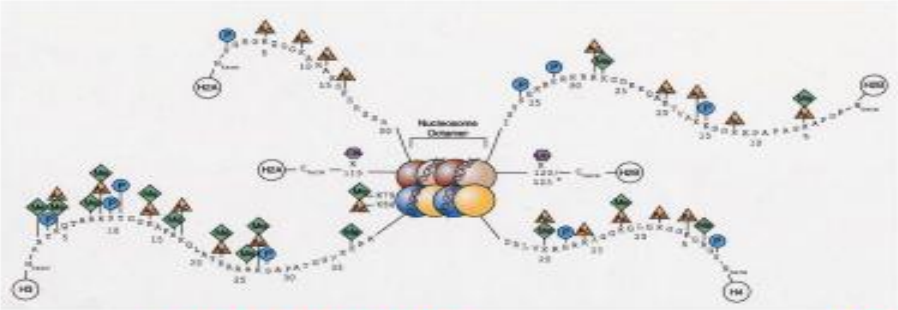
- The enzymes histone deacetylases (HDACs) catalyse the removal of acetyl groups from histones. Many HDAC inhibitors are currently involved in clinical trials as chemotherapeutic agents
- Dynamic process

Histone methylation

- The enzymes histone methyltransferases (HMTs) catalyse the transfer of methyl groups from S-adenosylmethionine (SAM) to the amino groups of conserved lysine residues located in N-terminal tails of the core histones

Note: Arginine residues in N-terminal tails of core histones can also be methylated

- Existence of a global histone demethylase seems unlikely as very little evidence exists for large-scale decreases in methylated histones from bulk chromatin
- Relatively stable, turnover of methyl groups is relatively less

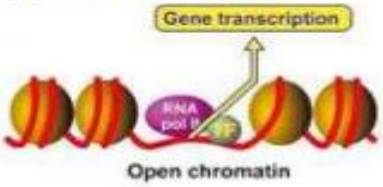


Conserved lysine residues in the N-terminal tails of H3 and H4 are the major targets for acetylation and methylation

However, lysine residues in N-terminal tails of H2a and H2b have been reported to also undergo these modifications

Histone acetylation

- Addition of acetyl groups (CH_3CO^-) to the lysine residues neutralizes the positive charge of the histone tails, decreasing their affinity for the negatively charged DNA, thereby 'opening' the chromatin, allowing access to the transcriptional machinery and facilitating gene transcription



Histone methylation

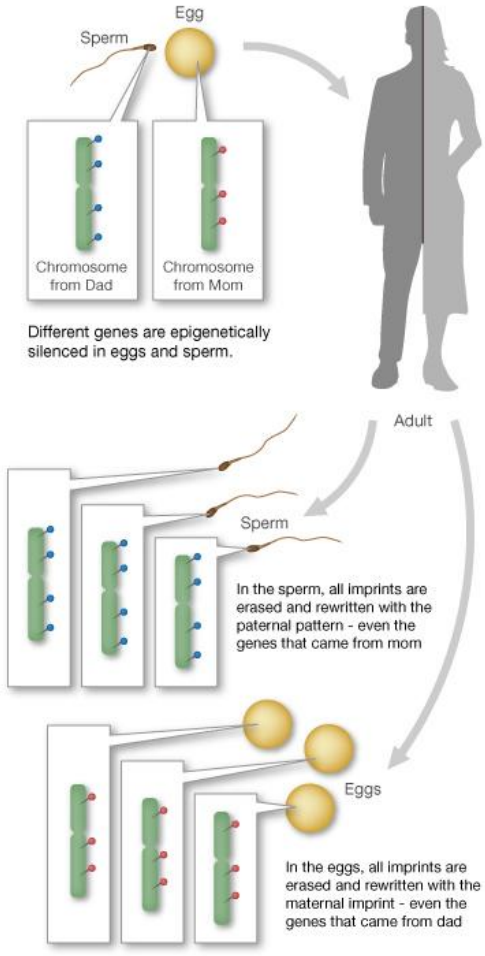
- Addition of methyl groups (CH_3+) to the lysine residues (particularly at positions 9 and 27 in H3) increases the affinity of the histone tails for anionic DNA, thereby 'closing' the chromatin, obstructing access to the transcriptional machinery and silencing gene transcription

Note: Methylation of lysine (4) and arginine (17) residues in the H3 tail = transcriptional activation

Genomic Imprinting

For most genes, we inherit two working copies -- one from mom and one from dad. But with imprinted genes, we inherit only one working copy. Depending on the gene, either the copy from mom or the copy from dad is epigenetically silenced. Silencing usually happens through the addition of methyl groups during egg or sperm formation.

The epigenetic tags on imprinted genes usually stay put for the life of the organism. But they are reset during egg and sperm formation. Regardless of whether they came from mom or dad, certain genes are always silenced in the egg, and others are always silenced in the sperm.



Imprinted Genes Bypass Epigenetic Reprogramming

Soon after egg and sperm meet, most of the epigenetic tags that activate and silence genes are stripped from the DNA. However, in mammals, imprinted genes keep their epigenetic tags. Imprinted genes begin the process of development with epigenetic tags in place.Imprinted genes are not the only genes that bypass epigenetic reprogramming in the early embryo. Studying imprinting may help researchers understand how other genes make it through reprogramming without losing their epigenetic tags. Imprinting is unique to mammals and flowering plants. In mammals, about 1% of genes are imprinted.

Imprinting is required for normal development

An individual normally has one active copy of an imprinted gene. Improper imprinting can result in an individual having two active copies or two inactive copies. This can lead to severe developmental abnormalities, cancer, and other problems.

Prader-Willi and Angelman syndrome are two very different disorders, but they are both linked to the same imprinted region of chromosome 15. Some of the genes in this region are silenced in the egg, and at least one gene is silenced in the sperm. So someone who inherits a defect on chromosome 15 is missing different active genes, depending on whether the chromosome came from mom or dad.

Prader-Willi syndrome

• Symptoms include learning difficulties, short stature, and compulsive eating.
• Individuals are missing gene activity that normally comes from dad.
• Happens when dad's copy is missing, or when there are two maternal copies.

Angelman syndrome

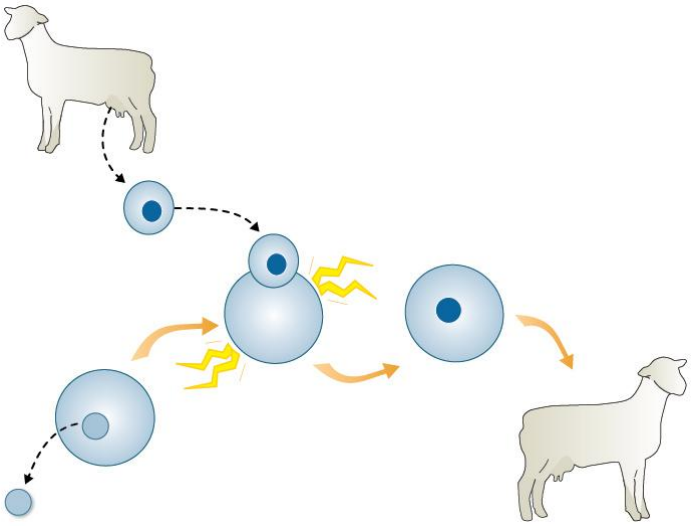
• Symptoms include learning difficulties, speech problems, seizures, jerky movements, and an unusually happy disposition.
• Individuals are missing gene activity that normally comes from mom.
• Happens when mom's copy is defective or missing, or when there are two paternal copies.

Mammals are notoriously difficult to clone. Researchers often need to go through the cloning procedure dozens or even hundreds of times in order to produce a single healthy clone. The epigenome, including problems with imprinted genes, is likely to be at the root of this difficulty.

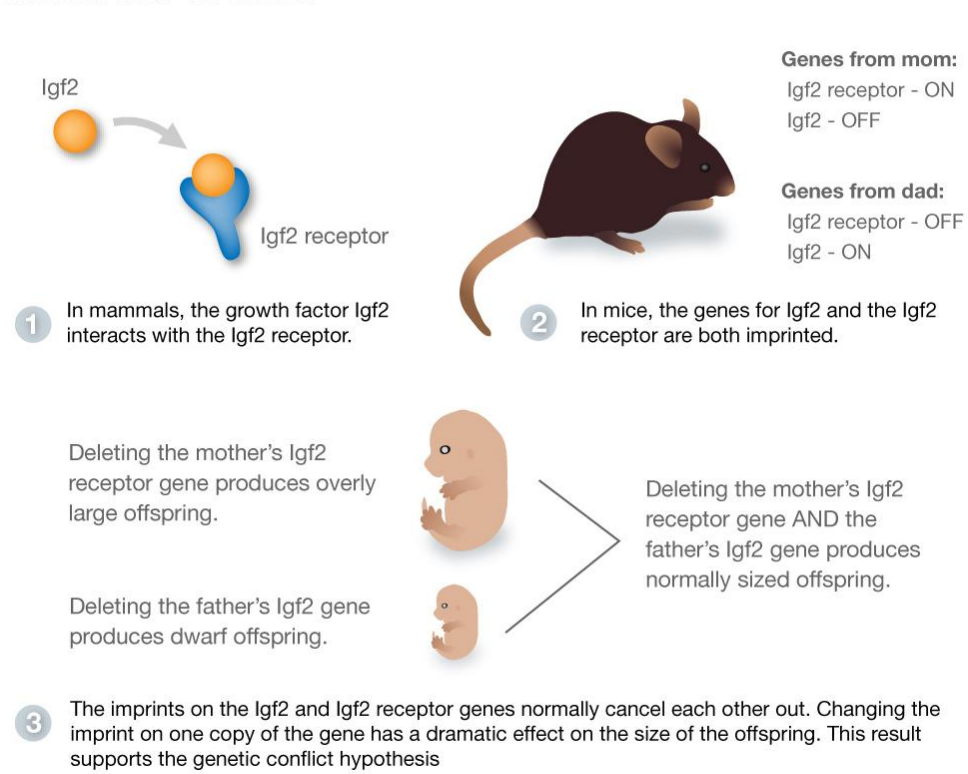
The most common method of cloning is called somatic cell nuclear transfer (SCNT). SCNT involves removing a donor nucleus from a non-reproductive cell (often a skin cell or mammary cell) and placing it into an egg cell that has had its nucleus removed.

Clones have abnormal epigenomes, which can lead to a variety of problems. The epigenetic problems with clones likely arise for two reasons. First, the donor nucleus comes from a differentiated cell with epigenetic tags already in place. These tags keep genes switched on or off and allow the cell to perform its responsibilities. After the donor nucleus is transferred, the egg does its best to erase the epigenetic tags. But the process is faulty, delayed and incomplete.

Second, the epigenetic tags in the donor nucleus have been copied several times over. While the machinery that copies the DNA code is faithful (it makes about one error in half a billion), the epigenetic copying machinery is sloppy. In some cases, its error rate can be as high as 1 in 25. Miscopied epigenetic tags on even a very small number of imprinted genes in the donor nucleus could have serious consequences during the development of the resulting embryo.



Dolly the sheep was the first mammal to be cloned by somatic cell nuclear transfer (SCNT).



Scientists have come up with a number of hypotheses to explain why imprinting happens in mammals. One of these, the Genetic Conflict hypothesis, supposes that imprinting grew out of a competition between males for maternal resources.

In some species, more than one male can father offspring from the same litter. A house cat, for example, can mate more than once during a heat and have a litter of kittens with two or more fathers. If one father's kittens grow larger than the rest, his offspring will be more likely to survive to

adulthood and pass along their genes. So it's in the interest of the father's genes to produce larger offspring. The larger kittens will be able to compete for maternal resources at the expense of the other father's kittens.

On the other hand, a better outcome for the mother's genes would be for all of her kittens to survive to adulthood and reproduce. The mother alone will provide nutrients and protection for her kittens throughout pregnancy and after birth. She needs to be able to divide her resources among several kittens, without compromising her own needs.

It turns out that many imprinted genes are involved in growth and metabolism. Paternal imprinting favors the production of larger offspring, and maternal imprinting favors smaller offspring. Often maternally and paternally imprinted genes work in the very same growth pathways. This conflict of interest sets up an epigenetic battle between the parents -- a sort of parental tug-of-war.

Beckwith-Wiedemann Syndrome

The Igf2 gene (but not the Igf2 receptor gene) is also imprinted in humans. The Igf2 gene codes for a hormone that stimulates growth during embryonic and fetal development. Methyl tags normally silence the maternal Igf2 gene. But a DNA mutation or an "epimutation" (missing methyl tags) can activate it, resulting in two active copies of the gene.

Activation of the maternal Igf2 gene during egg formation or very early in development causes Beckwith-Wiedemann Syndrome (BWS). While children with BWS have a variety of symptoms, the most common and obvious feature is overgrowth. Babies with BWS are born larger than 95% of their peers. They also have an increased risk of cancer, especially during childhood.

BWS occurs once in about 15,000 births. However, in babies that were conceived in the laboratory with the help of artificial reproductive technology (ART), the rate of BWS may be as high as 1 in 4,000. This and other evidence of imprinting errors is prompting some to call for further investigation into the safety of common ART laboratory procedures.

Ligers and Tigons

Imprinted genes are under greater selective pressure than normal genes. This is because only one copy is active at a time. Any variations in that copy will be expressed. There is no "back-up copy" to mask its effects. As a result, imprinted genes evolve more rapidly than other genes. And imprinting patterns -- which genes are silenced in the eggs and sperm -- also evolve quickly. They can be quite different in closely related species.

Lions and tigers don't normally meet in nature. But they can get along very well in captivity, where they sometimes produce hybrid offspring. The offspring look different, depending on who the mother is. A male lion and a female tiger produce a liger - the biggest of the big cats. A male tiger and a female lion produce a tigon, a cat that is about the same size as its parents.

The difference in size and appearance between ligers and tigons is due in part to the parents' differently imprinted genes. Other animals can also hybridize, with similar results. For example, a horse and a donkey can produce a mule or a hinny.

Variations in chromosome number and structure

- **Chromosomal mutation** = variations from the normal (wild-type) condition in chromosome structure and number.
  - In all organisms, can arise spontaneously or be induced experimentally by radiation or certain chemicals.
  - Detected usually by genetic analysis; results in changes in usual linkage arrangement of genes.

- Can also be detected under microscope during mitosis and meiosis.
- Chromosomal mutations are very common:
- Responsible for more than half of all spontaneous abortions.
- 0.6 % of live individuals have some type of chromosomal mutation.
- 11% of men with infertility have chromosomal mutations.
- 6% of institutionalized individuals with mental deficiencies have chromosomal mutations.
- Involved in some cancers.

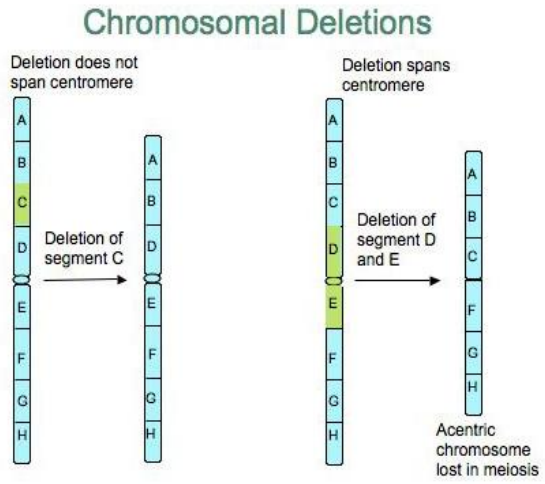
Variations in chromosome structure

- Four common types of chromosomal rearrangements are: deletions, duplications, inversions, and translocations.
- Chromosomal rearrangements begin with chromosomal breakage.
- if break occurs at a gene, function may be lost.
- broken ends of chromosomes are "sticky", and can adhere to other broken chromosome ends.
- **telomeres** are sequences at chromosome ends. Without telomeres, chromosome ends become degraded.
- **Polytene chromosomes** = special kinds of chromosomes found in cells of salivary glands of some insects (order *Diptera*). Consist of chromatid bundles resulting from repeated cycles of chromosome duplication without nuclear division (Fig 16.1). Can be 1000 times thicker than normal meiotic chromosomes.
- Their study has contributed much to our knowledge of chromosomal structure because chromosomal abnormalities are easily detected under light microscope.
- In each polytene chromosome, homologous chromosomes are tightly paired.
- Polytene chromosomes are joined together at their centromeres by a proteinaceous structure called a **chromocenter**.
- When stained, polytene chromosomes show a banding pattern. Each band contains an average of 30,000 bp, and may contain up to seven genes. Interbands also contain genes.
- In *Drosophila*, 5,000 bands have been observed.
- Learn more about polytene chromosomes.

Deletion

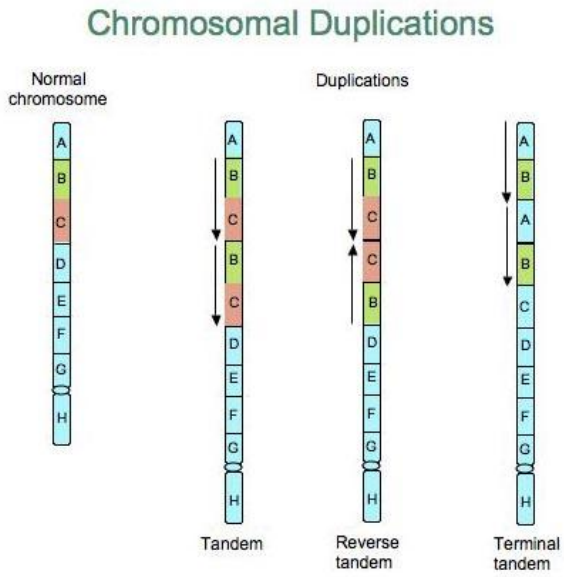
- **Deletion** = chromosomal mutation in which part of a chromosome is missing (Fig 16.2).
- Deletion starts at chromosomal breakage.
- Can be induced by heat, ionizing radiation, viruses, chemicals and transposable elements.

Deletion mutations do not revert to wildtype.
Consequences of a deletion depends on the particular DNA deleted.
Individuals heterozygous for a deletion may be normal. BUT if the homolog contains recessive genes with deleterious effects, consequences may be drastic.
If deletion involves loss of a centromere, it results in an **acentric** chromosome. Acentric chromosomes are lost during meiosis.
In most cases, deletion of an entire chromosome is lethal.
There are no living humans with an entire autosomal chromosome missing.



- In some organisms, deletions can be detected through karyotyping.
  - Heterozygotes will have mismatched homologous pair.
  - During meiosis, unmatched DNA can be seen looping out during synapsis.
- **Pseudodominance** = unexpected appearance of recessive phenotype caused by dominant allele being deleted.
- Deletions can be used to physically map genes to chromosomes. **Deletion mapping** is commonly done with *Drosophila* polytene chromosomes (Fig 16.3).
  - *Example*: in *Drosophila*, genes *y, ac*, and *sc* are located on left end of X chromosome. *Drosophila*strains that carried deletions in this region were analyzed: strain 260-1 was pseudodominant for *y*, *ac*, and *sc*, and had bands A1-7 and B1-4 deleted. Stain 260-2 was pseudodominant for *y* and *ac*, and had bands A1-7 and B1 deleted. Since *sc* was lost in strain 260-1 but not in strain 260-2, *sc* must be located in bands B2-4.
  - Deletion mapping analysis has been used to construct a detailed map of *Drosophila* polytene chromosomes.
- Some human disorders are caused by deletion of chromosome segments. If deletions are very large, it's almost always lethal. Many genes can be missing.
  - ***Cri-du-chat* syndrome** (Fig 16.4)
    - Heterozygous for deletion of part of short arm of chromosome 5.
    - Causes severe mental retardation, many physical abnormalities, and their cry sounds like the mew of a cat.
    - Affects 1 in 50,000 live births.
    - More on Cri-du-chat-syndrome
  - ***Prader-Willi* syndrome**
    - Heterozygous for deletion on long arm of chromosome 15.
    - Frequency estimated at 1 in 10,000-25,000 people (mostly males).
    - Infants are weak due to poor sucking reflex. By age 5-6, children become compulsive eaters. Eventually develop obesity and other related health problems. Left untreated, individuals can feed themselves to death. Other symptoms include mental retardation, poor sexual development, and behavior problems.

- More on Prader-willi-syndrome

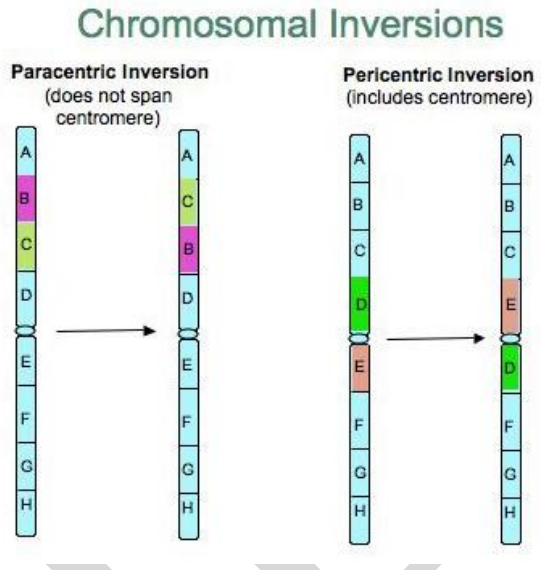


Duplication = chromosomal mutation that results in a doubling of a segment of a chromosome (Fig 16.5). Size of duplication varies considerably and may occur at different locations in the genome or in a tandem configuration. Tandem duplications can be:
Reverse tandem duplication: order of genes in duplicated segment is opposite of the original.
Terminal tandem duplication: duplicated segment located at end of chromosome

- Heterozygous duplications result in unpaired loops in prophase I and can be detected cytologically.
- The *Bar* mutant in *Drosophila* results from a duplicated region on the X chromosome (Fig 16.6). Flies homozygous for *Bar* eye have less eye facets than normal eyes, giving them a slit-like appearance, rather than being oval.
- Duplications have played an important role in evolution. Duplicated genes can diverge and aquire new functions. Most genes have evolved from gene duplication events with subsequent diversification to give rise to **multigene families**. Most genes belong to multigene families (e.g. actin, globin, collagen, etc…). Members of a family may perform slightly different functions. In humans, there are over 35 different collagen genes, each is suited for different purposes (i.e. making bone, cartilage, tendons, ligaments, etc…)

Inversion

- **Inversion** = a chromosomal mutation that arises when a segment of a chromosome is excised and then reintegrated at an orientation 180 degrees from the original.

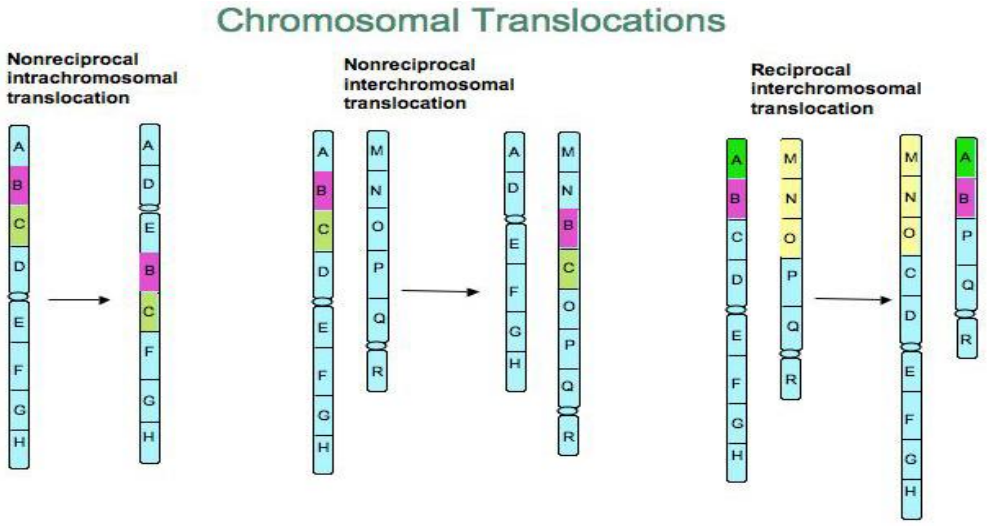


Two types of inversions:
Paracentric inversion: does not include the centromere.
Pericentric inversion: includes the centromere.
Does not involve loss of genetic material, but there can be phenotypic effects resulting if breakpoints are within a gene or its regulatory region.
Homozygous inversions can be detected because the linkage arrangement of genes can be altered (e.g. A**BCD**EFG can become A***DCB***EFG).

- Consequences of chromosome inversions:
  - If homozygous for inversion, meiosis is normal. No problems.
  - If heterozygous for inversion, AND crossover is absent in the inverted region, meiosis is normal. No problems.
  - If heterozygous for inversion, and there is a crossover in inverted region, then there are serious genetic consequences. See below
- ***Paracentric inversion heterozygote with a crossover*** (Fig 16.8):
  - Homologous chromosomes try to pair up as best they can. Requires formation of an **inversion loop**.
  - Crossover within inverted region causes formation of a **dicentric bridge** (results when the two centromeres of resultant chromosome migrate to opposite poles during anaphase I). This dicentric chromosome breaks apart. Crossover also yields acentric chromosome, which is lost .
  - After meiosis II, two of the gametes receive complete set of chromosomes. These are viable. The other two gametes receive DNA with many missing genes. These are not viable.
  - *The only gametes that can give rise to viable progeny are those containing the chromosomes that did not involve a crossover.* Thus, for paracentric inversion heterozygotes, viable recombinants are reduced or totally suppressed.
- ***Pericentric inversion heterozygote with a crossover*** (Fig 16.9):
  - Homologues also pair by forming inversion loop.
  - Results in two viable gametes with nonrecombinant chromosomes and two recombinant gametes that are inviable, each as a result of deletion of some genes and the duplication of other genes. No dicentric bridge formed nor acentric chromosomes.
- NOTE: if two crossovers occur within an inversion loop, all four gametes are viable.

Translocation

- **Translocation** = chromosomal mutation in which there is a change in position of chromosomal segments to a different location in the genome.
- Genetic material is neither gained nor lost.



- Translocations are of three types (Fig 16.10):
  - **Nonreciprocal intrachromosomal translocation**: chromosomal segment changes position within the same chromosome.
  - **Nonreciprocal interchromosomal translocation**: chromosomal segment changes position from one chromosome to another.
  - **Reciprocal interchromosomal translocation**: exchange of segments between two chromosomes.
- Organisms homozygous for a translocation have altered linkage arrangements of genes (some genes are now farther apart, others become closer).
- Translocations typically affect the products of meiosis.
  - In some cases, gametes are produced with either duplications or deletions and consequently are in viable.
  - In other cases, duplications steming from translocation can yield viable gametes (e.g. familial Down syndrome).
- We will focus on reciprocal translocations.
  - In homozygotes for reciprocal translocation, meiosis is normal and crossovers do not produce abnormal chromatids.
  - In heterozygotes for a reciprocal translocation, chromosomes pair up as best they can. Results in a cross-like configuration in prophase I. This cross-like configuration is made up of 4 paired-up chromosomes, each partially homologous to the others (Fig 16.11).
- Segregation at anaphase I occurs in 3 possible ways (ignoring crossovers):
- ***Alternate segregation***: alternate centromeres migrate to same pole (i.e. N1 and N2 to one pole, and T1 and T2 to the other; produces viable gametes because no information is deleted nor duplicated. Half the gametes have two normal chromosomes, the other half have two translocated chromosomes.
- ***Adjacent 1 segregation***: adjacent nonhomologous centromeres migrate to same pole (i.e. N1, T2 to one pole, and N2, T1 to the other; both gametes contain deletions and duplications and are *usually* inviable. Occurs as frequently as alternate segregation.
- ***Adjacent 2 segregation***: adjacent homologous centromeres migrate to same pole (i.e. N1, T1 to one pole, and N2, T2 to the other). Seldom occurs and always result in inviable gametes.

- Individuals heterozygous for translocations are semisterile, because only about half of the gametes produced are viable.
- In practice, although many animal gametes with duplications and deletions may be viable, the resultant zygotes are not. In plants, pollen grains with deletions or duplication do notdevelop and are nonfunctional.

Chromosomal mutations and Human tumors

- Most human tumors have chromosomal mutations (translocations most common). In many tumors, there is no specific associated chromosomal mutation. Other tumors are associated with a specific chromosomal mutation.
- **Chronic myelogenous leukemia (CML)**:
- Cancer involving the uncontrolled proleferation of myeloblasts (stem cells of white blood cells).
- 90% of CML patients have a **Philadelphia chromosome**, which results from a reciprocal translocation between chromosome 22 and 9 (Fig 16.12). The translocation converts a proto-oncogene to an **oncogene**.
- Specifically, *ABL* gene on chromosome 22 becomes fused to *BCR* gene on chromosome 9. The fusion protein causes cells to proliferate.
- *Gleevec* is a recently developed drug that targets the fusion protein and shows promise in treatment of this particular kind of cancer. This is a good example of targeted drug development (a.k.a rational drug design).
- **Burkitt Lymphoma**
- A viral-induced tumor that affects the B cells of the immune system (common in Africa).
- 90% of tumors in Burkitt lymphoma are associated with a reciprocal translocation between chromosome 8 and 14.
- Translocation brings *MYC* gene next to the regulatory region of an antibody gene, resulting in the overexpression of the *MYC* gene (an oncogene).

Position effect

- Inversions and translocations do not usually produce mutant phenotypes unless chromosomal breakpoints occur within a gene. Their effects are usually limited to meiosis in heterozygotes. In some cases, however, the inversion or translocation can change the position of a gene in the genome such that its expression is altered. This is known as a **position effect**.
- Position effect can occur if a gene that is normally in **euchromatin** (condensed during cell division , but uncoiled during interphase; transcriptionally active region) is moved to **heterochromatin** (stay condensesd during interphase; transcriptionally inactive).
- Some human genetic diseases are associated with position effects:
- **Aniridia**: a cogenital eye condition characterized by severe underdevelopment of the iris. Caused by lack of function of the *PAX6* gene, which is necessary for eye development. In some cases, expression of *PAX6* is supressed by a position effect resulting from a translocation.

Fragile sites and Fragile X syndrome

- Some human chromosomes develop narrowing or unstained areas (gaps) called **fragile sites** when grown in culture. Chromosomes may break at these fragile sites, causing deletions. More than 40 fragile sites have been identified.
- One particular fragile site on the X chromosome (Xq27.3) is associated with fragile X syndrome.
- **Fragile X syndrome** (Fig 16.13 and 14):

- Second leading genetic cause of mental retardation (1 in 1250 males and 1 in 2500 females).
- Fragile X chromosome is inherited in a Mendelian manner. However, only 80% of males with fragile X chromosome are mentally retarded. These phenotypically normal males are called *normal transmitting males*, and can pass on the condition to their daughters.
- Molecular basis of fragile X is somewhat understood:
- A gene at the fragile site called *FMR-1* contains an average of 29 CGG repeats.
- Individuals with fragile X have 200-1300 copies of this repeat. Amplification of the number of repeats is a result of slipped mispairing during DNA replication.
- Amplification of the number of these repeats changes the expression level of the *FMR-1* gene, which causes the mental retardation. The exact role of FMR-1 is unknown, althought it appears to be an mRNA binding protein.
- Triplet repeat amplification is associated with other human diseases, although it is not associated with a fragile site. However, like fragile X, there is a threshold number of reapeats necessary to cause disease. Examples include:
- myotonic dystrophy
- spinobulbar muscular atrophy (a.k.a Kennedy disease)
- Huntington's disease.

Variations in chromosome number

- **Euploid**s are cells or organisms with one complete set of chromosomes or an exact multiple of complete sets of chromosomes.
- **Aneuploids** are organisms with chromosome numbers that are not exact multiples of the haploid set of chromosomes.
- We now look at mutations that affect euploidy and aneuploidy.

Changes in one or a few chromosomes

- Generation of aneuploidy:
- Cells with abnormal numbers of chromosomes are typically generated from nondisjunction occurring either at meiosis I or II.
- **Nondisjunction** = failure of chromosomes to separate during meiosis.
- Types of aneuploidy (Fig 16.15)
- In aneuploidy, cells have more or less chromosomes than normal.
- In animals, aneuploidy is mostly lethal, so its only detected in aborted fetuses. Plants are more tolerant of aneuploidy.
- In diploids, there are 4 types of aneuploidy:
- **Nullisomy**: loss of one homologous chromosome pair (2N – 2). Results from nondisjunction for the same chromosome during meiosis in both parents.
- **Monosomy**: Loss of a single chromosome (2N – 1).
- **Trisomy**: single extra chromosome (2N + 1).
- **Tetrasomy**: extra chromosome pair (2N + 2).
- Aneuploidy may involve a loss or gain of more than one type of chromosome. Thus, can get *double monosomic* and *double tetrasomics*.
- Aneuploidy causes serious problems during meiosis:
- Monosomics produce N as well as N+1 gametes.
- Trisomics produce 4 types of gametes (Fig 16.16).

- Aneuploidy involving sex chromosomes is more tolerated than aneuploidy involving autosomes because of existing dosage compensation mechanisms for sex chromosomes. Table 16.1 lists some aneuploid abnormalities in the human population.
- In humans, autosomal monosomy is extremely rare. Embryos do not develop significantly and are lost in early pregnancy. Autosomal trisomy is more common and account for one half of all chromosomal abnormalities producing fetal deaths. Very few autosomal live births. Most of these are trisomy 8, -13, -18. Only trisomy 21 individuals survive into adulthood.
- **Trisomy 21** (Down Syndrome) (Fig 16.17)
- Frequency of 3.5/1000 conceptions. 1.4/1000 live births.
- Characterized by low IQ, epicanthal folds, short broad hands, short stature.
- Correlated with maternal age (Table 16.2). Probability of nondisjunction increases with the length of time primary oocyte is in the ovary.
- Mothers over age 35 are encouraged to undergo amniocentesis or chorionic villus sampling.
- There is a very small correlation with paternal age if mother is over 35.
- Down syndrome can also be caused by a **Robertsonian translocation** (Fig 16.18) (A.k.a. **familial Down syndrome**) involving chromosomes 14 and 21.
- Produces 3 copies of long arm of chromosome 21.
- When carrier parents mate with normal parents, 1/3 of viable zygotes produced give rise to familial Down syndrome, a much higher risk than nonfamilial.
- **Trisomy 13 (Patau syndrome)**
- 2/10,000 live births.
- Characteristics: cleft lip and palate, small eyes, extra fingers and toes, mental and developmental retardation, cardiac anomalies, and much more. Usually die within 3 months.
- **Trisomy 18 (Edwards syndrome)**
- 2.5/10,000 live births.
- For reasons unknown, 80% are female.
- Characteristics: small at birth, multiple congenital malformations affecting almost every organ, clenched fists, elongated skull, low-set malformed ears, mental and developmental retardation. 90% die within 6 months of cardiac problems.

Changes in complete sets of chromosomes

- Monoploidy and polyploidy (type of euploidy) involve variations in normal state in the number of complete set of chromosomes.
- **Monoploidy**: only one set of chromosomes.
- **Polyploidy**: more than the normal set of chromosomes:
- **Triploidy: 3N**
- **Tetraploidy: 4N**
- Lethal in most species, but less consequential in plants. In fact, euploidy has played important role in plant evolution.
- Euploidy occurs when either Meiosis I or II is aborted or when nondisjunction occurs for all chromosomes.
- Monoploidy usually results from unfertilized eggs.

Monoploidy:

- Found in haploid organisms (fungi, algae etc…)
- Rarely observed in adult diploid organisms because of recessive lethal genes.

- In certain species, it's a normal part of life cycle (e.g. some male bees, wasps, and ants)
- Very useful for isolating mutants because there is only dose for each gene.

Polyploidy:

- Arise spontaneously or experimentally induced.
- Often occur as a result in breakdown of spindle apparatus in meiosis or mitosis.
- Almost all plants and animals have some polyploidy tissues:
- plant endospem tissue is triploid.
- liver of mammals is polyploid.
- abdominal neuron of *Aplysia* has about 75,000 copies of the genome.
- Polyploid plants include: wheat (6N), strawberry (8N).
- Polyploid animals: N American sucker (fish), salmon, and some salamanders.
- Polyploids can either have an *even* number or *odd* number of chromosome sets.
- *Even number*: better chance of being partially fertile because homologues can pair up.
- *Odd number*: usually sterile because difficult for homologues to pair up and segregate. In triploids, the probability of producing a haploid gamete is $(1/2)^n$, where n is the number of haploid chromosomes.
- In humans, triploidy is most common type of polyploidy.
- Triploidy seen in 15-20% of spontaneous abortions and 1/10,000 live births. Always fatal.
- Tetraploidy in humans seen in 5% of spontaneous abortions. Live births very rare and always fatal.
- Polyploidy in plants is less consequential. In plants that self-fertilize, when even number polyploidy arise, fertile gametes can still be produced.
- Two types of polyploidy seen in plants:
- **Autopolyploidy**: all sets of chromosomes originate in same species.
- Results from defects in meiosis, producing diploid or triploid gametes.
- Bananas are triploid; seedless because seeds not fertile (problems with meiosis).
- Seedless fruits result from odd-number polyploidy (e.g. watermelons, grapes)
- **Allopolyploidy**: sets of chromosomes come from different (but usually related) species.
- Results from interbreeding of two different species to produce a diploid with two different sets of chromosomes. Usually sterile, but sometimes, each set of chromosomes doubles (ie cell contains diploid set of chromosomes from each parent, thus allowing pairing at meiosis), and meiosis results in viable gametes. Fusion of these gametes produced allotetraploid.
- Bread wheat is an allohexaploid (42 chromosomes).
- Derived from three distinct species, each with a diploid set of 14 chrosomomes.
- Many agricultural and horticultural plants are allopolyploids.

CLASS: II B.Sc., BIOCHEMISTRY **COURSE NAME: CONCEPTS IN GENETICS**
COURSE CODE: 17BCU304B **UNIT:IV-DEVELOPMENTAL GENETICS, EPIGENETICS AND CHROMOSOMAL ABERRATIONS (BATCH-2017-2020)**

POSSIBLE QUESTIONS

2 marks

1. Define morphogen
2. Define sex chromosomes.
3. Define epigenesis.
4. Explain on zygotic gene activity in development.
5. Name any two characteristics of drosophila as model organism.
6. Write any two points about mapping gene.

8 Marks

1. Elaborate in detail about dosage compensation of X-linked gene.
2. Comment on variations in chromosomal structure in detail.
3. Give a detail account on sex chromosomes and its determination.
4. Comment on variations in chromosomal number in detail.
5. Comment on morphogens and zygote gene activity in development.
6. Explain in detail about epigenetic mechanisms of transcriptional regulation.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
II-B.Sc., BIOCHEMISTRY
CONCEPTS IN GENETICS (17BCU304B)
MULTIPLE CHOICE QUESTIONS
UNIT IV

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	XY set of chromosomes are termed as	allosomes	autosomes	anosomes	hetrosomes	**allosomes**
2	The chromosomal arrangement in	3AAXX, 3AAXY	4AAXX, 4AAXY	22AAXX, 22AAXY	23AAXX, 23AAXY	**3AAXX, 3AAXY**
3	In marine worm, Bonellia, the sex is determined with the help of	Hormones	Metabolism	Parasite	Environmental factors	**Environmental factors**
4	In Drosophila, the genic balance mechanism was studied by	Morgon	Bridges	Punnet	Briggs	**Bridges**
5	The relative strength of X chromosomes for femaleness is	0.5	0.75	1	1.5	**1.5**
6	In man, the female are said to be	Hologametic	Homogametic	Heterogametic	Hypogametic	**Homogametic**
7	In man, each diploid body cell contains	42 chromosomes	44 chromosomes	46 chromosomes	48 chromosomes	**46 chromosomes**
8	In man, the chromosomal condition of 44AA + (XX) refers to the	Males	Females	Super males	Super females	**Females**
9	Which of the following type of sex determination occurs in man	XX - XO	XY - XO	XX - XY	XXX - XY	**XX - XY**
10	Which of the following sex is favoured by genic balance as 2 : 15	Maleness	Femaleness	Inter sex	Super sex	**Maleness**
11	In Bonellia, a marine worm, the male develops from a larva that	settles in sea water	attaches to the uterus of the female	settles in the proboscis of a female	secretes a chemical substance	**settles in the proboscis of a female**
12	The ratio of male and female relative strength in a normal male is	02:00.5	02:01.0	02:01.5	0.2:2.0	**02:01.5**
13	In man, chromosomal condition of male is	44AA + XO	44AA +XY	44 + XX	44AA + XXY	**44AA +XY**

14	The organism in which sex is determined by environmental factors is	Bonellia	Drosophila	Man	Crab	**Bonellia**
15	The number of autosomes in man are	40	42	44	46	**44**
16	Which chromosomal pair determines sex in Drosophila	I pair	II pair	III pair	IV pair	**IV pair**
17	In which of the following, similar type of sex determination occurs	Man and birds	Birds and insects	Drosophila and honey bee	Man and Drosophila	**Man and Drosophila**
18	The genic balance mechanism was first studied in	Man	Drosophila	Honey bee	Mice	**Drosophila**
19	In Bonellia, the larvae reared in sea water along with female worms develop into	Males	Females	Super males	Super females	**Males**
20	Which chromosomes has no role in the sex determination in Drosophila	x	y	Autosome	Both 1 and 2	**Y**
21	In Drosophila, sex is determined by the radio of	X-chromosome to autosome	Y-chromosome to autosome	Autosome to X-chromosome	Autosome to Y-chromosome	**X-chromosome to autosome**
22	Who found that the larvae reared from other Bonelia worm develop into female worms	Briggs	Bridges	Baltzer	Balinsky	**Baltzer**
23	Which chromosome has no sex influencing genes in Drosophila	x	y	Autosome	Both 1 and 2	**Y**
24	The chromosomal condition of human diploid cell is	44 autosomes + 2 sex chromosomes	42 autosomes + 4 sex chromosomes	43 autosomes + 3 sex chromosomes	41 autosomes + 5 sex chromosomes	**44 autosomes + 2 sex chromosomes**
25	The haemophilia diseases was first reported by	Briggs	Wilson	Baltzer	John cotto	**John cotto**
26	Which of the following is true about the inheritance of haemophilia and colour blindness	Father → F_1 son→ F_2 grand-son	Father → F_1 daughter→ F_2 grand-son	Mother → F_1 son→ F_2 grand-daughter	Both 2 and 3	**Both 2 and 3**
27	Which of the following shows criss-cross pattern of inheritance	Y-linked dominant gene	Y-linked recessive gene	X-linked recessive gene	X-linked dominant gene	**X-linked recessive gene**
28	Achromatopsia in man in otherwise known as	Haemophila	Muscular dystrophy	Colour blindness	Hydrocephalus	**Colour blindness**
29	The genotype of XX-cc refers to the	Colour blind daughter	Colour blind son	Colour blind grand-son	Colour blind grand-daughter	**Colour blind grand-daughter**
30	The experiments of sex linked genes were made by	Morgan in 1910	Morgan in 1920	Briggs in 1910	Briggs in 1920	**Morgan in 1910**
31	Persons suffering from protonopia cannot distinguish the	Green colour	Red colour	Yellow colour	Brown colour	**Red colour**
32	Who discovered the colour blindness in man	Briggs	Thomson	Wilson	Nelson	**Wilson**

33	The marriage between normal woman and colour blind man in F_2 generation produce	All normal grand-sons	All colour blind grand-sons	50% colour blind grand-sons and 50% normal grand-sons	50% normal grand-daughters and 50% colour blind grand-daughters	**50% colour blind grand-sons and 50% normal grand-sons**
34	Which are the holandric genes	'X' linked genes	'Y' linked genes	'XY' linked genes	'XX' linked genes	**'Y' linked genes**
35	How many linkage groups will be present in the case of human beings	23	46	44	22	**23**
36	In Drosophila, the pattern is established:	During production of the oocyte.	As a result of environmental interactions with the oocyte.	During gastrulation.	none of these	**During production of the oocyte.**
37	A morphogen:	Conveys positional information to cells within the embryo.	Is typically present in similar amounts in all cells.	Functions only at very low concentrations.	All of these	**Conveys positional information to cells within the embryo.**
38	Differential expression of same gene depending on parent of origin is referred to as:	Genomic imprinting	Mosaicism	Anticipation	Nonpenetrance	**Genomic imprinting**
39	The random abnormal number of chromosomes in the animals is called:	Polyploidy	Euploid	aneuploid	None	**aneuploid**
40	In Drosophila development, an embryonic segment:	Is the same thing as a parasegment.	Will contribute to two different segments in the adult fly.	Is a region of spatial control of gene expression.	None of the above.	**None of the above.**
41	Which of these would be an example of a homeotic phenotype?	The wings are shrunken and useless.	The eyes of the fly are brown instead of the normal red.	The first abdominal segment has legs.	None of the above	**The eyes of the fly are brown instead of the normal red.**
42	The process underlying differences in expression of a gene according to which parent has transmitted is called:	Anticipation	Mosaicism	Non penetrance	Genomic imprinting	**Genomic imprinting**
43	The interchange of parts between non-homologous chromosomes is called:	Duplication	translocation	Inversion	Deletion	**translocation**
44	Which of the following events is the primary mechanism used for sex determination in Drosophila?	Male-determining genes on the Y chromosome function as transcriptional activators of genes that control male development.	A cascade of transcription factor activation specifies development in the male pathway; lack of activation of this cascade will result in a female.	Alternative splicing of specific gene products leads to development of a female if two X chromosomes are present.	All of these.	**Alternative splicing of specific gene products leads to development of a female if two X chromosomes are present.**
45	Which of these mechanisms is common to sex determination in *Drosophila*	Transcriptional regulation	Alternative splicing	Environmental influences.	Y linked genes that determine maleness	**Alternative splicing**
46	In which of these species is an X chromosome required to produce a normal male? I. Drosophila II. C. elegans III. Mus musculus (mouse)	I only	I and II	I, II and III.	III only	**III only**
47	In which mode of inheritance do you expect more maternal influence among the offspring?	X-l inked	autosomal	cytoplasmic	Y-linked	**cytoplasmic**

48	Frameshift mutation is caused due to	Duplication	translocation	Inversion	Deletion	**Deletion**
49	Which, if any, of the following is not regularly an epigenetic phenomenon that depends on DNA methylation or chromatin modification?	X-chromosome inactivation.	A position effect in which a gene is silenced by an inversion where both breakpoints occur within a euchromatic environment.	Establishment of heterochromatin at a centromere.	Imprinting	**A position effect in which a gene is silenced by an inversion where both breakpoints occur within a euchromatic environment.**
50	In humans, males have set of chromosome	XX	XY	YX	YY	**XY**

UNIT 5
SYLLABUS

Complex traits inheritance, population & evolutionary genetics

Inheritance of complex trait, analysis of quantitative traits, narrow and broad sense heritability, quantitative trait loci (QTL) and their identification. Hardy- Weinberg law, predicting allele and genotype frequencies and exceptions to Hardy-Weinberg principle. Molecular evolution - analysis of nucleotide and amino acid sequences, molecular phylogenies, homologous sequences, phenotypic evolution and speciation.

Inheritance of complex trait

Mendelian Traits that are also Affected by the Environment. The effects of these mutations are usually only apparent at high temperatures. Siamese cats have a mutation in the C gene controlling dark pigment formation. The ch allele of this gene is heat sensitive. The ch allele can make dark pigment at low but not high temperatures. The permissive temperature for dark-color occurs at the extremities, and a restrictive temperature occurs in the body core.

Nutritional effects

Phenylketonuria is a human nutritional defect that can lead to severe physical and mental disorders in children, but only if they consume phenylalanine. The mutation prevents individuals from metabolizing this amino acid. The disease phenotype can be avoided by eliminating phenylalanine from the diet.

Most phenotypic traits in plants and animals are affected by many genes (size, weight, shape, lifespan, physiological traits, fecundity). Often, it is not feasible to determine the number of genes affecting a particular trait, and the individual effects of genes on the phenotype. Many of these traits can be measured on a quantitative, rather than a qualitative, scale. This is where the terms quantitative trait and quantitative genetics come from.

Analysis of Quantitative traits

1. Have continuous distributions, not discrete classes
2. Are usually affected by many genes (polygenic)
3. Are also affected by environmental factors

EXAMPLE: COLOR OF WHEAT KERNELS

This trait is determined by two genes that contribute “doses” of red pigment, and display partial dominance (heterozygotes intermediate). Each allele with the subscript "1" contributes 1 dose of red pigment. This trait demonstrates additive effects among different alleles at a single locus, and among alleles at different genetic loci. Genes with subscript "2 " don't contribute any red pigment. With two additive genes you have five phenotypic classes in the F2 offspring of true-breeding strains, instead of the three classes you would see if you had only one additive gene. Some classes are composed of several genotypes that are indistinguishable in phenotype. The number of phenotypic classes expected when there are n diallelic additive loci is (2 n+ 1). So, if have four additive genes, you should expect nine phenotypic classes in the F2 offspring of pure-breeding strains.

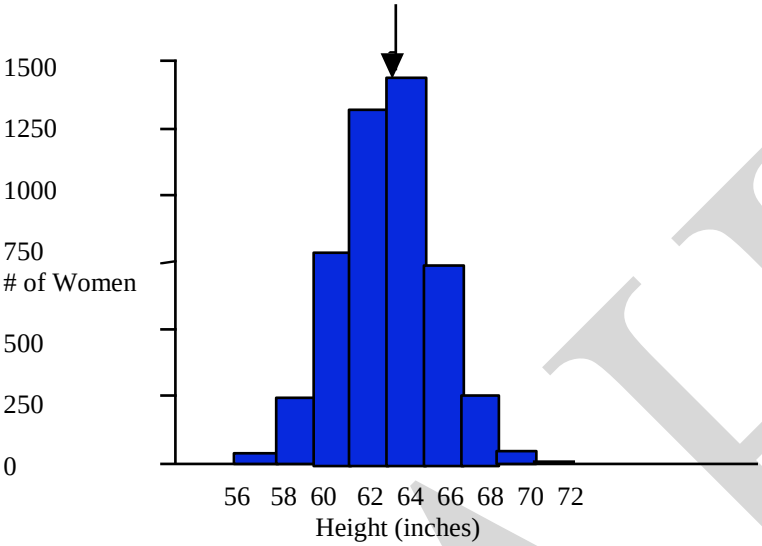
As the number of additive genes increases, the distribution of phenotypes becomes more continuous. In addition, as stated above, most quantitative traits are also affected by the

environment. Environmental effects may obscure genetically-caused differences between phenotypic classes. For example, nutrition affects the adult size in many organisms. The distribution of phenotypes

then becomes even more continuous. The distribution of quantitative traits often approximates a bell-shaped curve when you plot phenotypic value (height, for example) against the frequency of individuals in particular phenotypic classes. Such a plot is called a frequency histogram.

EXAMPLE: DISTRIBUTION OF HEIGHT IN 5000 BRITISH WOMEN

Mean = 63.1 inches



In this graph, the column designated "62" includes all individuals with heights between 61 and 63 inches, "64" includes all individuals with heights between 63 and 65 inches, and so on. This curve has two easily measured properties--the MEAN (average), and the VARIANCE (variation about the mean). Curves with the same mean may have very different variances. If you were to measure the heights of a sample of about 5000 different British women, you would get a curve similar to the one shown above.

Quantitative traits are influenced by genetics and by the environment. Under some circumstances, we can partition the phenotypic variance in quantitative traits into variance that is associated with genetic effects, and variance that is associated with environmental effects:

VP = VG + VE

Where VP is the total phenotypic variance, VG is the genetic variance, and VE is the environmental variance.

Narrow and broad sense heritability

Heritability in broad sense (H^2) is calculated using the following formula :

$H^2 = S_g^2/S_p^2$

Therefore, H^2 simply gives a measure of the proportion of phenotypic variance which is due to genotype. However, this does not tell us what proportion of an individual's phenotype is due to genotype. Moreover, H^2depends on the population in which it is measured and also depends on the set of environments in which the population was grown. For instance in a genotypically homogeneous population (e.g. homozygous parents or F_1 hybrids between them), $S_g^2 = 0$ and therefore $H^2 = 0$. It does not mean that genes do not control the phenotype, but no part of the variation can be ascribed to genetic

variation. Similarly lack of heterogeneity in environment may lead to a high estimate of H^2, but it does not mean that the trait is insensitive to all environments. Since H^2 can not tell us how will a particular trait behave in a particular environment, its utility is limited. In view of this, animal and plant breeders often prefer to estimate heritability in narrow sense (h^2), which is based on the knowledge about gene action. For this purpose genetic variance (S_a^2) is partitioned into additive genetic variance (S_a^2) and **dominance genetic variance** (S_d^2). If *A* and *a* are two alleles at a locus controlling a quantitative trait, the S_a^2 is associated with the average effect of substitution of *A* for *a* in a homozygote (aa →AA) and $S_d^2$is associated with partial dominance of *A* over *a* in heterozygotes. Once $S_a^2$is calculated from given data, S_d^2can be calculated as follows (The details of methods for computations of S_a^2 and $S_d^2$are beyond the scope of this section and the reader is advised to consult a book on Biometrical Genetics for further details) :

$S_d^2 = S_q^2 - S_a^2$.

The heritability in narrow sense (h^2) is used as a measure of relative proportion of S_a^2 as follows :

$h^2 = S_a^2 / S_p^2$ (as against $H^2 = S_q^2 / S_p^2$)

The usefulness of this partitioning of genetic variance lies in the prediction about the effect of selective breeding. Greater is the value of h^2, better are the chances that the selected individuals will maintain their superiority over the population as a whole. This kind of assurance and confidence is not available when we use H^2 (heritability in broad sense). Further, if h^2 is high, the selection will bring about an improvement, but if h^2 is low it means that S_d^2 should be high and heterozygotes will be superior to homozygotes and heterosis breeding should be preferred as done in maize.

Quantitative trait loci (QTL) and their identification

Traditional quantitative genetic research defined a quantitative trait in terms of variances. The total phenotypic was first partitioned into genetic and environmental variances. The genetic variance could then be further divided into additive, dominance and epistatic effects. From this information it was then possible to estimate the heritability of the trait and predict the response of the trait to selection. It was also possible to estimate the minimum number of genes which controlled the trait.

Mapping markers linked to QTLs identifies regions of the genome that may contain genes involved in the expression of the quantitative trait. But what functions could these genes be encoding. To answer this question we should consider a trait such as yield. What types of qualitative genes (genes inherited as simple genetic factors) could be involved in the expression of yield? The first event required for yield is meiosis. Therefore any gene that is involved in gamete formation could potentially be considered a QTL. Any of the genes involved in the protein and carbohydrate biosynthetic pathways could also affect the final yield of a plant and could also be considered to be QTLs. As we saw above, the markers associated with a QTL each account for only a portion of the genetic variance. Likewise each of these genes of known function may only account for a portion of the final yield. An important question that can now be posed is whether any known genes map as QTLs.

Beavis et al. (1991, TAG 83:141-145) analyzed four populations of maize and found molecular markers linked to plant height. No marker was consistently associated as a QTL with plant height in all four populations. Each of the ten maize chromosomes contained a marker linked to a QTL for at least one of the four populations. The authors further were able to demonstrate that a number of the QTLs identified by the molecular markers mapped to regions containing genes known to have a qualitative effect on plant

height. For example, on chromosome 9 the gene d3 resides within 10 cM of a plant height QTL. This gene is involved in gibberellic acid sensitivity. Mutants do not respond to the hormone and do not undergo the normal cell elongation. These mutants are phenotypically shorter than normal maize plants. The question that needs to be raised is if the QTL that was being identified by the molecular marker is actually the d3 gene. It could be possible that what is actually being measured by the marker is the linkage of the marker with the gene.

The statistical analysis of quantitative traits provided valuable information for the plant breeder. Molecular analysis of quantitative traits now provides new tools, not only as selection tools for plant breeding, but as starting points for the cloning of these genes. These objectives could not have been realized without molecular markers.

MODELS OF QUANTITATIVE INHERITANCE
HERITABILITY, GENETIC AND PHENOTYPIC VARIANCE

Some genetic variation is heritable because it can be passed from parent to offspring. Some genetic variation is not strictly heritable, because it is due to dominance or epistatic interactions that are not directly passed from parent to offspring. For example, if one allele is dominant to another, the phenotype of a heterozygous parent is determined in part by the dominance interaction between the two alleles. Since a sexually-reproducing parent only passes on a single allele to its offspring, the offspring does not inherit its whole genotype from a single parent. So it does not inherit the dominance interaction, just the effect of a single allele.

For example, assume that the size of a bird (measured as wing span) is determined by two genes, one with complete dominance of one allele over the other, and one with additive effects.

Interactions between the two different genes are additive. Birds of genotype aabb have 16 cm long wing spans, and birds with other genotypes are somewhat larger, depending upon which alleles they have at each locus. The following table shows the increase in wing span conferred by different genotypes at each locus:

	Dominance effects				**Additive effects**				
	AA +2	Aa +2	aa +0		BB +2		Bb +1	bb +0	
F1:	Aa Bb				x	Aa Bb 19 cm		19 cm	
F2 Genotypes: **Genotypic Effects**	**AABB**	AABb	AAbb	**AaBB**	AaBb	Aabb	**aaBB**	**aaBb**	**aabb**
Phenotype (cm)	**+4**	+3	+2	**+4**	+3	+2	**+2**	**+1**	**+0**
F2 proportions:	**20**	19	18	**20**	19	18	**18**	**17**	**16**
	1/16	2/16	1/16	**2/16**	4/16	2/16	**1/16**	**2/16**	**1/16**

What is the phenotypic mean and variance, assuming we have exactly 16 F2 offspring in the proportions given above?

Mean = [20*(3) + 19*(6) + 18*(4) + 17*(2) + 16*(1)]/16 = 296/16 = 18.5 inches

Variance in phenotypes = [(20-18.5)2*(3) + (19-18.5)2*(6) + (18-18.5)2*(4) + (17-18.5)2*(2) + (16-18.5)2*(1)]/(16-1) = 20/15 = 1.333 in 2

SOURCES OF VARIANCE

In this case, all the phenotypic variance is due to differences in genotypes, so it is all genetic variance. However, some of the difference between the genotypes is due to additive effects of alleles, and some are due to dominance effects of alleles. For example, the difference in phenotype between aabb, aaBb, and aaBB (16, 17, and 18 respectively) are only due to the additive interactions between different alleles at the B locus. However, the difference in phenotype between AABB, AaBB, and aaBB (20, 20, and 18, respectively) is due to both additive effects (difference between aa and AA is 2 "units" of the A allele and the difference in height is 2 inches--so the average effect of a "unit" of A is one inch) and dominance effects at the A locus. Therefore, some of the genetic variation is due to additive effects and some is due to dominance effects. In fact, although the derivation of the equation is beyond the scope of this course, the amount of additive (VA) vs. dominance variance (VD) is easily calculated in this example: VA= 1.0 and VD = 0.333; (VA=2∑piqiai2; VD=∑(2pqd2), where the sum is over each locus contributing to the trait). In this simplified example, there are no environmental effects, so the environmental variance (VE) is zero. If there were environmental effects, the phenotypic variance would be the sum of the genetic and environmental variance.

The additive and non- additive genetic variance (VA and VD) together are known as the genotypic or genetic variance, which is abbreviated VG. In the bird- wing example, VG = 1 . 333 , which is also the total phenotypic variance is abbreviated VP = VG + VE.

These considerations (and the mathematical fact that variances due to independent sources of variation can be summed) lead to the following equations:

VP = VG + VE

VP = VA + VD + VE

In the bird- wing example, all the phenotypic variance was due to genetic variance (there were no environmental effects on the trait). We can also imagine a trait that has no genetic variance, so that the differences between individuals are due only to environmental effects. For example, assume that you grow a clonal strain of corn in a field, so that every individual has exactly the same genotype. At the end of the growing season, there will be differences in height (VP) among different plants, but all the differences will be due to local soil, moisture, temperature, and light conditions. Thus, all the height differences are due to environmental variance (VE).

A useful thing for plant and animal breeders to know is, for any trait of interest, how much of the phenotypic variability of that trait is due to genetic variance, and how much is due to non-genetic

environmental factors.

This is the broad- sense heritability: H 2 = VG/VP

It is even more useful to know what proportion of the phenotypic variation is due to additive genetic effects.

The heritability (narrow- sense) of a trait is defined as the proportion of the total phenotypic variation that is due to heritable (additive genetic) effects:

VA/ VP = h2

h2 is the proportion of variability that can be passed on from parent to offspring, and this is why this quantity is of interest to animal and plant breeders. They want to know whether selection on the parents will produce inherited changes in the offspring. The degree to which selection on parents will produce inherited changes in the offspring is determined by the narrow sense heritability of the trait.

When h2 = 0, none of the phenotypic variance among individuals is due to additive genetic differences (VA= 0) and offspring will not closely resemble their parents for the trait of interest for genetic reasons. Can you think of any traits in humans, other animals, or plants that probably have zero heritability?

Note: Offspring can resemble their parents for non- genetic reasons. E. g., consider the fact that human offspring tend to have the same religious affiliation as their parents.

What do you think is the source of this kind of resemblance between parents and offspring?

When h2 = 1, all the variation among individuals is due to heritable genetic differences (VP= VA) and offspring will resemble their parents very closely. Can you think of any traits that probably have high heritability?

h2 is a very useful number, because it allows us to predict how a population will respond to artificial or natural selection. Animal and plant breeders rely heavily on the methods of quantitative genetics to improve economically-important traits such as milk yield in dairy cows, speed and endurance in race horses, protein and oil content in corn, or size, color, and ripening time in tomatoes.

The relationship between h2 and the response to selection, is given by:

R = h2 S

R is the response to selection, given by the difference between the population mean before selection and the mean of the offspring of selected parents after one generation of selection.

S is the selection coefficient, given by the difference between the unselected population mean, and the mean of the selected parents.

If we know the heritability of a trait, and the strength of artificial selection applied to it, we can predict the response to selection. For example, in a population of the tobacco plant, *Nicotiana longiflora*, the mean flower length of the unselected population is 83 cm, and the mean of selected parents is 90 cm. If the heritability of the trait is known from other studies to be 0.64, then the predicted response to selection is given by R = h2S, where S = 90-83=7 cm.

R = 0 . 64 * 7 cm = 4 . 48 cm.

The mean of the population after one generation of selection is therefore predicted to be:

New mean = Unselected mean + predicted response = 83 cm + 4 . 48 cm = 87 . 48 cm

The actual result of this selection experiment was that the mean corolla length after one generation of

selection was 87.9 cm.

Because of this relationship between R, S, and h2, one way to estimate an unknown heritability is to measure the response to a known amount of artificial selection.

$$R = h2\ S \qquad h2 = R/S$$

In fact, the value of h2 = 0.64 used above was calculated from the results of a previous episode of artificial selection in tobacco. The unselected population mean was 70 cm, and the mean of the selected parents was 81 cm. Thus, the difference between the mean of selected parents and the unselected mean (S) was 81-70=11 cm. The mean of the offspring of the selected parents was 77 cm. So the response (R) to a single generation of selection was 77-70=7 cm. The estimate of h2 is therefore

h2 = R/ S = 7 /11 = 0.636.

So about 64 percent of the total phenotypic variation in corolla length is due to heritable genetic differences. The remainder of the variation (36 percent) is due to environmental (or possibly to non-additive genetic) sources of variation.

QUANTITATIVE TRAITS IN HUMANS

TWIN STUDIES

Obviously, selection experiments cannot be performed in humans, but in some cases we would like to know if diseases or other traits are affected by genetic factors or by environmental factors, or both. One method that has frequently been used to try to address these questions is the twin- study approach.

Identical twins arise from the splitting of a single fertilized egg, and are genetically identical. They have the same alleles at all genetic loci. Theoretically, any phenotypic differences between identical twins are environmental.

Fraternal twins arise from two fertilized eggs, and have the same genetic relatedness as ordinary siblings (on average they share 50% of all alleles at all loci because of independent assortment of chromosome pairs at meiosis--review meiosis if you don't understand this). Therefore, phenotypic differences between fraternal twins can be due to both environmental and genetic differences.

One way to estimate heritability from twin data is to measure a trait in a large number of identical and fraternal twin pairs. One can then use the correlation between twin pairs for the trait concerned. In this case the correlation between co-twins gives an estimate of the heritability of a trait. If the heritability is high, identical twins will normally be very similar for a trait, and fraternal twins will be less similar. If the heritability is low, identical twins may not be much more similar than fraternal twins. If variation for a trait is completely heritable, then identical twins should be have a correlation near 1, and fraternal twins should have a correlation near 0.5, since they will be similar for about 50% of the genetically variable loci. If we can measure a trait in many sets of identical and fraternal twins, we can calculate the correlation for both kinds of twins

A measure of the broad-sense heritability, H2, is given by the difference between the two correlations: H 2 = 2 *(ri- rf). Actually, this calculation overestimates H2, by an amount equal to 1/2 (VD/VP), so the heritabilities calculated by this formula should be interpreted as maximal possible values. The table below gives correlations between identical and fraternal human twins for several different traits.

Human Traits	Correlations		Heritability = H 2
	Identical Twins	Fraternal Twins	
Fingerprint ridge count	0.96	0.47	0.98
Height	0.90	0.57	0.66
IQ score	0.83	0.66	0.34
Social maturity score	0.97	0.89	0.16

PROBLEMS WITH USING TWIN STUDIES TO ESTIMATE HERITABILITY IN HUMANS

In addition to the fact that 2 *(ri- rf) overestimates H2 by a factor that is usually unknown, there are several additional sources of error in such studies.

Genotype-environment interaction (some genotypes are affected differently by the environment than are other genotypes--remember the rat maze-brightness example). Since different genotypes are lumped together in the category 'fraternal twins', but not in 'identical twins', GxE interaction lowers the correlation between fraternal twins, but not identical twins (rf is underestimated).

An important assumption of these calculations is that identical twins and fraternal twins experience the same degree of environmental similarity. This assumption is likely to be violated because of:

Greater similarity in the treatment of identical twins by parents, teachers, and peers, resulting in inflated correlations between identical twins (ri is overestimated).

Frequent sharing of embryonic membranes between identical twins, resulting in a more similar intrauterine environment (ri is overestimated).

Different sexes in half of the pairs of fraternal twins, in contrast with the same sex of identical twins (rf is underestimated).

Finally, Many of these studies are based on small samples, and so the estimates are subject to large deviations from the 'true' value due to chance.

Therefore, estimates of heritability derived from human twin studies should be considered very approximate, and probably too high.

The Hardy-Weinberg law

-A *population* is a group of interbreeding individuals, of the same species that inhabit a certain area at a certain time.

-The *gene pool* is the sum of all the alleles present in a population; alterations to this pool give rise to *microevolution*.

-The Hardy-Weinberg Principle says that heredity itself cannot cause changes in the frequencies of alternate forms of the same gene (alleles).

-If certain conditions are met, then the proportions of genotypes that make up a population of organisms should remain constant generation after generation according to Hardy- Weinberg equilibrium:

$$(p + q)^2 = p^2 + 2pq + q^2 = 1.0$$ (for two alleles)

-However, this formula can be used for more than two alleles. Use the binomial theorem to expand such an expression eg. $(p + q + r)^2 = 1.0\ldots\ldots = p^2 + q^2 + r^2 + 2pq + 2qr + 2pr = 1.0$

-If p is the frequency of one allele (A), and q is the frequency of the other allele (a), then

$p + q = 1.0$

-If two alleles for coat colour exist in a population of mice, and the allele for white coats is present 70% of the time, then the alternate allele (black) must be present 30% of the time. The Hardy-Weinberg Principle can be used to determine the proportions of phenotypes present in succeeding generations, as long as conditions do not change. In our example, since $p = 0.7$, we would expect 49% (p^2) of the mice in our population to be homozygous for white coats.

-Forty-two percent (*2pq*) would have one of each allele and would appear gray if the alleles are co-dominants (that is, both alleles have equal expression in the phenotype).

-In nature, however, the frequencies of genes in populations are not static (that is, not unchanging). Natural populations never meet all of the assumptions for Hardy-Weinberg equilibrium.

Predicting allele and genotype frequencies

The assumptions for Hardy-Weinberg equilibrium are:

1. The organism in question is diploid.
2. Reproduction is sexual.
3. Generations are non-overlapping.
4. Mating is random.
5. Population size is very large – virtually infinite.
6. Migration is negligible. (i.e., no immigration or emigration occurs)
7. Mutation does not occur.
8. Natural selection does not affect the locus under consideration (i.e., all genotypes are equally likely to reproduce).

-*Evolution is a process resulting in changes in the genetic makeup of populations through time;* therefore, factors that disrupt Hardy-Weinberg equilibrium are referred to as **evolutionary agents**. In random mating populations, natural selection, gene flow, genetic drift, and mutation can all result in a shift in gene frequencies predicted by the Hardy-Weinberg formula.

-Nonrandom mating can also result in such changes.

-Even though genotypic frequencies change slightly from generation to generation, allele frequencies remain static. If not, something 'interesting' is occurring in the population.

An example of a HW problem (including hypothesis testing with the chi-squared test):

-Use of the Hardy–Weinberg law is exemplified in Table I, which gives the results of MN blood typing of 6129 American Caucasians:

1787 *MM*,
3037 *MN*,
1305 *NN*.

The allele frequencies of *M* and *N* are calculated and the Hardy–Weinberg frequencies obtained as shown.

• The expected number of each genotype is obtained by multiplying the Hardy-Weinberg frequencies by the sample size (6129).

• The chi-squared value (χ^2) for goodness of fit is then calculated as the sum of (*observed number* – *expected number*)^2 / *expected number* for each genotypic class, producing χ^2 =

.04887.

-For the following example, keep in mind that frequencies in first generation are calculated as follows [$2(N_{RR} + N_{Rr} + N_{rr})$ is the total number of chromosomes]:

Frequency of M (p)	Frequency of m [q = 1-p]
$p = \frac{2N_{MM} + N_{Rr}}{2(N_{RR} + N_{Rr} + N_{rr})}$	$q = \frac{2N_{rr} + N_{Rr}}{2(N_{RR} + N_{Rr} + N_{rr})}$

	Genotype MM	MN	NN	Total
No. of individuals	1787	3037	1305	6129
No. of *M* alleles	3574	3037	0	6611
No. of *N* alleles	0	3037	2610	5647
No of *M* + *N* alleles	3574	6074	2610	12 258
	Allele frequency of $M = 6611/12\,258 = 0.53932 = p$			
	Allele frequency of $N = 5647/12\,258 = 0.46068 = q$			
Expected frequency	$p^2 = 0.29087$	$2pq = 0.49691$	$q^2 = 0.21222$	1.000
Expected no. (frequency x 6129)	1782.7	3045.6	1300.7	6129
Chi-squared value	$\frac{(1787-1782.7)^2}{1782.7}$	$\frac{(3037-3045.6)^2}{3045.6}$	$\frac{(1305-1300.7)^2}{1300.7}$	0.04887
	Degrees of freedom = 3 - 1 - 1 = 1			
	Probabilty associated with χ^2 (df = 1) of 0.04887 is about 0.90			

-For the chi-sq. test, the data themselves are used to calculate the expectations; specifically, the data are first used to estimate the allele frequency of *M* (*p*=0.53932; note that only *p* need be estimated from the data because *q* can be calculated from the relation *q* = 1 – p), and then this value of *p* is used in the Hardy–Weinberg formula along with the total sample size in order to obtain the expected number in each genotypic class.

-Since the value of *p* used in calculating the expectations is obtained from the data themselves, one should expect a better fit than would be the case were *p* obtained from some other source.

-Thus a smaller p^2 should suffice to reject the hypothesis that the genotypes are in Hardy– Weinberg proportions, and the test should somehow take this into account. freedom for estimating *p* from the data.

-For the chi-squared test, therefore, the appropriate number of degrees of freedom is 3 (the number of classes) minus 1 (for using the sample size in calculating expectations) minus 1 (for estimating *p* from the data), or 3 – 1 – 1 = 1.

-*[The general rule for chi-squared tests of the type in question is that the number of degrees of freedom equals the number of classes of data minus 1 (for using the sample size) minus the number of parameters estimated from the data.]*

-The probability value associated with a χ^2 of .04887 with 1 degree of freedom is about 0.90 – a remarkably good fit. A χ^2 as large or larger would be obtained by chance 90 percent of the time in a population having its actual genotype frequencies as calculated from the Hardy–Weinberg law; thus there is no reason for thinking this population does not obey the Hardy–Weinberg law for the MN locus.

-The **chi-square** test is one way to test an hypothesis in an experiment in which the data collected are **frequency data**, rather than continuous data (see *t* test above). Note that the χ^2 analysis uses raw data

only, not percentages or proportions [this will be discussed in tutorials this week].

The χ^2 formula is

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

Degrees of freedom	Level of probability		
	0.10	0.05	0.01
1	2.71	3.84	6.64
2	4.61	5.99	9.21
3	6.25	7.82	11.34
4	7.78	9.49	13.28

-However, lets say you had a population of 1 in 4000 Asians that are affected by alpha-thalassemia (this is an autosomal recessive disease). Then, how will you calculate the *genotypic* frequency of alpha-thalassemia cases? [it would be 1/4000 = q^2]. How about the frequency of the recessive allele? [q] and the frequency of heterozygotes? [2pq].

Exceptions to the Hardy-Weinberg principle

Mutations:

Mutation in only one direction can cause one allele slowly to replace another. Mutation in both directions results in an equilibrium with frequencies determined by the mutation rates.

Migrations:

Migration between populations always causes the gene frequencies of the receiving population to shift towards those of the immigrants.

Nonrandom mating:

Inbreeding (increased mating among related individuals) results in more homozygotes.

Assortative mating is mating according to phenotype, with mating between phenotypically similar individuals being either more frequent (positive assortment) or less frequent (negative assortment).

Genetic drift:

If the population is not large, gene frequencies can fluctuate randomly in either direction simply by chance. Selection occurs whenever different genotypes contribute genes unequally to the next generation. This is the most important and most frequent departure from equilibrium conditions.

Influence of natural selection and mutations on the HW equilibrium:

-*Natural selection* disturbs Hardy-Weinberg equilibrium by discriminating between individuals with respect to their ability to produce young. Those individuals that survive and reproduce will perpetuate more of their genes in the population. These individuals are said to exhibit greater fitness than those who leave no offspring or fewer offspring.

-Environment is usually what selects for certain alleles of a gene.

-Two components: viability of individuals and reproductive success. It acts with no regards for individuals but rather the population as a whole. Mutations usually are deleterious and are thus

selected against by natural selection.
-Selection that favors one extreme phenotype over the other and causes allele frequencies to change in a predictable direction is known as directional selection**.** When selection favors an intermediate phenotype rather than one at the extremes, it is known as stabilizing selection**.**
-Selection that operates against the intermediate phenotype and favors the extreme ones is called disruptive selection**.**
-It is important to realize that selection operates on the entire phenotype so that the overall fitness of an organism is based on the result of interactions of thousands of genes.
=However, some disease alleles still prevail in the population, even though natural selection purges these out. Why? Because these alleles are present in heterozygous individuals that may not express the disease. Also, sometimes heterozygous have a higher fitness than either homozygote, a situation referred to as the heterozygous advantage, eg. Malaria and blood cells...
-Pattern of expression of alleles can also influence the influence natural selection has on them. If, for example, as disease is expressed at a late age, when the individual has reproduced, there is little natural selection acting here, so the disease alleles are likely to prevail in population.

Influence of gene flow on the HW equilibrium:
- The frequencies of alleles in a population also change if new organisms immigrate and interbreed, or when old breeding members emigrate. *Gene flow* due to migration may be a powerful force in evolution.
- Some level of gene flow is necessary to keep local populations of the same species from becoming more and more different from each other. Things that serve as barriers to gene flow may accelerate the production of new species. Migration may also introduce new genes into a population and produce new genetic combinations.

Genetic Drift and the Founder Effect

-Chance is another factor that affects the kind of gametes in a population that are involved in fertilization. As a result, shifts in gene frequencies can occur between generations just because of the random aspects of fertilization. This phenomenon is known as *genetic drift*.
-Generally, the larger the breeding population, the smaller the sampling effect that we call genetic drifts. In small populations, genetic drift can cause fluctuations in gene frequencies that are great enough to eliminate an allele from a population, such that p becomes 0.0 and the other allele becomes fixed (q = 1.0). Genetic variation in such a population is reduced.
Populations that become very small may lose much of their genetic variation. This is known as a *bottleneck effect*.
-Another way in which chance affects allele frequencies in a population is when new populations are established by migrants from old populations.
-The genetic makeup in future generations in the new population will more closely resemble the six migrants than the population from which the migrants came. This effect is known as the *founder effect*. The founder effect may not be an entirely random process because organisms that

migrate from a population may be genetically different from the rest of the population to begin with. For example, if wing length in a population of insects is variable, one might expect insects with longer wings to be better at founding new populations because they may be carried farther by winds.

MOLECULAR EVOLUTION

For evolutionists the revolution in DNA technology has been a major advance. The reason is that the very nature of DNA allows it to be used as a "document" of evolutionary history: comparisons of the DNA sequences of various genes between different organisms can tell us a lot about the relationships of organisms that cannot be correctly inferred from morphology. One definite problem is that the DNA itself is a scattered and fragmentary "document" of history and we have to beware of the effects of changes in the genome that can bias our picture of organismal evolution.

Analysis of nucleotide and amino acid sequences

Two general approaches to molecular evolution are to 1) use DNA to study the evolution of organisms (such as population structure, geographic variation and systematics) and to 2) to use different organisms to study the evolution of DNA. To the hard-core molecular evolutionist of the latter type, organisms are just another source of DNA. Our general goal in all this is to infer process from pattern and this applies to the processes of organisms evolution deduced from patterns of DNA variation, and processes of molecular evolution inferred from the patterns of variation in the DNA itself. An important issue is that there are processes of DNA change within the genome that can alter the picture we infer about both organisms and DNA evolution: the genome is fluid and some of the very processes that make genomes "fluid" are of great interest to evolutionary biologists. Thus molecular evolution might be called the "natural history of DNA". The points that follow are some interesting observations interspersed with some basic concepts.

Some important background: DNA has many different roles in terms of function. Most of our DNA does not code for proteins (more below) and thus is quite a different type of character/trait than DNA that does code for protein. In eukaryotes, genes are frequently broken up into exons (expressed) and introns (spliced out of the RNA before becoming a true messenger RNA). The genes also have regulatory sequences that indicate when and where to transcribe the DNA into RNA for protein synthesis. The genetic code is the information system for translating the sequence of RNA into the sequence of amino acids. Within this triplet code some of the nucleotide positions are silent or synonymous because any nucleotide in that position will do. This "universal" code is not completely universal because the mitochondrial genome uses some of the codons in different ways (e.g., some termination codons in the universal code specify amino acids in the mitochondrial code). Thus even the genetic code can evolve.

Some important theoretical background: we want to develop a picture of what happens to a new mutant in a population, lets say a single nucleotide change a one position in the DNA. This is the starting point for molecular evolution. If the new mutant is governed by genetic drift, its fate should be quite different than another nucleotide mutation that is governed by selection (see below). To describe molecular evolution Kimura formulated the Neutral theory of molecular evolution which is remarkably simple. If:

u = mutation rate / gene / generation, N = population size, then the number of new mutations occurring per generation in a population = 2Nu (2 because we are considering diploid organisms).

Now, when a new mutation occurs in a population its initial frequency = 1/2N because it is the one new variant out of a total of 2N genes in the population. This is also its probability of fixation because the

probability of you reaching into a barrel of 2N marbles and getting the one new marble is 1/2N. Thus taking these two values (the number of new mutations/generation and the probability of fixation), the rate of substitution, K is just their product: 2Nu 1/2N = u (substitution means that the new mutant goes to fixation in the population and substitutes the original nucleotide or gene). Kimura got very famous for this simple bit of algebra. It tells us that the neutral rate of molecular evolution is equal to the neutral mutation rate. Now a question arises: what is the distribution of the types of mutations? Are most neutral? Are some deleterious?; beneficial?. One consequence of the neutral theory is that genes with different mutation rates will have different rates of evolution.

The rate of evolution of a gene or mutation that is under selection will be very different. Similarly different genes with different functions, or different parts of a gene with different functions will have different rates of evolution. Thus, different regions of DNA with different functional constraints will evolve at different rates One prediction of the neutral theory is that silent (synonymous) sites in protein coding regions will evolve faster than replacement (nonsynonymous) sites (due to different functional constraints). This provides a null hypothesis about DNA evolution. Most sequences fit this neutral model; however, the histocompatibility loci appear to deviate from a neutral model in that there are more nonsynonymous substitutions than synonymous substitutions. This holds only for the antigen binding region; the rest of the molecule is consistent with neutral expectations.

Another prediction of the neutral theory is that amount of sequence divergence will be correlated with the level of heterozygosity; heterozygosity is measured as 2pq for a two allele situation or (2pq+2pr+2qr) for a three allele situation, or 1- x_i^2 for i alleles; see figure 5.2 pg. 96 for a two allele view). Loci with high heterozygosity should evolve at a faster rate under the assumption that these loci have a higher rate of neutral mutation (thus more variation within species and more substitution between species; se Kimura's proof above). In general, loci fit this relationship, however, balancing selection at a locus will introduce more heterozygosity than expected. What about purifying selection?

Different rates of substitution have also been observed in different lineages of organisms: For the human - chimp divergence, the rate = 1.3×10^{-9} substitutions/nucleotide site/year

for the human - Old World monkey split, the rate = 2.2×10^{-9} substitution/site/year

for the mouse - rat split (=rodents), the rate = 7.9×10^{-9} substitutions/site/year. Thus, rodents appear to have a faster rate of molecular evolution. It has been argued that a shorter generation time in rodents accounts for the faster rate of evolution, the so called generation time effect. There are examples where short generation species have slower rates of evolution; the point is that rates differ, the cause(s) of these rate differences have not been unambiguously identified.

Most discussions of the rates of DNA evolution have been with respect to the molecular clock hypothesis which states that there is a positive linear relationship between time since two species diverged and amount of genetic divergence (e.g., DNA sequence difference) between those species. These observations stated above indicate that there is not one molecular clock but probably many molecular clocks that "tick" at different rates.

Lets say we identify a reliable molecular clock (e.g., number of amino acid substitutions in the cytochrome C gene), we can use this to date, or corroborate, evolutionary events of interest (e.g., the divergence times for species that do not have good fossil data). For example: we know that there are K_{XY} substitutions between species X and Y and we know that they diverged T years ago (from fossil data). Thus the rate of molecular evolution is $r = K_{XY}/2T$. The denominator has a 2 in it because there

are two paths of evolution on which the divergence can accumulate (ancestor to X and ancestor to Y). Now lets say we obtain the sequence of cytochrome C from species A, B and C. From these data we count up the number of substitutions for all three pair wise comparisons (the K's). We are given the date of divergence between A and C (=T1) and we want to date the divergence of A and B (T2 = ?) but don't have any fossil data. If one assumes that the rate of evolution is the same in species A, B and C as it was measured to be between species X and Y, we can use the amount of sequence divergence between species A, B and C to estimates their dates of divergence (measured in millions of years before present, MYBP).

Repetitive DNA Studies of many organisms has revealed that a large proportion of eukaryotic genomes consists of repetitive DNA. Some of this is short localized repeats: in the kangaroo rat the sequence (AAG) is repeated 2.4 billion times, the sequence (TTAGGG) is repeated 2.2 billion times and the sequence (ACACAGCGGG) is repeated 1.2 billion times. What it does is unclear. Sequences like this have been called junk DNA. Note that junk is stuff you don't throw away because it might be useful some day; garbage is stuff you don't want so you throw it away. These sequences might have some function we don't know about so they have been called junk DNA. The fact that such sequences seem to accumulate in genomes has lead to the notion that repetitive DNA is selfish DNA, since the sequence makes additional copies of itself within the genome decoupled from the reproduction rate of the host (i.e., the kangaroo rat).

Another form of repetitive DNA are transposable elements. These are sequences of DNA that generally code for certain proteins and have the ability to move around the genome in a process called transposition. There are quite a number of different types of such elements (we will not review them all). The point is that they (like other repetitive DNA) are governed by intragenomic dynamics as well as organismal population dynamics. An example is the P element in *Drosophila melanogaster*. There are strains of flies that have P elements (P strains) and strains that do not (M strains). When a P male is crossed to an M female the P elements enter the genome of the offspring and jump around causing mutations (this is what we mean by a fluid genome). A curious observation about P elements is that strains of flies collected from natural populations before 1950 do not have P elements whereas flies collected from the wild after that do have them. A variety of observations indicate that P elements invaded *D. melanogaster* recently. The best evidence is that *D. melanogaster*'s close relative do not have P elements, but a more distantly related fly *D. willistoni* does have them and they differ by only a few nucleotides over 2900 base pairs of DNA. These data suggest that P elements in *D. melanogaster* are the result of a horizontal gene transfer (horizontal as opposed to "vertical" as one inherits DNA from ones parents or ancestor "above"). Thus, not only can DNA move around the fluid genome, but if DNA from one species can enter the gene pool of another without the species fusing into one, one has to be very aware of what DNA sequence one is using to determine phylogenies, etc.

If multiple copies of a DNA sequence are present in a genome we can think of each sequence as a single "species" evolving on its own "line of descent" because each repeat will be mutated at random. Thus if we had the complete DNA sequence of all the repeated P elements within a genome, we would find that they are not identical and thus we could build a cladogram of these elements much like we can build a cladogram of birds. When this sort of analysis was done on different kinds of repeated elements (many copies of ribosomal DNA, for example) it was found that the copies showed almost no variation. This observation suggested that all the repeats of this family (ribosomal DNA family) were evolving in concert, i.e., together. This pattern of homogeneity of repeats is called concerted evolution . The process(es) that

generate this pattern could be unequal crossing over or gene conversion. Gene conversion is when the sequence of one region of DNA is used as a "template" to "correct" or modify the sequence of another region of DNA. We do not need to go into the molecular details of these processes, but that DNA can evolve in concert with other sequences in the genome again indicates that intragenomic dynamics can influence the pattern of DNA variation we see within and between species.

Gene duplication is a minimalist version of repetitive DNA. Many genes in the genome are duplicated and when this happens one of the copies may be "freed" from constraints and evolve a new function. The best understood case of this phenomenon is the evolution of the globin genes myoglobin, □-hemoglobin, ß-hemoglobin. The existence of duplicated genes forces us to recognize different kinds of homology because there are two ways to have a common ancestor: by gene duplication and by speciation. When two genes share a common ancestor due to a duplication event we call them paralogous (□-hemoglobin and ß-hemoglobin in you are paralogous as are the □-hemoglobin in you and the ß-hemoglobin in chimps). When two genes share a common ancestor due to a speciation event we call them orthologous (□-hemoglobin in you and □-hemoglobin in chimps). Obviously when constructing a cladogram from molecular data one should use orthologous genes if one wants to build a tree of organisms.

A further example of the fluid genome is exon shuffling This is the pattern observed when exons or functional domains of genes are shuffled together to form new or modified genes. Some genes have very distinct domains that have clear relationships to other domains in very different genes. It is thought that these domains have been moved around the genome by transposition or illegitimate recombination events in evolution, accidentally forming new associations that happen to have novel functions. Wally Gilbert proposed the idea of exon shuffling and argued that such a phenomenon might accelerate evolution by creating new material for adaptive evolution.

One of the more interesting observations to ponder in molecular evolution is the C-value paradox. The C value of a species is the Characteristic or Constant amount of DNA in a haploid genome of that species. If we look at the diversity of organisms from viruses to humans we see a clear trend in biological complexity. If we compare the C values across this range of organisms (assuming viruses are "organisms") some of the less complex organisms have much more DNA than the more complex organisms. This presents a paradox. If DNA codes for proteins that give us form and function, what is a lowly alga doing with all that DNA? We just don't know. Presumable most of it is not "functional" in the sense of coding for proteins and RNAs. Much of it may be "junk DNA", but maybe this "junk" helps in aligning chromosomes properly during mitosis an meiosis.

Molecular Phylogenies

If genomes evolve by the gradual accumulation of mutations, then the amount of difference in nucleotide sequence between a pair of genomes should indicate how recently those two genomes shared a common ancestor. Two genomes that diverged in the recent past would be expected to have fewer differences than a pair of genomes whose common ancestor is more ancient. This means that by comparing three or more genomes it should be possible to work out the evolutionary relationships between them. These are the objectives of molecular phylogenetics.

The Origins of Molecular Phylogenetics

Molecular phylogenetics predates DNA sequencing by several decades. It is derived from the traditional method for classifying organisms according to their similarities and differences, as first practiced in a comprehensive fashion by Linnaeus in the 18th century. Linnaeus was a systematicist not an evolutionist,

his objective being to place all known organisms into a logical classification which he believed would reveal the great plan used by the Creator - the *Systema Naturae*. However, he unwittingly laid the framework for later evolutionary schemes by dividing organisms into a hierarchic series of taxonomic categories, starting with kingdom and progressing down through phylum, class, order, family and genus to species. The naturalists of the 18th and early 19th centuries likened this hierarchy to a 'tree of life', an analogy that was adopted by Darwin (1859) in *The Origin of Species* as a means of describing the interconnected evolutionary histories of living organisms. The classificatory scheme devised by Linnaeus therefore became reinterpreted as a phylogeny indicating not just the similarities between species but also their evolutionary relationships.

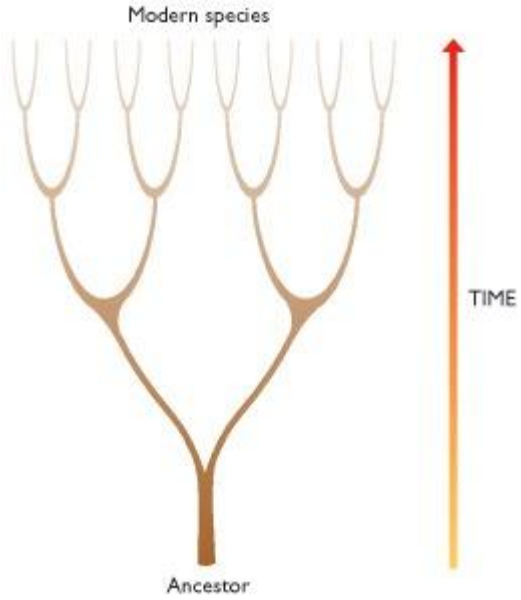
Nuttall's work showed that molecular data can be used in phylogenetics, but the approach was not widely adopted until the late 1950s, the delay being due largely to technical limitations, but also partly because classification and phylogenetics had to undergo their own evolutionary changes before the value of molecular data could be fully appreciated. These changes came about with the introduction of phenetics and cladistics, two novel phylogenetic methods which, although quite different in their approach, both place emphasis on large datasets that can be analyzed by rigorous mathematical procedures. The difficulty in obtaining large mathematical datasets when morphological characters are used was one of the main driving forces behind the gradual shift towards molecular data, which have three advantages compared with other types of phylogenetic information:

- When molecular data are used, a single experiment can provide information on many different characters: in a DNA sequence, for example, every nucleotide position is a character with four character states, A, C, G and T. Large molecular datasets can therefore be generated relatively quickly.
- Molecular character states are unambiguous: A, C, G and T are easily recognizable and one cannot be confused with another. Some morphological characters, such as those based on the shape of a structure, can be less easy to distinguish because of overlaps between different character states.
- Molecular data are easily converted to numerical form and hence are amenable to mathematical and statistical analysis.

The sequences of protein and DNA molecules provide the most detailed and unambiguous data for molecular phylogenetics, but techniques for protein sequencing did not become routine until the late 1960s, and rapid DNA sequencing was not developed until 10 years after that. Early studies therefore depended largely on indirect assessments of DNA or protein variations, using one of three methods:

- **Immunological data**, such as those obtained by Nuttall (1904), involve measurements of the amount of cross-reactivity seen when an antibody specific for a protein from one organism is mixed with the same protein from a different organism. The antibodies are immunoglobulin proteins that help to protect the body against invasion by bacteria, viruses and other unwanted substances by binding to these 'antigens'. Proteins also act as antigens, so if human β-globin, for example, is injected into a rabbit then the rabbit makes an antibody that binds specifically to that protein. The antibody will also cross-react with β-globins from other vertebrates, because these β-globins have similar structures to the human version. The degree of cross-reactivity depends on how similar the β-globin being tested is to the human protein, providing the similarity data used in the phylogenetic analysis.

- Protein electrophoresis is used to compare the electrophoretic properties, and hence degree of similarity, of proteins from different organisms. This technique has proved useful for comparing closely related species and variations between members of a single species.
- **DNA-DNA hybridization data** are obtained by hybridizing DNA samples from the two organisms being compared. The DNA samples are denatured and mixed together so that hybrid molecules form. The stability of these hybrid molecules depends on the degree of similarity between the nucleotide sequences of the two DNAs, and is measured by determining the melting temperature, a stable hybrid having a higher melting temperature than a less stable one. The melting temperatures obtained with DNAs from different pairs of organisms provide the data used in the phylogenetic analysis.



By the end of the 1960s these indirect methods had been supplemented with an increasing number of protein sequence studies (e.g. Fitch and Margoliash, 1967) and during the 1980s DNA-based phylogenetics began to be carried out on a large scale. Protein sequences are still used today in some contexts, but DNA has now become by far the predominant molecule. This is mainly because DNA yields more phylogenetic information than protein, the nucleotide sequences of a pair of homologous genes having a higher information content than the amino acid sequences of the corresponding proteins, because mutations that result in non-synonymous changes alter the DNA sequence but do not affect the amino acid sequence. Entirely novel information can also be obtained by DNA sequence analysis because variability in both the coding and non-coding regions of the genome can be examined. The ease with which DNA samples for sequence analysis can be prepared by PCR is another key reason behind the predominance of DNA in modern molecular phylogenetics.

As well as DNA sequences, molecular phylogenetics also makes use of DNA markers such as RFLPs, SSLPs and SNPs, particularly for intra specific studies such as those aimed at understanding migrations of prehistoric human populations. Later in this chapter we will consider various examples of the use of both DNA sequences and DNA markers in molecular phylogenetics, but first we must make a more detailed study of the methodology used in this area of genome research.

The Reconstruction of DNA-based Phylogenetic Trees

The objective of most phylogenetic studies is to reconstruct the tree-like pattern that describes the evolutionary relationships between the organisms being studied. Before examining the methodology for doing this we must first take a closer look at a typical tree in order to familiarize ourselves with the basic terminology used in phylogenetic analysis.

The key features of DNA-based phylogenetic trees

A typical phylogenetic tree could have been reconstructed from any type of comparative data, but as we are interested in DNA sequences we will assume that the tree shows the relationships between four homologous genes, called *A*, *B*, *C* and *D*. The topology of this tree comprises four external nodes, each representing one of the four genes that we have compared, and two internal nodes representing ancestral genes. The lengths of the **branches** indicate the degree of difference between the genes represented by the nodes. The degree of difference is calculated when the sequences are compared.

The tree is unrooted, which means that it is only an illustration of the relationships between *A*, *B*, *C* and *D* and does not tell us anything about the series of evolutionary events that led to these genes. Five different evolutionary pathways are possible, each depicted by a different rooted tree. To distinguish between them the phylogenetic analysis must include at least one outgroup, this being a homologous gene that we know is less closely related to *A*, *B*, *C* and *D* than these four genes are to each other. The outgroup enables the root of the tree to be located and the correct evolutionary pathway to be identified. The criteria used when choosing an outgroup depend very much on the type of analysis that is being carried out. As an example, let us say that the four homologous genes in our tree come from human, chimpanzee, gorilla and orangutan. We could then use as an outgroup the homologous gene from another primate, such as the baboon, which we know from paleontological evidence branched away from the lineage leading to human, chimpanzee, gorilla and orangutan before the time of the common ancestor of those four species.

We refer to the rooted tree that we obtain by phylogenetic analysis as an inferred tree. This is to emphasize that it depicts the series of evolutionary events that are inferred from the data that were analyzed, and may not be the same as the true tree, the one that depicts the actual series of events that occurred. Sometimes we can be fairly confident that the inferred tree is the true tree, but most phylogenetic data analyses are prone to uncertainties which are likely to result in the inferred tree differing in some respects from the true tree. The various methods used to assign degrees of confidence to the branching pattern in an inferred tree, and later in the chapter we will examine some of the controversies that have arisen as a result of the imprecise nature of phylogenetic analysis.

Gene trees are not the same as species trees

Gene tree, reconstructed from comparisons between the sequences of orthologous genes to make inferences about the evolutionary history of the species from which the genes are obtained. The assumption is that the gene tree, based on molecular data with all its advantages, will be a more accurate and less ambiguous representation of the species tree than that obtainable by morphological comparisons. This assumption is often correct, but it does not mean that the gene tree is the *same* as the species tree. For that to be the case, the internal nodes in the gene and species trees would have to be precisely equivalent. However, they are not equivalent, because:

- An *internal node in a gene tree* represents the divergence of an ancestral gene into two genes with different DNA sequences: this occurs by mutation.

- An *internal node in a species tree* represents a speciation event: this occurs by the population of the ancestral species splitting into two groups that are unable to interbreed, for example, because they are geographically isolated.

The important point is that these two events - mutation and speciation - are not expected to occur at the same time. For example, the mutation event could precede the speciation. This would mean that, to begin with, both alleles of the gene are present in the unsplit population of the ancestral species.

- If a molecular clock is used to date the time at which the gene divergence took place, then it cannot be assumed that this is also the time of the speciation event. If the node being dated is ancient, say 50 million or more years ago, then the error may not be noticeable. But if the speciation event is recent, as when primates are being compared, then the date for the gene divergence might be significantly different to that for the speciation event.
- If the first speciation event is quickly followed by a second speciation event in one of the two resulting populations, then the branching order of the gene tree might be different from that of the species tree. This can occur if the genes in the modern species are derived from alleles that had already appeared before the first of the two speciation events.

Tree reconstruction

Tree reconstruction is carried out with DNA sequences, concentrating on the four steps in the procedure:

- Aligning the DNA sequences and obtaining the comparative data that will be used to reconstruct the tree;
- Converting the comparative data into a reconstructed tree;
- Assessing the accuracy of the reconstructed tree;
- Using a molecular clock to assign dates to branch points within the tree.

Sequence alignment is the essential preliminary to tree reconstruction

The data used in reconstruction of a DNA-based phylogenetic tree are obtained by comparing nucleotide sequences. These comparisons are made by aligning the sequences so that nucleotide differences can be scored. This is the critical part of the entire enterprise because if the alignment is incorrect then the resulting tree will definitely not be the true tree.

The first issue to consider is whether the sequences being aligned are homologous. If they are homologous then they must, by definition, be derived from a common ancestral sequence and so there is a sound basis for the phylogenetic study. If they are not homologous then they do not share a common ancestor. The phylogenetic analysis will find a common ancestor because the methods used for tree reconstruction always produce a tree of some description, even if the data are completely erroneous, but the resulting tree will have no biological relevance. With some DNA sequences - for example, the β-globin genes of different vertebrates - there is no difficulty in being sure that the sequences being compared are homologous, but this is not always the case, and one of the commonest errors that arises during phylogenetic analysis is the inadvertent inclusion of a non-homologous sequence.

Once it has been established that two DNA sequences are indeed homologous, the next step is to align the sequences so that homologous nucleotides can be compared. With some pairs of sequences this is a trivial exercise, but it is not so easy if the sequences are relatively dissimilar and/or have diverged by the accumulation of insertions and deletions as well as point mutations. Insertions and deletions cannot be distinguished when pairs of sequences are compared so we refer to them as indels. Placing indels at their

correct positions is often the most difficult part of sequence alignment.

Some pairs of sequences can be aligned reliably by eye. For more complex pairs, alignment might be possible by the dot matrix method. The two sequences are written out on the x- and y-axes of a graph, and dots placed in the squares of the graph paper at positions corresponding to identical nucleotides in the two sequences. The alignment is indicated by a diagonal series of dots, broken by empty squares where the sequences have nucleotide differences, and shifting from one column to another at places where indels occur.

More rigorous mathematical approaches to sequence alignment have also been devised. The first of these is the similarity approach, which aims to maximize the number of matched nucleotides - those that are identical in the two sequences. The complementary approach is the distance method, in which the objective is to minimize the number of mismatches. Often the two procedures will identify the same alignment as being the best one.

Usually the comparison involves more than just two sequences, meaning that a multiple alignment is required. This can rarely be done effectively with pen and paper so, as in all steps in a phylogenetic analysis, a computer program is used. For multiple alignments, Clustal is often the most popular choice. Clustal and other software packages for phylogenetic analysis.

Converting alignment data into a phylogenetic tree

Once the sequences have been aligned accurately, an attempt can be made to reconstruct the phylogenetic tree. To date nobody has devised a perfect method for tree reconstruction, and several different procedures are used routinely. Comparative tests have been run with artificial data, for which the true tree is known, but these have failed to identify any particular method as being better than any of the others.

The main distinction between the different tree-building methods is the way in which the multiple sequence alignment is converted into numerical data that can be analyzed mathematically in order to reconstruct the tree. The simplest approach is to convert the sequence information into a distance matrix, which is simply a table showing the evolutionary distances between all pairs of sequences in the dataset. The evolutionary distance is calculated from the number of nucleotide differences between a pair of sequences and is used to establish the lengths of the branches connecting these two sequences in the reconstructed tree.

The neighbor-joining method is a popular tree-building procedure that uses the distance matrix approach. To begin the reconstruction, it is initially assumed that there is just one internal node from which branches leading to all the DNA sequences radiate in a star-like pattern. This is virtually impossible in evolutionary terms but the pattern is just a starting point. Next, a pair of sequences is chosen at random, removed from the star, and attached to a second internal node, connected by a branch to the center of the star. The distance matrix is then used to calculate the total branch length in this new 'tree'. The sequences are then returned to their original positions and another pair attached to the second internal node, and again the total branch length is calculated. This operation is repeated until all the possible pairs have been examined, enabling the combination that gives the tree with the shortest total branch length to be identified. This pair of sequences will be neighbors in the final tree; in the interim, they are combined into a single unit, creating a new star with one branch fewer than the original one. The whole process of pair selection and tree-length calculation is now repeated so that a second pair of neighboring sequences is identified, and then repeated again so that a third pair is located, and so on. The result is a complete reconstructed tree.

The advantage of the neighbor-joining method is that the data handling is relatively easy to carry out, largely because the information content of the multiple alignment has been reduced to its simplest form. The disadvantage is that some of the information is lost, in particular that pertaining to the identities of the ancestral and derived nucleotides at each position in the multiple alignment. The **maximum parsimony** method takes account of this information, utilizing it to recreate the series of nucleotide changes that resulted in the pattern of variation revealed by the multiple alignment. The assumption, possibly erroneous, is that evolution follows the shortest possible route and that the correct phylogenetic tree is therefore the one that requires the minimum number of nucleotide changes to produce the observed differences between the sequences. Trees are therefore constructed at random and the number of nucleotide changes that they involve calculated until all possible topologies have been examined and the one requiring the smallest number of steps identified. This is presented as the most likely inferred tree.

The maximum parsimony method is more rigorous in its approach compared with the neighbor-joining method, but this increase in rigor inevitably extends the amount of data handling that is involved. This is a significant problem because the number of possible trees that must be scrutinized increases rapidly as more sequences are added to the dataset. With just five sequences there are only 15 possible unrooted trees, but for ten sequences there are 2 027 025 unrooted trees and for 50 sequences the number exceeds the number of atoms in the universe. Even with a high-speed computer it is not possible to check every one of these trees in a reasonable time, if at all, so often the maximum parsimony method is unable to carry out a comprehensive analysis. The same is true with many of the other more sophisticated methods for tree reconstruction.

The Applications of Molecular Phylogenetics

Molecular phylogenetics has grown in stature since the start of the 1990s, largely because of the development of more rigorous methods for tree building, combined with the explosion of DNA sequence information obtained initially by PCR analysis and more recently by genome projects. The importance of molecular phylogenetics has also been enhanced by the successful application of tree reconstruction and other phylogenetic techniques to some of the more perplexing issues in biology.

Homologous sequences

Homologous sequences is the biological homology between protein or DNA sequences, defined in terms of shared ancestry in the evolutionary history of life. Two segments of DNA can have shared ancestry either because of a speciationevent (orthologs), or because of a duplication event (paralogs).

Homology among proteins or DNA is typically inferred from their sequence similarity. Significant similarity is strong evidence that two sequences are related by divergent evolution of a common ancestor. Alignments of multiple sequences are used to indicate which regions of each sequence are homologous.

The term "percent homology" is often used to mean "sequence similarity". The percentage of identical residues (*percent identity*) or the percentage of residues conserved with similar physicochemical properties (*percent similarity*), e.g. leucine and isoleucine, is usually used to "quantify the homology". Based on the definition of homology specified above this terminology is incorrect since sequence similarity is the observation, homology is the conclusion. Sequences are either homologous or not. As with anatomical structures, high sequence similarity might occur because of convergent evolution, or, as with shorter sequences, by chance, meaning that they are not homologous. Homologous sequence regions

are also called conserved. This is not to be confused with conservation in amino acid sequences, where the amino acid at a specific position has been substituted with a different one that has functionally equivalent physicochemical properties.

Orthology

Partial homology can occur where a segment of the compared sequences has a shared origin, while the rest does not. Such partial homology may result from a gene fusion event.

Homologous sequences are **orthologous** if they are inferred to be descended from the same ancestral sequence separated by a speciation event: when a species diverges into two separate species, the copies of a single gene in the two resulting species are said to be orthologous. **Orthologs**, or orthologous genes, are genes in different species that originated by vertical descent from a single gene of the last common ancestor. The term "ortholog" was coined in 1970 by the molecular evolutionist Walter Fitch.

For instance, the plant Flu regulatory protein is present both in *Arabidopsis* (multicellular higher plant) and *Chlamydomonas* (single cell green algae). The *Chlamydomonas* version is more complex: it crosses the membrane twice rather than once, contains additional domains and undergoes alternative splicing. However it can fully substitute the much simpler *Arabidopsis* protein, if transferred from algae to plant genome by means of genetic engineering. Significant sequence similarity and shared functional domains indicate that these two genes are orthologous genes, inherited from the shared ancestor.

Orthology is strictly defined in terms of ancestry. Given that the exact ancestry of genes in different organisms is difficult to ascertain due to gene duplication and genome rearrangement events, the strongest evidence that two similar genes are orthologous is usually found by carrying out phylogenetic analysis of the gene lineage. Orthologs often, but not always, have the same function.

Orthologous sequences provide useful information in taxonomic classification and phylogenetic studies of organisms.

Paralogy

Paralogous genes often belong to the same species, but this is not necessary: for example, the hemoglobin gene of humans and the myoglobin gene of chimpanzees are paralogs. Paralogs can be split into in-paralogs (paralogous pairs that arose after a speciation event) and out-paralogs (paralogous pairs that arose before a speciation event). Between-species out-paralogs are pairs of paralogs that exist between two organisms due to duplication before speciation, whereas within-species out-paralogs are pairs of paralogs that exist in the same organism, but whose duplication event happened before speciation. Paralogs typically have the same or similar function, but sometimes do not: due to lack of the original selective pressure upon one copy of the duplicated gene, this copy is free to mutate and acquire new functions.

Paralogous genes can shape the structure of whole genomes and thus explain genome evolution to a large extent. Examples include the Homeobox (Hox) genes in animals. These genes not only underwent gene duplications within chromosomes but also whole genome duplications. As a result Hox genes in most vertebrates are clustered across multiple chromosomes with the HoxA-D clusters being the best studied.[20]

Another example are the globin genes which encode myoglobin and hemoglobin are considered to be ancient paralogs. Similarly, the four known classes of hemoglobins (hemoglobin A, hemoglobin A2, hemoglobin B, and hemoglobin F) are paralogs of each other. While each of these proteins serves the same basic function of oxygen transport, they have already diverged slightly in function: fetal hemoglobin

(hemoglobin F) has a higher affinity for oxygen than adult hemoglobin. Function is not always conserved, however. Human angiogenin diverged from ribonuclease, for example, and while the two paralogs remain similar in tertiary structure, their functions within the cell are now quite different.

It is often asserted that orthologs are more functionally similar than paralogs of similar divergence, but several papers have challenged this notion.

Regulation. Paralogs are often regulated differently, e.g. by having different tissue-specific expression patterns. However, they can also be regulated differently on the protein level. For instance, *Bacillus subtilis* encodes two paralogues of glutamate dehydrogenase: GudB is constitutively transcribed whereas RocG is tightly regulated. In their active, oligomeric states, both enzymes show similar enzymatic rates. However, swaps of enzymes and promoters cause severe fitness losses, thus indicating promoter–enzyme coevolution. Characterization of the proteins shows that, compared to RocG, GudB's enzymatic activity is highly dependent on glutamate and pH.

Paralogous regions

Sometimes, large chromosomal regions share gene content similar to other chromosomal regions within the same genome. They are well characterized in the human genome, where they have been used as evidence to support the 2R hypothesis. Sets of duplicated, triplicated and quadruplicated genes, with the related genes on different chromosomes, are deduced to be remnants from genome or chromosomal duplications. A set of paralogy regions is together called a paralogon. Well-studied sets of paralogy regions include regions of human chromosome 2, 7, 12 and 17 containing Hox gene clusters, collagen genes, keratin genes and other duplicated genes, regions of human chromosomes 4, 5, 8 and 10 containing neuropeptide receptor genes, NK class homeobox genes and many more gene families, and parts of human chromosomes 13, 4, 5 and X containing the ParaHox genes and their neighbors. The Major histocompatibility complex (MHC) on human chromosome 6 has paralogy regions on chromosomes 1, 9 and 19. Much of the human genome seems to be assignable to paralogy regions

Ohnology

Ohnologous genes are paralogous genes that have originated by a process of whole-genome duplication. The name was first given in honour of Susumu Ohno by Ken Wolfe.[34]Ohnologues are useful for evolutionary analysis because all ohnologues in a genome have been diverging for the same length of time (since their common origin in the whole genome duplication).

Xenology

Homologs resulting from horizontal gene transfer between two organisms are termed xenologs. Xenologs can have different functions, if the new environment is vastly different for the horizontally moving gene. In general, though, xenologs typically have similar function in both organisms. The term was coined by Walter Fitch.

Gametology

Gametology denotes the relationship between homologous genes on non-recombining, opposite sex chromosomes. The term was coined by García-Moreno and Mindell in 2000. Gametologs result from the origination of genetic sex determination and barriers to recombination between sex chromosomes. Examples of gametologs include CHDW and CHDZ in birds.

Phenotypic evolution

The process of phenotype construction-including development of multicellular organisms-and the multiple

interactions and feedbacks between DNA, organism, and environment at various levels and timescales in the evolutionary process. During the construction of an individual's phenotype, DNA is recruited as a template for building blocks within the cellular context and may in addition be involved in dynamical feedback loops that depend on the environmental and organismal context. In the production of phenotypic variation among individuals, stochastic, environmental, genetic, and parental sources of variation act jointly. While in controlled laboratory settings, various genetic and environmental factors can be tested one at a time or in various combinations, they cannot be separated in natural populations because the environment is not controlled and the genotype can rarely be replicated. Along generations, genotype and environment each have specific properties concerning the origin of their variation, the hereditary transmission of this variation, and the evolutionary feedbacks. Natural selection acts as a feedback from phenotype and environment to genotype.

Speciation

According to the most widely used species definition, the **biological species concept**, a species is a group of organisms that can potentially interbreed, or mate, with one another to produce viable, fertile offspring. In this definition, members of the same species must have the potential to interbreed. However, that doesn't mean they have to be part of the same interbreeding group in real life. For instance, a dog living in Australia and a dog living in Africa are unlikely to meet but *could* have puppies if they did.

In order to be considered to be a single species in the biological species concept, a group of organisms must produce healthy, fertile offspring when they interbreed. In some case, organisms of different species can mate and produce healthy offspring, but the offspring are infertile, can't reproduce.

The biological species concept defines organisms as being, or not being, of the same species based on whether they can interbreed to make fertile offspring. But why is it that different species can't successfully interbreed? This question may seem silly for very different species (like a plant and an animal), but for others like the horse and the donkey above, it's much less obvious.

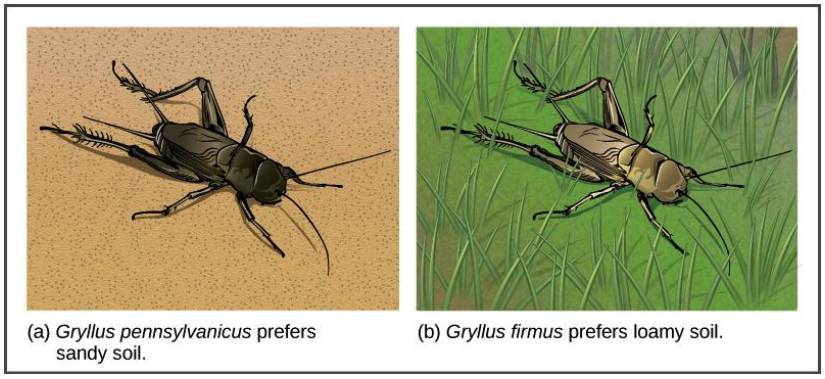
Broadly speaking, different species are unable to interbreed and produce healthy, fertile offspring due to barriers called **mechanisms of reproductive isolation**.

These barriers can be split into two categories based on when they act: prezygotic and postzygotic.

Prezygotic barriers

Prezygotic barriers prevent members of different species from mating to produce a zygote, a single-celled embryo. Some example scenarios are below:

- Two species might prefer different habitats and thus be unlikely to encounter one another. This is called **habitat isolation**.


(a) *Gryllus pennsylvanicus* prefers sandy soil.

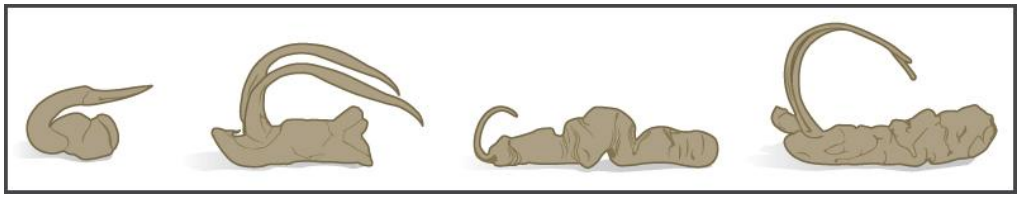
(b) *Gryllus firmus* prefers loamy soil.

(a) shows the black Gryllus pennsylvanicus cricket on sandy soil, and (b) shows the beige Gryllus firmus cricket in grass.

- Two species might reproduce at different times of the day or year and thus be unlikely to meet up when seeking mates. This is called **temporal isolation**.

(a) shows Rana aurora, a beige frog with green spots. (b) shows Rana boylii, a brown frog.

- Two species might have different courtship behaviors or mate preferences and thus find each other "unattractive". This is known as **behavioral isolation**.
- Two species might produce egg and sperm cells that can't combine in fertilization, even if they meet up through mating. This is known as **gametic isolation**.
  ^2start superscript, 2, end superscript
- Two species might have bodies or reproductive structures that simply don't fit together. This is called **mechanical isolation**



These are all examples of prezygotic barriers because they prevent a hybrid zygote from ever forming.

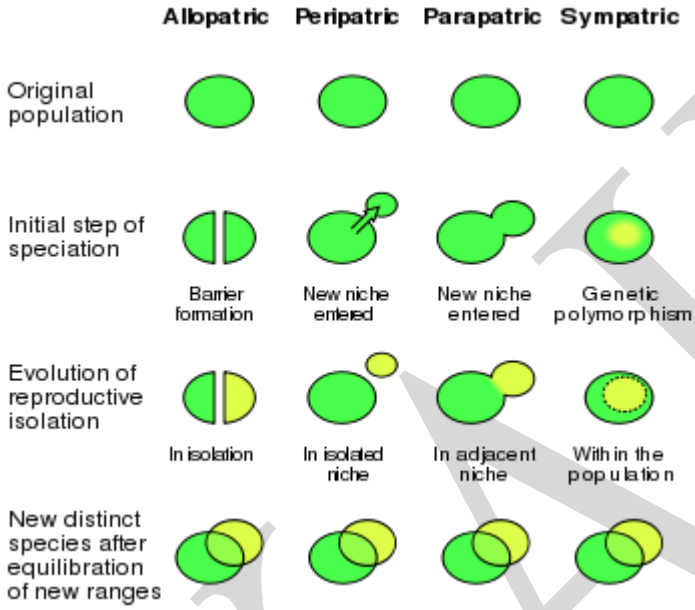
Postzygotic barriers

Postzygotic barriers keep hybrid zygotes—one-celled embryos with parents of two different species—from developing into healthy, fertile adults. Postzygotic barriers are often related to the hybrid embryo's mixed set of chromosomes, which may not match up correctly or carry a complete set of information. In some cases, the chromosomal mismatch is lethal to the embryo or results in an individual that can survive but is unhealthy. In other cases, a hybrid can survive to adulthood in good health but is infertile

because it can't split its mismatched chromosomes evenly into eggs and sperm. For example, this type of mismatch explains why mules are sterile, unable to reproduce

Types of Speciation

Speciation can take place in two general ways. A single species may change over time into a new form that is different enough to be considered a new species. This process is known as anagenesis. More commonly, a species may become split into two groups that no longer share the same gene pool. This process is known as cladogenesis. There are several ways in which anagenesis and cladogenesis may take place. In all cases, reproductive isolation occurs.



Sympatric Speciation

Sympatric speciation occurs when populations of a species that share the same habitat become reproductively isolated from each other. This speciation phenomenon most commonly occurs through polyploidy, in which an offspring or group of offspring will be produced with twice the normal number of chromosomes. Where a normal individual has two copies of each chromosome (diploidy), these offspring may have four copies (tetraploidy). A tetraploid individual cannot mate with a diploid individual, creating reproductive isolation.

Sympatric speciation is rare. It occurs more often among plants than animals, since it is so much easier for plants to self-fertilize than it is for animals. A tetraploidy plant can fertilize itself and create offspring. For a tetraploidy animal to reproduce, it must find another animal of the same species but of opposite sex that has also randomly undergone polyploidy.

Allopatric Speciation

Allopatric speciation, the most common form of speciation, occurs when populations of a species become geographically isolated. When populations become separated, gene flow between them ceases. Over time, the populations may become genetically different in response to the natural selection imposed by their different environments. If the populations are relatively small, they may experience a founder effect: the

populations may have contained different allelic frequencies when they were separated. Selection and genetic drift will act differently on these two different genetic backgrounds, creating genetic differences between the two new species.

Parapatric Speciation

Parapatric speciation is extremely rare. It occurs when populations are separated not by a geographical barrier, such as a body of water, but by an extreme change in habitat. While populations in these areas may interbreed, they often develop distinct characteristics and lifestyles. Reproductive isolation in these cases is not geographic but rather temporal or behavioral. For example, plants that live on boundaries between very distinct climates may flower at different times in response to their different environments, making them unable to interbreed.

POSSIBLE QUESTIONS

2 marks

1. State Hardy-Weinberg equilibrium.
2. Define QTL.
3. Define evolution.
4. Define phylogenetics.
5. Define complex trait with suitable example.
6. What is positive selection?
7. What is negative selection?
8. Define species.

8 Marks

1. Discuss on Inheritance of complex trait.
2. Elaborate on the Hardy-Weinberg law with predicting allele and genotype frequencies.
3. Elaborate on the analysis of nucleotide and aminoacid sequencing.
4. Write in detail about molecular phylogenesis.
5. Discuss on quantitative trait loci and their identification.
6. Write in detail about speciation.
7. How will you describe positive and negative selection in the mode of speciation? Give some examples.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
II-B.Sc., BIOCHEMISTRY
CONCEPTS IN GENETICS (17BCU304B)
MULTIPLE CHOICE QUESTIONS
UNIT V

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	The mold used by Beadle and Tatum in their series of experiments _____.	can normally produce all the enzymes it needs to grow on minimal medium	grew on minimal medium after treatment with X-rays	were unable to produce any enzymes after treatment with X-rays	all of the above	**can normally produce all the enzymes it needs to grow on minimal medium**
2	Since the X-ray treated mold was able to grow on media enriched with metabolites C and D of this metabolic pathway: 1 2 3 A -------->B -------> C -------> D where the numbers are enzymes and the letters are metabolites, Beadle and Tatum concluded that the mold lacked enzyme _____.	1	2	3	1 and 2	**2**
3	DNA specifies for the production of _____.	carbohydrates	nucleotides	proteins	lipids	**Proteins**
4	Proteins are synthesized _____.	at the SER	at the ribosomes	in the lysosomes	in the nucleus	**at the ribosomes**
5	Which of the following is confined to the nucleus in eukaryotes?	DNA	RNA	Proteins	tRNA	**DNA**
6	RNA is a _____.	polymer	monomer	trimer	tetramer	**Polymer**
7	RNA nucleotides contain _____.	a phosphate group	the sugar ribose	the base uracil	all of the above	**all of the above**
8	Which of the following is not found in RNA?	cytosine	guanine	thymine	adenine	**Thymine**
9	There is/are _____ class(es) of RNA.	one	two	three	four	**Three**
10	During transcription DNA serves as a template for _____ synthesis.	mRNA	tRNA	rRNA	DNA	**mRNA**
11	Amino acids are carried to the ribosomes by _____.	mRNA	tRNA	rRNA	DNA	**rRNA**
12	A piece of mRNA directs the sequence of amino acids in a polypeptide during _____.	replication	transcription	translation	transduction	**Translation**
13	Each codon is made up of _____ base(s).	one	two	three	four	**Three**
14	The order of amino acids in a polypeptide is specified by _____.	a random set of nucleotides	an exact sequence of DNA nucleotides	DNA polymerase	RNA polymerase	**an exact sequence of DNA nucleotides**
15	The genetic code _____.	is degenerate	unambiguous	has only one start signal	all of the above	**all of the above**
16	The potentially harmful effects of mutations are minimized by the _____ of the genetic code.	degeneracy	unambiguity	three stop signals	one start signal	**Degeneracy**

17	Protein synthesis involves _____.	initiation	elongation	termination	all of the above	**all of the above**
18	The three bases on the tRNA molecule that are complementary to the triplet of mRNA are referred to as a _____	anticodon	codon	exon	intron	**Anticodon**
19	The amino acid alanine would be carried by a tRNA with the anticodon _____.	CGC	CGU	CGA	all of the above	**all of the above**
20	A cDNA library:	Can also be called an expressed sequence tag (EST) library.	Consists of coding sequences from genes that are expressed.	Is specific to the set of conditions under which the original mRNA was generated.	All of these.	**All of these.**
21	The collection of proteins that can be produced by a given species is:	Considered that species' genetic complement.	Correlated with the size of the organism.	Called the proteome.	All of these.	**Called the proteome.**
22	Which of these questions could be answered using subtractive hybridization?	What genes are required for basic cellular functions?	What site is bound by a kidney specific transcription factor?	What genes are expressed in cardiac but not skeletal muscle?	All of these	**What genes are expressed in cardiac but not skeletal muscle?**
23	A _____ analysis is the study of groups of genes that seem to be regulated together.	cluster	multiple	base pair	revelant	**Cluster**
24	The two most common processes that lead to production of multiple functional proteins from the same DNA sequence are:	RNA editing and alternative splicing.	Protein folding and posttranslational covalent modifications.	Alternative splicing and posttranslational covalent modifications.	Posttranslational covalent modification and transcriptional regulation.	**Alternative splicing and posttranslational covalent modifications.**
25	Which of the following statements regarding the proteome is the most correct?	A large proportion of the proteome is expressed by each cell of a species.	Levels of gene expression at the mRNA level generally correlate highly with levels of functional protein.	The proteins produced by a specific cell depend on cell type and environmental conditions.	All of the above.	**The proteins produced by a specific cell depend on cell type and environmental conditions.**
26	If computers were not able to access the entire genomic sequence of an organism, which of the following techniques might allow determination of the gene sequence that encodes a particular protein?	PCR amplification of related gene sequences.	Hybridization of a genomic library with a degenerate probe.	Production of synthetic peptides.	All of these.	**Hybridization of a genomic library with a degenerate probe.**
27	Which of these would not be an example of sequence element?	A recognition site for binding of a particular transcription factor.	A restriction endonuclease cut site.	An open reading frame.	A stop codon.	**An open reading frame.**
28	In an analysis of eukaryotic gene, you identify several nonoverlapping open reading frames, but they are not all in the same frame. Which explanation makes the most sense?	By random chance, a second reading frame within the gene also	This gene includes introns which are not multiples of three.	This is a mutant allele that has had several small	All of these.	**This gene includes introns which are not multiples of three.**

		has an open reading frame.		insertions.		
29	Homologous genes:	Would be expected to have very similar sequences in related organisms.	Would be expected to be more similar in distantly related organisms than in organisms that are closely related.	May have become similar to each other by random mutation.	All of these.	**Would be expected to have very similar sequences in related organisms.**
30	You clone a novel gene from your favorite experimental organism, and identify the sequence of the gene itself and the protein. You would like to know if this gene is regulated by any known transcription factors. Which technique would be best suited to provide you evidence to address this question?	BLAST analysis of the protein sequence.	BLAST analysis of the DNA from the coding sequence.	BLAST analysis of the DNA from the upstream regions.	None of these.	**None of these.**
31	Secondary structure of RNA molecules:	Depends on complementary base pairing.	Is generated by covalent bonding between sections of the RNA molecule.	Can be described as interactions between portions of the backbone of the molecule.	Does not have an impact on function of the molecule.	**Depends on complementary base pairing.**
32	How many potential open reading frames are present in a DNA sequence?	one	three	six	more than six	**Six**
33	Which of the following pieces of information would not be required for a computer program to determine exon/intron structure of a gene?	A table correlating codons to amino acids.	The nucleotide sequence of the mRNa.	Splice site signals for the species under study.	All of these pieces of information would be required	**All of these pieces of information would be required**
34	During chain elongation amino acids are attached by _____ to the polypeptide being formed	hydrolysis	peptide bonds	hydrogen bonds	tight junctions	**peptide bonds**
35	The belief that RNA could have served as both the genetic material and as the first enzymes in the earliest living organisms was supported by the discovery of _____.	promoters	introns	ribozymes	spliceosomes	**Ribozymes**
36	Inheritance of complex traits differs from inheritance of quantitative traits in which of the following ways?	Traits run in families	No clear Mendelian pattern of inheritance is observed	More than one gene is involved	There is no difference.	**No clear Mendelian pattern of inheritance is observed**
37	Allele and genotype frequencies in a population will remain constant from generation to generation in absence of other evolutionary influences state law	of evolution	of replication	of genetic variation	Hardy-Weinberg	**Hardy-Weinberg**
38	Hardy-Weinberg proportions are	p^2	q^2	2pq	all of above	**all of above**
39	Red-green color blindness trait in Western- European males affects about	1 in 12	1 in 12	3 in 12	4 in 12	**1 in 12**

40	The random abnormal number of chromosomes in the animals is called:	Polyploidy	Euploid	aneuploid	None	**aneuploid**
41	Inheritance of complex traits differs from inheritance of quantitative traits in which of the following ways?	Traits run in families	No clear Mendelian pattern of inheritance is observed	More than one gene is involved	There is no difference.	**No clear Mendelian pattern of inheritance is observed**
42	If union of gametes to produce next generation is random, then allele frequencies between generation are	increasing	decreasing	constant	zero	**constant**
43	Assumptions underlying Hardy- Weinberg equilibrium are	organisms are diploid	mating is random	population size is infinitely large	all of above	**all of above**
44	The interchange of parts between non-homologous chromosomes is called:	Duplication	translocation	Inversion	Deletion	**translocation**
45	Which one is used for knowing whether or not a population is evolving?	Degree of evolution	Genetic drift	Proportion between acquired variations	Hardy Weinberg equation	**Hardy Weinberg equation**
46	The evolution of a species is based upon sum total of adaptive changes preserved by	natural selection	speciation	human conseravation	isolation	**isolation**
47	Genes are linked if their loci are	far from each other	nearer to each other	lie on same chromosome arm	none of above	**nearer to each other**
48	Fill in the missing portion of this flow diagram: DNA's nucleotide sequence --> amino acid sequence --> proteins' polypeptides --> _____ --> an organism's structures.	steroids	enzymes	polysaccharides	RNA	**enzymes**

49	The genotype of an organism is linked to the organism's phenotype by _____.	lipids	water	proteins	carbohydrates	**proteins**
50	The relationship between inheritance and metabolic diseases is suggested by the phrase "_____."	inborn error of metabolism	in number error of metabolism	unborn error of metabolism	inborn problem of metabolism	**inborn error of metabolism**

Reg. No. _____

KARPAGAM ACADEMY OF HIGHER EDUCATION
COIMBATORE – 21
DEPARTMENT OF BIOCHEMISTRY
II B.Sc., BIOCHEMISTRY
Thrid Semester
First Internal Examination - July 2018
CONCEPTS IN GENETICS (17BCU304-B)

Date : .07.2018
Time: 2 hours **Maximum: 50 marks**

PART A **(20x1=20 marks)**
Answer ALL the questions

1. Mendel's experimental model was
   a) *Pisum sativum* b) *Lathyrus odaratus*
   c) *Oryza sativa* d) *Mirabilis jalappa*
2. The phenomenon of 'like begets like' is due to
   a) genetics b) heredity
   c) germplasm d) variation
3. Genotype is the
   a) genetic constitution b) physical expression
   c) trait expressed d) genes
4. The physical expression or appearance of a character is called as
   a) morphology b) geneoype
   c) phenotype d) ecotype
5. Johansen replaced the term ……. for factors
   a) chromosome b) gene
   c)allele d) trait
6. Number of characters studied by Mendel in Pea was
   (a) 5 (b) 7 (c) 6 (d) 4
7. The geometrical device that helps to find out all the possible combinations of male and female gametes is called
   a) Punnete square b) Bateson square
   c) Mendel square d) Morgan square
8. Monohybrid F2 generation ratio is
   a) 3:1 b) 2:1 c) 1:1 d) 9:3:3:1
9. 5. The phenotypic dihybrid ratio is
   a) 3:1 b) 2:1 c) 1:1 d) 9:3:3:1
10. During experiments, Mendel called genes by the term
    a) factors b) traits c) characters d) qualities
11. The title of Mendel's paper while presenting at Brunn Natural History Society in 1865 was
    a) Laws of inheritance b) Laws of heredity
    c) Experiments on pea plants d) Experiments in plant hybridization
12. Offspring generation of P1 (parental generation) is classified as
    a) F2 generation b) G1 generation
    c) F1 generation d) P2 generation
13. Law which states that in F1 generation only dominant character is expressed is
    a) law of segregation b) law of assortment
    c) law of dominance d) law of dependent assortment
14. The scientist(s) who identified chromosome as the carrier of genetic material is
    a) Sutton and boveri b) Mendel
    c) Johansen d) Correns
15. Boveri conducted his experiments on
    a) grasshopper b) sea urchins
    c) pea plant d) drosophila
16. Number of chromosomes present in human
    a) 23 pairs b) 33 pairs c) 26 pairs d) 42pairs
17. If the genotype consists of only one type of allele. It is called
    a) homozygous b) heterozygous c) monoallelic d) unialleli
18. Which of the rules in probability is ued for two mutually exclusive events
    a) sum rule b) multiplication rule
    c) binomial expansion d) multinomial expansion
19. The alternate forms of a gene is called
    a) recessive character b) dominant character
    c) alleles d) chromosome
20. Changes in gene due to various reason is called
    a) abnormality b) heredity
    c)mutation d) inheritance

PART B **(3x2=6 marks)**

Answer ALL the questions

21. Define heredity.
22. Define law of segregation.
23. Name any four characteristic features of a model organism.

PART C **(3x8=24 marks)**

Answer ALL the questions

24. a) Discuss in detail about the model organisms used in genetic study

or

b) Give a note on laws of probability and binomial expansion

25. a) Explain the basic principles of heredity using Mendel's experiments.

or

b) Write in detail about alleles.

26.a) Elaborate on Sutton-Boveri hypothesis and experiments.

or

b) Discuss in detail about multiple allele with suitable examples.

****