

2 Edition

Core Concepts in Biology:

GENETICS

Kate Fickas

INCLUDES

Access code to enrich learning experience with interactive e-lectures, quiz, videos and much more

BIBLIOTEX Digital Library

CORE CONCEPTS IN BIOLOGY:

GENETICS

Kate Fickas

2nd Edition

BIBLIOTEX
Digital Library

www.bibliotex.com

CORE CONCEPTS IN BIOLOGY: GENETICS

2nd Edition

Kate Fickas

BIBLIOTEX
Digital Library

www.bibliotex.com

email: info@bibliotex.com

e-book Edition 2024

ISBN: 978-1-98469-144-6 (e-book)

This book contains information obtained from highly regarded resources. Reprinted material sources are indicated. Copyright for individual articles remains with the authors as indicated and published under Creative Commons License. A Wide variety of references are listed. Reasonable efforts have been made to publish reliable data and views articulated in the chapters are those of the individual contributors, and not necessarily those of the editors or publishers. Editors or publishers are not responsible for the accuracy of the information in the published chapters or consequences of their use. The publisher assumes no responsibility for any damage or grievance to the persons or property arising out of the use of any materials, instructions, methods or thoughts in the book. The editors and the publisher have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission has not been obtained. If any copyright holder has not been acknowledged, please write to us so we may rectify.

Notice: Registered trademark of products or corporate names are used only for explanation and identification without intent of infringement.

© 2024 3G E-learning LLC

In Collaboration with 3G E-Learning LLC. Originally Published in printed book format by 3G E-Learning LLC with ISBN 978-1-98468-927-6

EDITORIAL BOARD



Aleksandar Mratinković was born on May 5, 1988. in Arandjelovac, Serbia. He has graduated on Economic high school (2007), The College of Tourism in Belgrade (2013), and also has a master degree of Psychology (Faculty of Philosophy, University of Novi Sad). He has been engaged in different fields of psychology (Developmental Psychology, Clinical Psychology, Educational Psychology and Industrial Psychology) and has published several scientific works



Dan Piestun (PhD) is currently a startup entrepreneur in Israel working on the interface of Agriculture and Biomedical Sciences and was formerly president-CEO of the National Institute of Agricultural Research (INIA) in Uruguay. Dan is a widely published scientist who has received many honours during his career including being a two-time recipient of the Amit Golda Meir Prize from the Hebrew University of Jerusalem, his areas of expertise includes stem cell molecular biology, plant and animal genetics and bioinformatics. Dan's passion for applied science and technological solutions did not stop him from pursuing a deep connection to the farmer, his family and nature. Among some of his interest and practices counts enjoying working as a beekeeper and onboard fishing.



Felecia Killings is the Founder and CEO of LiyahAmore Publishing, a publishing company committed to providing technical and educational services and products to Christian Authors. She operates as the Senior Editor and Writer, the Senior Writing Coach, the Content Marketing Specialist, Editor-in-Chief to the company's quarterly magazine, the Executive and Host of an international virtual network, and the Executive Director of the company's online school for Authors. She is a former high-school English instructor and professional development professor. She possesses a Master of Arts degree in Education and a Bachelor's degree in English and African American studies.



Dr. Sandra El Hajj, Ph.D. in Health Sciences from Nova Southeastern University, Florida, USA is a health professional specialized in Preventive and Global Health. With her 12 years of education obtained from one of the most prominent universities in Beirut, in addition to two leading universities in the State of Florida (USA), Dr. Sandra made sure to incorporate interdisciplinary and multicultural approaches in her work. Her long years of studies helped her create her own miniature world of knowledge linking together the healthcare field with Medical Research, Statistics, Food Technology, Environmental & Occupational Health, Preventive Health and most noteworthy her precious last degree of Global Health. Till today, she is the first and only doctor specialized in Global Health in the Middle East area.



Hazem Shawky Fouda has a PhD. in Agriculture Sciences, obtained his PhD. From the Faculty of Agriculture, Alexandria University in 2008, He is working in Cotton Arbitration & Testing General Organization (CATGO).



Fozia Parveen has a Dphil in Sustainable Water Engineering from the University of Oxford. Prior to this she has received MS in Environmental Sciences from National University of Science and Technology (NUST), Islamabad Pakistan and BS in Environmental Sciences from Fatima Jinnah Women University (FJWU), Rawalpindi.



Igor Krunic 2003-2007 in the School of Economics. After graduating in 2007, he went on to study at The College of Tourism, at the University of Belgrade where he got his bachelor degree in 2010. He was active as a third-year student representative in the student parliament.Then he went on the Faculty of science, at the University of Novi Sad where he successfully defended his master's thesis in 2013. The crown of his study was the work titled Opportunities for development of cultural tourism in Cacak". Later on, he became part of a multinational company where he got promoted to a deputy director of logistic. Nowadays he is a consultant and writer of academic subjects in the field of tourism.



Dr. Jovan Pehcevski obtained his PhD in Computer Science from RMIT University in Melbourne, Australia in 2007. His research interests include big data, business intelligence and predictive analytics, data and information science, information retrieval, XML, web services and service-oriented architectures, and relational and NoSQL database systems. He has published over 30 journal and conference papers and he also serves as a journal and conference reviewer. He is currently working as a Dean and Associate Professor at European University in Skopje, Macedonia.



Dr. Tanjina Nur finished her PhD in Civil and Environmental Engineering in 2014 from University of Technology Sydney (UTS). Now she is working as Post-Doctoral Researcher in the Centre for Technology in Water and Wastewater (CTWW) and published about eight International journal papers with 80 citations. Her research interest is wastewater treatment technology using adsorption process.



Stephen obtained his PhD from the University of North Carolina at Charlotte in 2013 where his graduate research focused on cancer immunology and the tumor microenvironment. He received postdoctoral training in regenerative and translational medicine, specifically gastrointestinal tissue engineering, at the Wake Forest Institute of Regenerative Medicine. Currently, Stephen is an instructor for anatomy and physiology and biology at Forsyth Technical Community College.



Michelle holds a Masters of Business Administration from the University of Phoenix, with a concentration in Human Resources Management. She is a professional author and has had numerous articles published in the Henry County Times and has written and revised several employee handbooks for various YMCA organizations throughout the United States.

Table of Contents

Preface *ix*

1 Introduction to Genetics 1

Introduction 2
Core Concept: Principles of Genetics 4
Core Concept: Linkage and Crossing Over 14
Summary 35
Multiple Choice Questions 36
Review Questions 37
References 38

2 Principles of Genetic Transmission 39

Introduction 40
Core Concept: Mendel's Experiments 41
Core Concept: Segregation 52
Core Concept: Dominance 57
Core Concept: Nomenclature 59
Core Concept: Multiple Alleles 67
Core Concept: Biochemical Genetics 77
Summary 82
Multiple Choice Questions 83
Review Questions 84
References 85

3 Population Genetics 87

Introduction 88
Core Concept: Factor of Population Genetics 89
Core Concept: Problems of Population Genetics 104

Summary 122
Multiple Choice Questions 123
Review Questions 124
References 125

4 Mitosis and Meiosis 127

Introduction 128
Core Concept: Mitosis 129
Core Concept: Meiosis 140
Summary 162
Multiple Choice Questions 163
Review Questions 164
References 165

5 DNA Structure and Analysis 167

Introduction 168
Core Concept: Properties of DNA 169
Core Concept: DNA Structure 171
Core Concept: DNA Functions 183
Summary 190
Multiple Choice Questions 191
Review Questions 192
References 193

6 Linkage and Chromosome Mapping 195

Introduction 196
Core Concept: Genetic Linkage 197
Core Concept: Chromosome Mapping 216
Summary 226
Multiple Choice Questions 227
Review Questions 228
References 229

7 Genetic Laboratory Techniques 231

Introduction 232
Core Concept: Culturing Drosophila Stock 233
Core Concept: Recombinant DNA Techniques Cloning 237
Core Concept: Blood Typing 244
Core Concept: DNA Fingerprinting 248
Core Concept: Cryogenic Preservation of Animal Germplasm 252
Summary 261
Multiple Choice Questions 262
Review Questions 263
References 264

Index 265

Preface

Genetics refers to the study of heredity and variation in living organisms. It has captivated scientists and enthusiasts alike for centuries. It offers a fascinating glimpse into how traits and characteristics are passed down from one generation to the next, shedding light on the fundamental building blocks of life itself. Genetics plays a crucial role in the development of society. Understanding how genes shape our traits, behaviors, and susceptibility to diseases opens up new pathways for medical advancements and personalized treatments. The field of genetics is vast and ever-evolving, comprising several key areas including molecular genetics, population genetics, and genetic engineering. At the core of genetic research lies the study of DNA - the blueprint of life. Examining the structure, function, and expression of genes enables scientists to comprehend their influence on human development and inheritance patterns. By deciphering these complex genetic codes, researchers are able to diagnose various diseases at an early stage when interventions can be more effective. Furthermore, insights derived from population genetics provide valuable information about our evolutionary history and diversity within different groups. This knowledge promotes inclusivity by highlighting the importance of preserving genetic variety for future generations. Additionally, it aids in understanding how populations adapt to environmental changes over time. The concept of genetic engineering has revolutionized agriculture by enhancing crop yields through manipulation of plant genes for improved characteristics such as disease resistance or enhanced nutrition content. These innovations have profound implications for food security worldwide as agriculture becomes increasingly critical due to rapid global population growth. Moreover, advancements in pharmacogenomics allow tailor-made medications based on individual variations in gene expressions or mutations. This approach optimizes drug therapies while minimizing adverse side effects through a better understanding of each person's unique genomic profile.

In this second edition of Core Concepts in Genetics, we embark on an intellectual journey through the intricate world of genetic principles. The overarching goal of this book is to provide readers with a comprehensive understanding of genetics, drawing upon both classical theories and modern advancements. By delving into concepts such as DNA replication, gene expression, Mendelian inheritance patterns, genetic engineering techniques, and molecular genetics research methods – among others – we aim to equip individuals with a thorough grasp of these indispensable foundations.

Organized into concise chapters carefully structured for educational efficacy, Core Concepts in Genetics serves as an invaluable resource not only for students pursuing degrees or careers related to biological sciences but also for anyone seeking a deeper comprehension

of our genetic blueprint. Each chapter seamlessly blends critical concepts with illustrative examples and thought-provoking exercises designed to enhance understanding while honing analytical skills.

Moreover, this revised edition incorporates up-to-date content reflecting recent breakthroughs that have reshaped our understanding of genetics. The constantly evolving nature of this field necessitates avoiding complacency when it comes to knowledge acquisition; therefore seizing upon cutting-edge developments allows us to fortify students' grasp on contemporary genetic principles.

CHAPTER 1

INTRODUCTION TO GENETICS



Genetics is the scientific study of inherited variation. Human genetics is the scientific study of inherited human variation. This field has been energized in recent years by the Human Genome Project. Scientists expect that the project will lead to the development of new drugs targeted to specific genetic disorders. Increasingly, modern genetics involves genetic engineering; a technique used to manipulate genes and has produced many advances in medicine.

After studying this chapter, you will understand the core concept of:

- Principles of Genetics
- Linkage and Crossing Over

INTRODUCTION

Genetics is a captivating and crucial branch of biology that explores the inheritance and variation of traits in living organisms. It delves into the mechanisms by which genetic information is passed from one generation to the next, shaping the characteristics and diversity of all living beings.

Genes are made from a long molecule called DNA, which is copied and inherited across generations. DNA is made of simple units that line up in a particular order within it, carrying genetic information. The language used by DNA is called genetic code, which lets organisms read the information in the genes. This information is the instructions for the construction and operation of a living organism.

Genetics arose out of the identification of genes, the fundamental units responsible for heredity. Genetics may be defined as the study of genes at all levels, including the ways in which they act in the cell and the ways in which they are transmitted from parents to offspring. Modern genetics focuses on the chemical substance that genes are made of, called deoxyribonucleic acid, or DNA, and the ways in which it affects the chemical reactions that constitute the living processes within the cell. Gene action depends on interaction with the environment. Green plants, for example, have genes containing the information necessary to synthesize the photosynthetic pigment chlorophyll that gives them their green colour. Chlorophyll is synthesized in an environment containing light because the gene for chlorophyll is expressed only when it interacts with light. If a plant is placed in a dark environment, chlorophyll synthesis stops because the gene is no longer expressed.

Central to genetics is the concept of genes, the hereditary units that encode specific traits and characteristics within an organism's DNA. Genes are composed of DNA sequences and serve as the instructions for building proteins, which play vital roles in the structure and functioning of cells and organisms. The arrangement and sequence of genes along the DNA molecule create the genetic code unique to each individual, defining traits like eye color, height, and susceptibility to certain diseases.

Inheritance, the passing of genetic information from parents to offspring, occurs during sexual reproduction. Each parent contributes half of their genetic material to form the offspring's complete genetic makeup. The process of meiosis ensures the reduction of the chromosome number in sex cells (gametes) to half, ensuring genetic diversity and variation in offspring.

The field of genetics is vast and diverse, encompassing various sub-disciplines. Classical genetics focuses on the principles of inheritance, while molecular genetics explores the structure, function, and regulation of genes at the molecular level. Population genetics studies the genetic composition of populations and how genetic variation changes over time.

Genetics is the study of biological inheritance. Geneticists examine a wide variety of inherited traits, from the ability to bear large numbers of fruit in trees, to eye color in mammals. This discipline is a major cornerstone of the larger field of biology.

People who study it do so in a wide variety of situations, from research laboratories to pharmaceutical companies, and new breakthroughs are consistently being made.



Humans have understood genetics on some level for thousands of years. They have been improving crops and animals through breeding for quite some time, selecting desirable traits and attempting to propagate them. Casual observations have also been an important part of human life for thousands of years, with people noting that members of certain families tend to have distinctive anatomical features. The Hapsburgs of Austria, for example, are famous for their prognathic jaws. In the 1800s, a monk named Gregor Mendel began doing controlled experiments with peas which were designed to unravel the mysteries of gene inheritance. He is often regarded as the father of modern genetics, because although he was not able to identify individual genes, he did demonstrate that a combination of dominant and recessive traits determine the physical appearance of an organism. In the 1950s, research finally revealed that DNA in the chromosomes carried the genetic code of the host organism.



With the confirmation that DNA held the secret to inheritance, people also started charting the genetic code, finding the specific locations on the chromosomes where certain genes appeared. Along the way, they learned a great deal about DNA and inheritance, and the DNA code is continually being studied to identify new genes. Understanding this coding has allowed researchers to manipulate it, removing or altering sections to create a specific desired outcome, such as a crop that is resistant

to pests. Knowledge of the human genetic code has also led to gene therapy, in which specific problems are addressed or prevented by manipulating the code of an individual. The study of this field does not just explain the manifestation of physical traits like hair color. Genetics is also used to explore inherited disease, and to determine whether or not people are more susceptible to certain diseases because of their genes. The discipline also involves what happens when things go wrong, as in the case of an organism that inherits an extra chromosome. Some people have also theorized that genes may play a role in behavior, with certain people being predisposed to depression, addiction, and other behavioral issues.



CORE CONCEPT: PRINCIPLES OF GENETICS

It is common observation that seeds of mango trees grow into mango plants, dogs give birth to puppies and not into the young ones of any other animal. Humans give birth to human beings. The tendency of offspring to inherit parental characteristics is termed 'heredity' and the science of heredity in 'Genetics'. Genetics also seeks to answer questions like why two offspring of same parents look different, why some people dark and others are light skinned etc. In other words, why is there variation among individuals of the same kind? This section deals with heredity and variation. It also includes a section on hereditary disorders and gives an idea of the human genome too.

Mendel's studies have provided scientists with the basis for mathematically predicting the probabilities of genotypes and phenotypes in the offspring of a genetic cross. But not all genetic observations can be explained and predicted based on Mendelian genetics. Other complex and distinct genetic phenomena may also occur. Several complex genetic concepts, described in this section, explain such distinct genetic phenomena as blood types and skin color.

Heredity and Variation

Whenever an infant is born in a family, the relatives begin to wonder about the resemblance of the infant's eyes, facial features, complexion, and color of hair with those of the parents, siblings and grandparents. The source of such resemblances and differences are in the "genes" that are passed down from parents to children and so on

generation after generation. This inheritance of genes is termed 'heredity' the study of **heredity** is 'Genetics'. New individual develop according to the genes inherited from their parents.

Heredity is the passing on of traits from parents to their offspring, either through asexual reproduction or sexual reproduction, the offspring cells or organisms acquire the genetic information of their parents.



Variation in a population is very important. It has survival value for the population. This is because if the environment changes, some individuals (variants) may be able to adapt to new situations and save the population from dying out. Variation arises due to mutation or sudden change in the genes. Variation also arises because genes get shifted and exchanged during meiosis at formation of gametes, giving rise to new gene combinations. At fertilization, there is random mixing of parental chromosomes with different gene combinations. Such a source of variation which is most common is called recombination.

Mendel's Experiments on the Garden Pea and Principles of Inheritance

Sir Gregor Johann Mendel (1822 to 1884) was Austrian monk who used garden pea (Pisum sativum) for his experiments and published his results in 1865. His work, however, was rediscovered in 1900, long after Mendel's death, by Tschermak, Correns and DeVries. But since Mendel was the first to suggest principles underlying inheritance he is regarded as the founder or father of genetics.

Mendel designed his experiments such that a pure tall variety of pea plants could be crossed to a pure dwarf variety. The anthers from flowers of tall plants were removed and their stigmas dusted with pollen from flowers of dwarf plants. The reverse experiment was also carried out. In the following spring, seeds from the new plants were collected and sown. He found that all the plants of this generation called first filial

generation or F_1 grew to be tall plants. He allowed them to self-pollinate. Again he collected the seeds. The following year, after the seeds had been sown, he found that three quarters of these plants were tall and the rest dwarf. He repeated the experiment several times and found that the ratio of tall to dwarf plants was 3: 1 (Figure 1). In this way he tried to cross pea plants with seven such contrasting characters.



These were

- red flowered and white flowered plants;
- axillary flowered (flower arising in the axial of the leaf) and terminal flowered (flower arising at tip of stalk);
- yellow seeded versus green seeded;
- round seeded verses wrinkled seeded;
- green pod versus yellow pod
- plants with inflated pods versus those with constricted pod and
- pure tall plants versus pure dwarf plants.

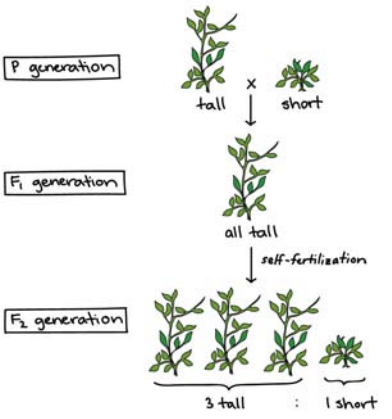


Figure 1: Mendel's experiment with fall and dwarf peas.

Plants with such contrasting characters exist in varieties that are 'self-pollinating' so that generation after generation they express only one type of feature (Figure 2). Crosses considering the inheritance of one feature only are called monohybrid crosses.

Mendel also tried crosses involving two contrasting features, such as tall and red flowered with dwarf and white flowered plant such a cross is termed dihybrid cross

Mendel's Principles (laws) of Inheritance

Basing on the results of his experiments, Mendel postulated the following laws of heredity

Law of Segregation or Purity of Gametes

At formation of gametes, the two chromosomes of each pair separate (segregate) into two different cell which form the **gametes**. This is a universal law and always during gamete formation in all sexually reproducing organisms, the two factors of a pair pass into different gametes. Each gamete receives one member of a pair of factors and the gametes are pure.

Gamete is a haploid cell that fuses with another haploid cell during fertilization (conception) in organisms that sexually reproduce.

Trait	Dominant vs. recessive	F_2 generations Dominant form	Recessive form	Ratio
Flower color	Purple x White	705	224	3.15:1
Seed color	Yellow x Green	6022	2001	3.01:1
Seed shape	Round x Wrinkled	5474	1850	2.96:1
Pod color	Green x Yellow	428	152	2.82:1
Pod shape	Round x Constricted	882	299	2.95:1
Flower position	Axial x Top	651	207	3.14:1
Plant height	Tall x Dwarf	787	277	2.84:1

Figure 2: Seven traits studies by Mendel.

Law of Dominance

Inheritance of many features (e.g. eye color, flower color, seed shape etc.) is controlled by one pair of genes. When the two genes are of the same kind (e.g. brown color of eyes, red color of flower) the condition is termed homozygous. When a pair of chromosomes has the gene controlling the same feature (flower color) in two different forms (red flower gene on one chromosome and white flower gene on its pair (termed its homologue,) the condition is termed heterozygous. The second law of inheritance maintains that of the two contrasting

genes of a pair, the expression of one is dominant over that of the other. Thus if both genes are for tallness (represented on TT) that is homozygous or for tallness and one for dwarfness (Tt), the pea plants will be tall. The opposite of dominant gene is termed recessive gene. The recessive feature (e.g. dwarfness of the plant) shows up only when its gene is in the homozygous condition (tt).

> The transmission of characters from one generation to the next, that is from parents to offsprings is known .as heredity

The law of dominance was found to be true in both monohybrid and dihybrid crosses in cases of the characteristics studied by Mendel in the garden pea. Law of independent assortment meaning whereby that in the inheritance of two features (each feature controlled by a pair of genes), genes for the two different features are passed down into the offspring independently (Figure 3) i.e. the segregation of one pair of factors is independent of the segregation of the other.

	Red Tall		White Dwarf	
Parents	TTRR	×	ttrr	
gametes	TR		tr	
F_1	Tt		Rr	Tall, red (self)

F_2 progeny worked out below.

Genes in male and female gametes	TR	Tr	tR	tr
TR	TTRR Tall red	TTRr Tall red	TtRR Tall red	TtRr Tall red
Tr	TTRr Tall red	TTrr Tall white	TtRr Tall red	Ttrr Tall white
tR	TtRr Tall red	TtRr Tall red	ttRR White red	ttRr White red
tr	TtRr Tall red	TtRr Tall red	ttRr Dwarf red	ttrr Dwarf white

Figure 3: Dihybrid phenotypic ratio.

Figure 3 Results showing independent assortment in two pairs of genes. R stands for red flower color, r for white flower color, T for tall gene and t for gene for dwarfness. You would have noticed that the composition of genes or genotype controls the outside expression which we can be see, that is the phenotype. The ratio of progeny in the crosses is therefore, the phenotypic ratio. However, as more and more scientists began to devise genetic experiments, it became clear that Mendel's laws do not hold true in all cases. We shall see about the deviations from Mendel's laws such as incomplete dominance, codominance etc.

Reasons for Mendel's Success

- His choice of material was good. He selected garden pea which has a short life cycle, has self-pollinated bisexual flowers with closed corolla so

pollination can be easily controlled. Also it is easy to cultivate pea plants and a large number of pure line plants with several pairs of contrasting characters were available.

- His selection of traits: The seven pairs of contrasting characters of pea plant considered by Mendel in his experiments were responsible for the desired results that helped Mendel postulate the laws of inheritance.
- Mendel's Technique: His technique of experimentation was excellent and was as follows:
  - Homozygous pure line plants with contrasting characters were crossed.
  - Self-pollination was prevented by removing stamens.
  - Female plants were dusted with pollen grains from another plant with the
  - Opposite feature and were tied in a bag to prevent any further pollination.
  - Seeds were collected and shown in time.
  - He considered the inheritance of one character at a time.
  - The results were analyzed statistically.
  - He performed reciprocal crosses and test crosses to confirm the results.

Important Terms in Genetics

The important terms in genetics are:

- *Factor:* The unit responsible for the inheritance and expression of a particular character. Now replaced by the term gene.
- *Gene:* It is a particular segment of a DNA molecule which determines the inheritance and expression of a particular character.
- *Alleles or Allelomorphs:* Two or more alternative firms of a factor or a gene are called alleles. For example in pea plant, the gene for producing seed shape may occur in two alternative forms: round (R) and wrinkled (r). Genes for round and wrinkled seeds are alleles of each other. Similarly, there are three alleles for gene controlling blood group in man IA, IB, i (I = immunoglobulin gene).
- Alleles occupy same locus on homologous chromosomes.
- *Trait:* is the expressed character, e.g. color of flower, shape of seed etc.
- *Dominant trait:* Out of the two alternating forms (allelomorphs) of a trait, the one which expresses itself in a heterogygous organism in the F1 hybrid is called the dominant trait (dominant allele) and the phenomenon is called dominance. For example in an organism with Tt, T (tallness) expresses itself and t (dwarfness) cannot, so T is the dominant allele.

- *Recessive trait:* Out of the two alternating alleles for a trait, the one which is suppressed (does not express) in the F_1 hybrid is called the recessive trait (recessive allele). Recessive allele expresses only in the homozygous state (e.g. tt).
- *Genotype:* The genetic constitution of an individual (which he/she inherits from the parents) is called the genotype, e.g., the genotype of pure round seeded parent pea plant is RR.
- *Phenotype:* The outward (morphological) appearance of an individual for any trait or traits is called the phenotype, e.g. for seeds, round shape or wrinkled shape is the phenotype.
- *Homozygous:* an individual possessing identical alleles for a trait is called **homozygous** or pure for that trait, e.g. plant with RR alleles is homozygous for the seed shape.
- *Heterozygous:* An individual receiving dissimilar alleles for a trait is called (heterozygous or impure for that trait, e.g. a plant with Rr alleles is heterozygous for the seed shape.
- *Parent generations:* The parents used for the first cross represent the parent (or P_1) generation.
- *F_1 generation:* The progeny produced from a cross between two parents (P_1) is called First filial or F_1 generation.
- *F_2 generation:* The progeny resulting from self-hybridization or inbreeding of F_1 individuals is called Second Filial or F_2 generation.
- *Monohybrid cross:* The cross between two parents differing in a single pair of contrasting characters is called monohybrid cross and the F_1 offspring as the hybrid. The phenotypic ratio of 3 dominants: 1 recessive obtained in the F_2 generation from the monohybrid cross is called monohybrid phenotype ratio (eg. 3: 1 in Mendelian Inheritance).
- *Dihybrid cross:* The cross between two parents in which two pairs of contrasting characters are studied simultaneously for the inheritance pattern is called dihybrid cross. The phenotypic ratio obtained in the F_2 generation from a dihybrid cross is called dihybrid ratio phenotypic (e.g. 9 : 3 : 3 : 1 in Mendelian crosses.

Homozygous is a word that refers to a particular gene that has identical alleles on both homologous chromosomes.

- *Hybridisation:* Crossing organisms belonging to different species for getting favorable qualities in the offspring
- *Test cross:* Crossing of the F_1 progeny with the homozygous recessive parent. If F_1 progeny is heterozygous, then test cross always yields the ratio of 1: 1.
- *Reciprocal cross:* Is the cross in which the sex of the parents is reversed. That is if in the first cross father was dwarf and mother tall, then in the reciprocal cross, dwarf parent will be female and tall parent male.

Deviation from Mendel's Laws Incomplete Dominance

In the four O'clock plant Mirabilis jalapa and Snapdragon or Antirrhinum law of dominance is not followed. Thus when a homozygous red flowered plant (RR) is crossed to a homozygous white flowered plant, all flowers in the F_1 are pink when F_1 plants are self-pollinated, the phenotypic ratio is found to be 1: 2: 1.

Parents	RR × rr
Gametes	R, R × r, r
F_1	Rr Pink
F_2	1 Red : 2 Pink : 1 White

You will find that the heterozygous (Rr) plants have an intermediate colour pink. You must have also noticed that the genotypic ratio 1 RR: 2 Rr: 1 rr and phenotypic ratio 1 Red: 2 Pink: 1 white are both (1: 2: 1) same.

Multiple alleles and codominance

Tall or dwarf, red or white, brown eyes or blue eyes, these are just two varieties or alleles of genes for plant height and flower color and eye color. Most genes, however, have more than two alleles or multiple alleles. An example of multiple alleles is the gene for blood group.

The four blood groups of humans are determined by a single gene. The alleles IA for A group, IB for B blood group are both dominant. Therefore person with alleles I A and IB has the blood group AB as both the genes IA and IB are co-dominant. The gene i when homozygous (ii) gives the blood group O. Genotype and phenotype of blood groups in humans are given in Table 1.

Table 1: Genotypes and Phenotypes of human blood groups

Genotype	Blood group
I^AI^A and I^A i	A
I^B I^B and I^B i	B
I^A I^B	AB
ii	O

Lethal genes

Have you ever seen a yellow mouse? Probably not. The yellow coat colour in mice is due to the presence of the gene which is also responsible for killing the mouse born with that gene. Such genes are termed lethal genes. Some lethal genes kill only in the homozygous condition and are recessive lethals. Dominant lethals cause death even when present in the heterozygous condition.

Pleiotropy

While a gene may have multiple alleles and thus give multiple genotypes, one gene may control several genotypes. For example the recessive gene for white eye in Drosophila when present in the homozygous condition affects several other features such as wing shape and shape of abdomen. Thus a white eyed Drosophila is also born with vestigeal wings and curled **abdomen**.

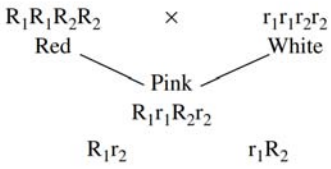
Abdomen constitutes the part of the body between the thorax (chest) and pelvis, in humans and in other vertebrates.

Polygenic or Quantitative Inheritance

When a trait (feature or character) is controlled by a single gene it is termed monogenic inheritance. Many traits or features are controlled by a number of different genes. For example the skin color of humans and the kernel color of wheat results from the combined effect of several genes, none of which are singly dominant. Polygenes affecting a particular trait are found on many chromosomes. Each of these genes has equal contribution and cumulative the total effect. Three to four genes contribute towards formation of the pigment in the skin of humans. So there is a continuous variation in skin color from very fair to very dark. Such inheritance controlled by many genes is termed quantitative inheritance or polygenic (poly meaning due many genes) inheritance.

In polygenic interlace:

- Each dominant gene controls equally the intensity of the character.
- The effect of the dominant genes in cumulative.
- The intensity of character or trait depend upon the number of dominant genes.



	R_1R_2	R_1r_2	r_1R_2	r_1r_2
R_1R_2	$R_1R_1R_2R_2$ Red	$R_1R_1R_2r_2$ Dark Pink	$R_1r_1R_2R_2$ Dark Pink	$R_1r_1R_2r_2$ Pink
R_1r_2	$R_1R_1R_2r_2$ Dark Pink	$R_1R_1r_2r_2$ Pink	$R_1r_1R_2r_2$ Pink	$R_1r_1r_2r_2$ Light Pink
r_1R_2	$R_1r_1R_2R_2$ Dark Pink	$R_1r_1R_2R_2$ Pink	$r_1r_1R_2R_2$ Pink	$r_1r_1R_2r_2$ Light Pink
r_1r_2	$R_1r_1R_2r_2$ Pink	$R_1r_1r_2r_2$ Link Pink	$r_1r_1R_2r_2$ Light Pink	$r_1r_1r_2r_2$ White

1 Red: 4 Dark Pink: 6 Pink: 1 White

Chromosomal Theory of Inheritance

It was observed by Sutton and Boveri in 1902 that

- Maternal (from mother) and paternal (from father) characters come together in the progeny which is diploid or 2n and has chromosomes in pairs and later on segregate during the formation of gametes. The gametes have a single chromosome from each pair and is haploid or n.
- Chromosomes from two parents come together in the same zygote as a result of the fusion of two gametes and again separate out during the formation of gametes. You have already learnt that charomosomes are filamentous bodies present in the nucleus and seen only during cell division.

The above two observations proved that there is a remarkable similarity between the behavior of characters during inheritance and that of chromosomes during meiosis. This led Sutton and Boveri to propose ‘chromosomal theory of inheritance’ and its salient features are as follows.

- The somatic (body) cells of an organism, which are derived by the repeated division of zygote have two identical sets of chromosomes i.e. they are diploid. Out of these, one set of chromosomes is received from the mother (maternal chromosomes) and one set from the father (paternal chromosomes). Two chromosomes of one type (carrying same genes) constitute a homologous pair. Humans have 23 pairs of chromosomes.
- The chromosomes of homologous pair separate out during meiosis at the time of gamete formation.
- The behavior of chromosomes during meiosis indicates that Medelian factors or genes are located linearly on the chromosomes. With progress

in molecular biology it is now known that a chromosome is made of a molecule of DNA and segments of DNA are the genes.



CORE CONCEPT: LINKAGE AND CROSSING OVER

Genetic linkage is the tendency of DNA sequences that are close together on a chromosome to be inherited together during the meiosis phase of sexual reproduction. Two genetic markers that are physically near to each other are unlikely to be separated onto different chromatids during chromosomal crossover, and are therefore said to be more linked than markers that are far apart. In other words, the nearer two genes are on a chromosome, the lower the chance of recombination between them, and the more likely they are to be inherited together. Markers on different chromosomes are perfectly unlinked.

Bateson and Punnett performed a dihybrid cross with true breeding varieties of sweet pea (Lathyns sativus) and instead of 9: 3: 3: 1 ratio in F_2 they got the ratio 7: 1:1: 7. So the parental forms were more than the new combinations of characters (recombinants). It means the two genes chosen for the experiment do not follow the principle of independent assortment. Instead they tend to be inherited together or are linked together. Thus genes present on same chromosome tend to be inherited together and are said to be linked. This phenomenon is called linkage. Crossing over occurs during meiosis I prophase I at the time of gamete formation. The point where crossing over occurs is called chiasma. (plural : chiasmata) See Figure 3. Linked genes get separated from each other by crossing over. Because of linkage and crossing over a heterozygous individual can produce four types of gametes as shown in Figure 4. The figure 4 shows linked genes of the parents and recombinants due to crossing over.

All the genes present on the same pair of chromosomes and with a tendency to be inherited together forms a linkage group.

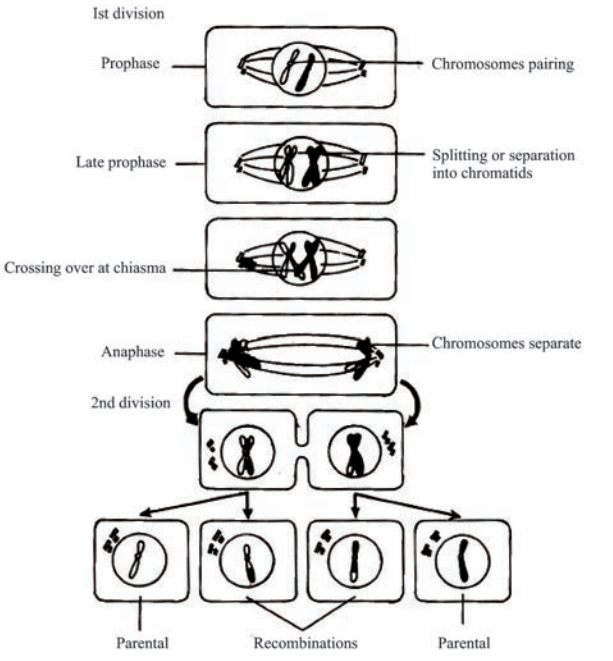


Figure 4: Schematic diagram showing recombination by crossing over.

Chromosomes and Sex Determination

In some diploid organisms, specific chromosomes have a role in **sex determination**. Such chromosomes some are called sex chromosomes and the rest of the chromosomes of a set are called autosomes.

Sex determination is the establishment of the sex of an organism, usually by the inheritance at the time of fertilization of certain genes commonly localized on a particular chromosome.

- If sex chromosomes are morphologically similar (i.e. XX) in an individual, the individual is termed homogametic. Such individuals, produce only one kind of gamete (containing X). For example: all eggs of the human female contain an X chromosome. So human female is termed homogametic.

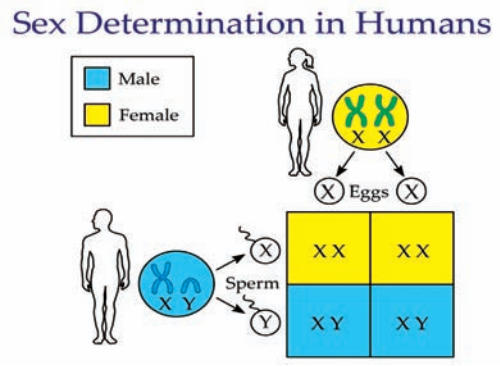


Figure 5: Chromosomal basis of sex determination in humans.

- If sex chromosomes are morphologically dissimilar (i.e. XY) in an individual, is termed heterogametic. Such individuals produce two types of gametes. (One containing X and the other containing Y). For example: human male produces two kinds of sperms, X bearing and Y bearing sperms. When the human egg is fertilized by an X bearing sperm a boy is born (Figure 6). Whether the unborn will be a male or female is purely a matter of chance and no parent can be blamed for the sex of the progeny

Cris-Cross Inheritance (X-Linked Inheritance)

We already know that genes are located on chromosomes. The genes which are located on X chromosome (sex chromosome), are called sex linked genes. These genes show cris-cross inheritance as shown in Figure 6. When a male has a defective sex linked gene located on X chromosome during reproduction, he transmits the defective X chromosome to his daughter only. The female who has this gene transmits it to her sons and daughter both in equal probability. So the male passes on his recessive sex linked trait to 50% of his grandsons through his daughter. The sex linked trait being recessive is not expressed in female but is expressed in males. Therefore males suffer from genetic defect while females are only carriers of these defective genes. Females express the trait if they are homozygous for the recessive gene.

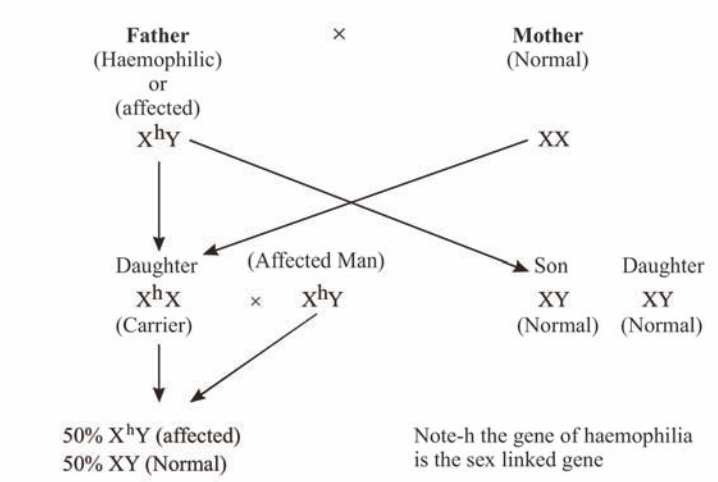


Figure 6: Cris cross inheritance or X-linked inheritance.

Cris Cross Inheritance in humans

Red green color blindness and Haemophilia are examples of sex linked inheritance in humans. The defective gene is located on X chromosome. Thus a single defective gene causes disease in male while two defective genes (homozygous condition) only can cause the disease in female. Females in heterozygous condition are apparently normal but actually the carriers of the disease. Carrier females pass this defective gene to 50% of her sons. The disease is expressed only in males because male does not have the partners of the genes on Y chromosome. See Figure 7 (a), (b) and (c).

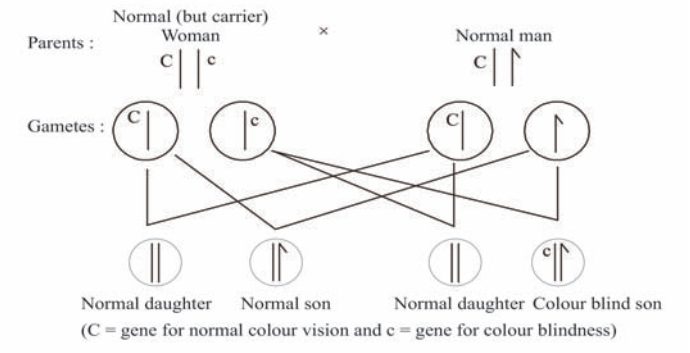


Figure 7: (a) Progeny of carrier female for color blindness and a man with normal color vision.

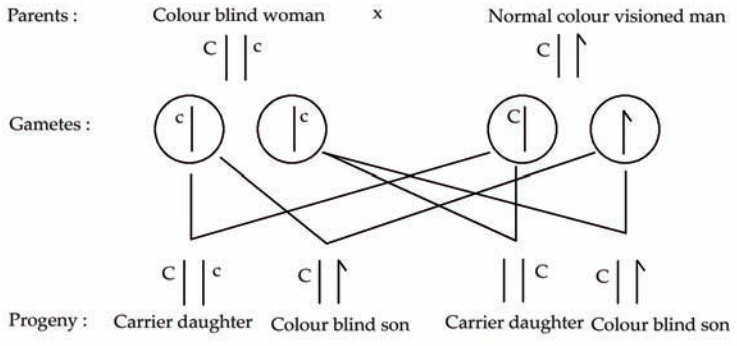


Figure 7: (b) Progeny of color blind female and a man with normal vision.

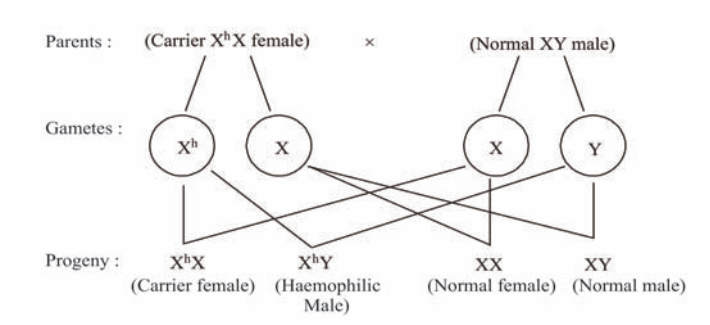


Figure 7: (c) Progeny of an haemophilic carrier female and normal male.

Mitochondrial Inheritance as a Case of Maternal Inheritance

Apart from the nucleus, mitochondria and chloroplasts also possess DNA and you have just learnt that genes are segments of DNA. Till now you have studied about genes on the chromosomes present in the nucleus. Since mitochondria come into the **zygote** from the egg, inheritance of mitochondrial DNA or DNA is said to be a case of maternal inheritance. In fact, certain diseases and therefore the genes responsible for them are due to defects in mitochondrial DNA and can be traced to the mother's family.

Zygote is a eukaryotic cell formed by a fertilization event between two gametes.

Human Karyotype

Human karyotype is the arrangement of chromosomes in seven groups according to the types of chromosomes and their size. It is prepared by arranging chromosomes seen at mitotic metaphase.

- Total no. of chromosomes or 2n = 46 (23 pairs).
- Number of autosomes = 44 (22 pairs).
- Sex chromosomes = X and Y = 2
- Depending on size, location of centromere, and bands obtained by special staining methods, human chromosomes are grouped into 7 groups A to G as shown in Figure 8.

Sex determination in humans, as you have already learnt is as follows:

- Normal male 22 pairs of autosomes + one X chromosome and one Y chromosome
- Normal female 22 pairs of autosomes + two X chromosome

Presence of Y is necessary for maleness.

Absence of Y chromosomes makes the individual a female.

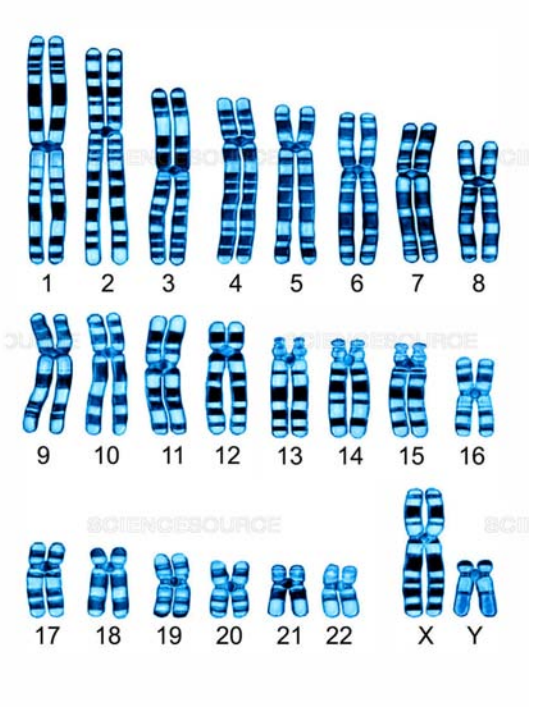
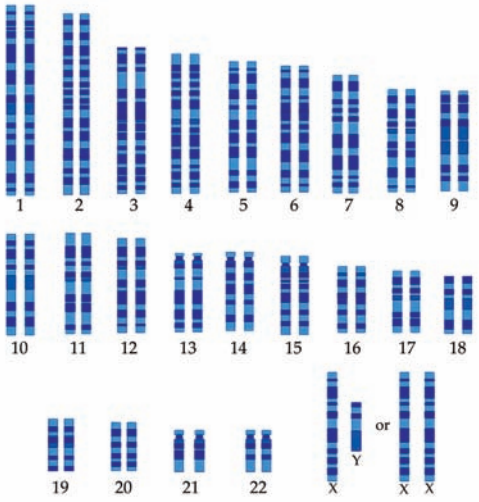


Figure 8: Karyotype showing chromosomes of normal male. Female has the same autosomes but two X-chromosomes instead of XY.

How a karyotype is made:

- Cells are blocked during mitosis.
- The chromosomes are taken from those cells, attached to a slide and are stained with Giesma dye (pronounced Jeem-suh). The dye stains regions of chromosomes that are rich in the base pairs Adenine (A) and Thymine (T) producing a dark band.
- Pictures of the chromosomes are taken, cut out and matched up according to size, banding pattern and centromere position as guides.



Interpreting a karyotype:

Scientists use a specific notation to characterize the karyotype.

This notation includes:

- the total number of chromosomes
- the sex chromosomes
- any extra or missing autosomal chromosomes

For example,

47, XY, +18 indicates that the patient has 47 chromosomes, is a male, and has an extra autosomal chromosome 18.

46, XX is a female with a normal number of chromosomes

47, XXY is a patient with an extra sex chromosome.

Making a diagnosis with a karyotype:

- Scientists can diagnose or predict genetic disorders by looking at the chromosomes in a karyotype.

DNA is made up of long chains of nucleotides. To make a complete DNA molecule, single nucleotides join to make chains that come together as matched pairs and form long double strands.

- Karyotype analysis is used in prenatal testing and in diagnosing certain disorders, such as Down syndrome.

Some chromosome aberrations in humans:

Chromosome abnormality	Syndrome name (condition)	Effects on individual (variable)
X	Turner's	• sterile female • underdeveloped female characteristics • short stature • sometimes below normal intelligence
XXY	Klinefelter's	• sterile male • abnormal male characteristics • abnormally long arms and legs • below normal intelligence
Trisomy #21	Down's	• low mental ability • short arms and legs • internal defects • distinct facial features
Trisomy #18	Edwards	• severe abnormalities • life expectancy about 10 weeks
Trisomy #13	Patans	• severe defects including small, non-functioning eyes • survive only a few weeks after birth
Defective #5 (part of chromosome missing)	Cri du chat	• severe defects, both physical and mental • sound, when crying, similar to a cat's meow

Chromosomal Abnormalities and Genetic Disorders in Humans

Chromosomes are the structures that hold genes. Genes are the individual instructions that tell our bodies how to develop and function; they govern physical and medical characteristics, such as hair color, blood type and susceptibility to disease.



There are many types of chromosome abnormalities. However, they can be organized into two basic groups: numerical abnormalities and structural abnormalities.

- Numerical Abnormalities: When an individual is missing one of the chromosomes from a pair, the condition is called monosomy. When an individual has more than two chromosomes instead of a pair, the condition is called trisomy.

An example of a condition caused by numerical abnormalities is Down syndrome, which is marked by mental retardation, learning difficulties, a characteristic facial appearance and poor muscle tone (hypotonia) in infancy. An individual with Down syndrome has three copies of chromosome 21 rather than two; for that reason, the condition is also known as Trisomy 21. An example of monosomy, in which an individual lacks a chromosome, is Turner syndrome. In Turner syndrome, a female is born with only one sex chromosome, an X, and is usually shorter than average and unable to have children, among other difficulties.

- Structural Abnormalities: A chromosome's structure can be altered in several ways.
  - Deletions: A portion of the chromosome is missing or deleted.
  - Duplications: A portion of the chromosome is duplicated, resulting in extra genetic material.
  - Translocations: A portion of one chromosome is transferred to another chromosome. There are two main types of translocation. In a reciprocal translocation, segments from two different chromosomes have been exchanged. In a Robertsonian translocation, an entire chromosome has attached to another at the centromere.
  - Inversions: A portion of the chromosome has broken off, turned upside down, and reattached. As a result, the genetic material is inverted.
  - Rings: A portion of a chromosome has broken off and formed a circle or ring. This can happen with or without loss of genetic material.

Most chromosome abnormalities occur as an accident in the egg or sperm. In these cases, the abnormality is present in every cell of the body. Some abnormalities, however, happen after conception; then some cells have the abnormality and some do not. Chromosome abnormalities can be inherited from a parent (such as a translocation) or be "*de novo*" (new to the individual). This is why, when a child is found to have

Chromosome is a DNA molecule with part or all of the genetic material of an organism.

an abnormality, **chromosome** studies are often performed on the parents. Any change from the normal number or structure of chromosomes causes abnormalities. Following are some of them.

Mongolism or Down's syndrome

The individual has 47 chromosome because of one extra chromosome in the 21st pair (Trisomy of chromosome 21). The outcome is that they are

- mentally retarded
- have a thick tongue
- and a drooping (false expression of pleasure) face. Figure 9.

The possibility of giving birth to a Mongol child is far greater in pregnant mothers above the age of forty.

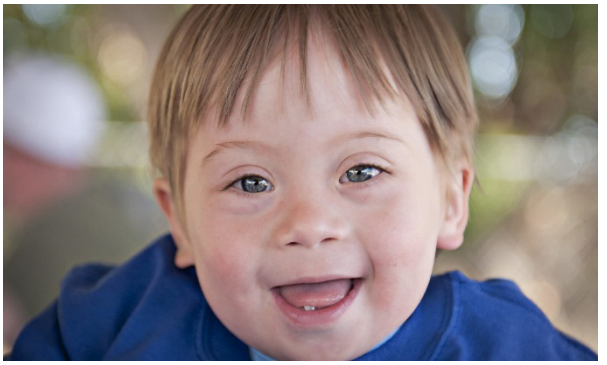


Figure 9: Mongolism or Down's syndrome.

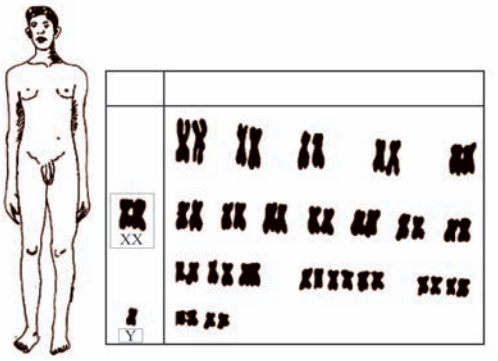


Figure 10: Man showing Klinefelter's syndrome.

Klinefelter's syndrome

Individual is a male with 47 chromosome with one extra X chromosome. (44 autosomes + XXY). Typical features of a

Klinefelter are

- Tall, mentally retarded male;
- Sterile and shows breast development or gynaecomastia (gynae : female; massere : mammary glands). Figure 10.

Turner's syndrome

Individual is a female with 45 chromosomes with only one X, chromosome (22 pairs of autosomes +XO). The characteristic features are

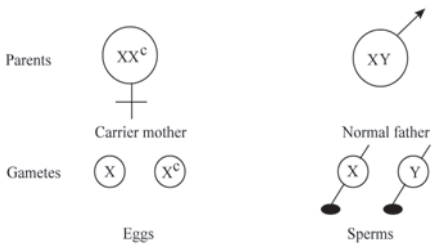
- Mentally retarded
- web like skin on neck.
- incompletely developed breasts. Fig 11



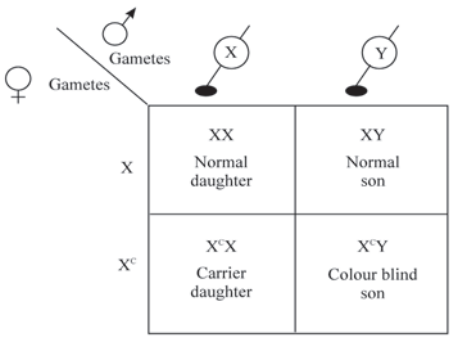
Figure 11: A woman suffering from Turner's syndrome.

Color blindness and Hemophilia (Bleeder's diseases)

Both these are sex linked disorders. (Figures 6, 7)



The inheritance is as follows:



Note: X = normal allele; X^C = recessive mutant

Figure 12: Inheritance of color blindness

In male, the single X-chromosome is received from the mother. Hence a defective, gene (for color blindness or haemophilia) on X chromosome of the mother, is passed on to the son and expressed as a defect. The daughter receives one X-chromosome from the mother and the other X from the father. The defective gene received from the mother may be masked by normal allele on the other X. The daughter is then called a carrier. (Figure 12). Color blind males are unable to distinguish between red-green colors. In haemophilia afflicted male, blood does not clot easily and the patient may bleed to death. Its mode of inheritance is exactly like that of color blindness

Thalassemia

It is a disorder in which haemoglobin is not synthesized properly. So, frequent blood transfusions are required for survival. The defective gene is recessive and therefore heterozygous parents may not show the disorder. The child who gets the defective genes 'from both the parents (homozygous recessive) suffers from Thalassemia.

Sickle Cell Anemia

This is another hereditary abnormality due to mutation of a single gene in which. Red blood corpuscles lose their shape and become sickle shaped because of defective Haemoglobin. Individuals possessing two defective genes (homozygous recessive), cannot survive. In the heterozygous individuals, one gene is normal and so half the number of total red blood corpuscles are normal containing normal haemoglobin while the others are defective. A boon in disguise against malaria for children with one defective Haemoglobin gene can survive as they are less affected by malaria because the malaria parasite cannot thrive inside the defective RBC's.

Rh factor

Rh factor is an antigen (a protein) present on the surface of red blood corpuscles. About 15% of all women do not have the gene for Rh antigen. They are Rh-negative. Men can also be Rh-negative, but the problem which this trait creates is in Rhnegative women. A pregnant Rh-negative women whose husband: is Rh + may bear a child who may have inherited the Rh + gene from the father. If the foetal blood enters mother's body stream, her immune system produces antibodies against Rh antigen of the embryo. Rh antigens cause no problems in first pregnancy. But in a subsequent early pregnancy, the mother's antibodies against Rh antigen may enter the foetal blood stream and destroy its red blood corpuscles causing severe anemia which may even be fatal. (erythroblastosis foetalis) Now a days Rh-negative mother of a Rh-positive foetus is treated immediately after delivery, to destroy Rh antigens in her blood stream. (Figure 13)

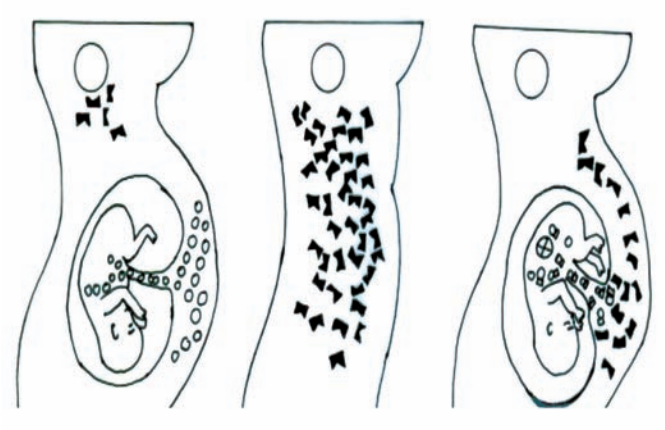


Figure 13: The mechanism of Rh inheritance.

- Shows the first pregnancy where the mother is Rh (-) and foetus Rh (+). Protein (empty circles) of the foetus stimulates the production of anti-factors (black blocks) in the mother.
- Shows the retention of anti-factors in the mother's body.
- Shows the Rh (+) foetus in the womb of the same mother during the second pregnancy. The anti-factors from the same mother during the second pregnancy.

The anti-factors from the mother's body will destroy the infant's red blood cells.

Amniocentesis

Amniocentesis in a technique by which hereditary disorders due to defects in genes can be detected. In this technique (Figure 14)

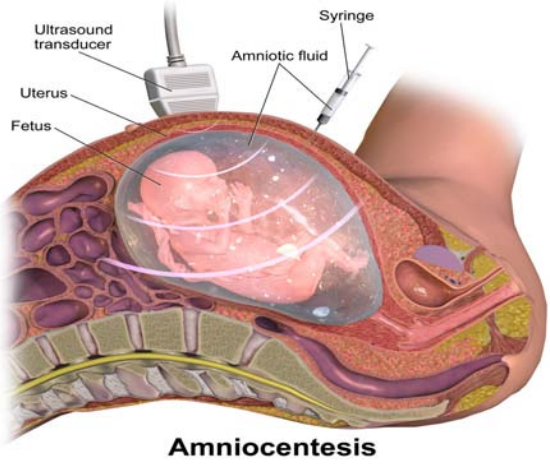


Figure14: Techniques of Aminocentesis.

- a small sample of amniotic fluid which surrounds the foetus is syringed out.
- This fluid has cells which break off from the skin of the foetus.

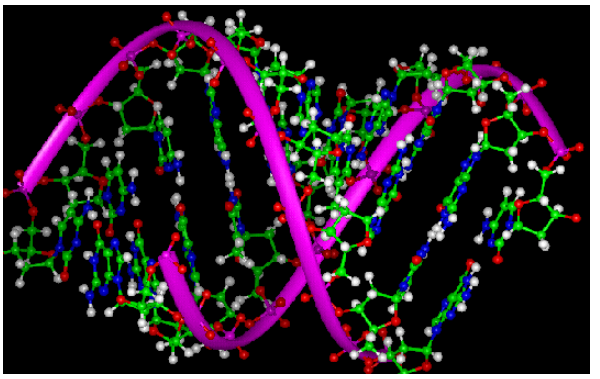
- Foetal cells are picked up and cultured.
- Chromosomes in the dividing cells are analyzed for genetic defects.

If incurable genetic defects are detected, pregnancy can be terminated. It is illegal to use amniocentesis for detecting the sex of the unborn.

The Human Genome

Human genome, all of the approximately three billion base pairs of deoxyribonucleic acid (DNA) that make up the entire set of chromosomes of the human organism. The human genome includes the coding regions of DNA, which encode all the genes (between 20,000 and 25,000) of the human organism, as well as the noncoding regions of DNA, which do not encode any genes. By 2003 the DNA sequence of the entire human genome was known.

In the last over hundred years genetics and molecular biology have gene far ahead and the progress has been very rapid. You have just read about genetic disorders and today there is hope for their cure through gene therapy. This is because in 2003, most of the genes on human chromosomes in have been mapped or located on the 23(n) chromosomes. The genes responsible for inheritance of various structural features, genes that control various enzymes that catalyse the various biochemical reactions in the body, genes responsible for genetic disorders have been located. Genome means genes of a particular organism. Since genes are present in pairs (one inherited from mother and other from the father), all kinds of genes present in a particular type of organism are present on its haploid set (n). Thus human genome may be defined as all the genes present in the haploid set of chromosomes in humans. There are an estimated 20,000 to 25,000 genes and 3 billion base pairs in the total human DNA. Each human chromosome has apart from protein coding genes regulating base sequences, non-coding DNA, promoter sequences (TATA boy) in between genes that code for proteins. There are genes that code for the production of ribosomal RNA and the many tRNAs. It is estimated that only 1.5% of the human genome has protein coding sequences and 20,000 to 25,000 genes.



The human genome, like the genomes of all other living animals, is a collection of long polymers of DNA. These polymers are maintained in duplicate copy in the form of chromosomes in every human cell and encode in their sequence of constituent bases

(guanine [G], adenine [A], thymine [T], and cytosine [C]) the details of the molecular and physical characteristics that form the corresponding organism. The sequence of these polymers, their organization and structure, and the chemical modifications they contain not only provide the machinery needed to express the information held within the genome but also provide the genome with the capability to replicate, repair, package, and otherwise maintain itself. In addition, the genome is essential for the survival of the human organism; without it no cell or tissue could live beyond a short period of time. For example, red blood cells (erythrocytes), which live for only about 120 days, and skin cells, which on average live for only about 17 days, must be renewed to maintain the viability of the human body, and it is within the **genome** that the fundamental information for the renewal of these cells, and many other types of cells, is found.

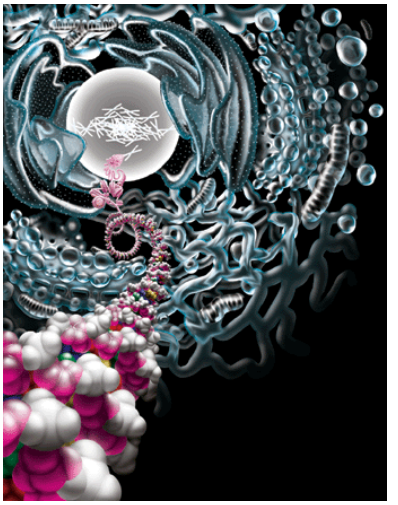
The content of the human genome is commonly divided into coding and noncoding DNA sequences. Coding DNA is defined as those sequences that can be transcribed into mRNA and translated into proteins during the human life cycle; these sequences occupy only a small fraction of the genome (<2%). Noncoding DNA is made up of all of those sequences (ca. 98% of the genome) that are not used to encode proteins.

Genome is the genetic material of an organism. It consists of DNA. The genome includes both the genes and the noncoding DNA, as well as the genetic material of the mitochondria and chloroplasts.

The human genome is not uniform. Excepting identical (monozygous) twins, no two humans on Earth share exactly the same genomic sequence. Further, the human genome is not static. Subtle and sometimes not so subtle changes arise with startling frequency.

Some of these changes are neutral or even advantageous; these are passed from parent to child and eventually become commonplace in the population. Other changes may be detrimental, resulting in reduced survival or decreased fertility of those individuals who harbour them; these changes tend to be rare in the population.

The genome of modern humans, therefore, is a record of the trials and successes of the generations that have come before. Reflected in the variation of the modern genome is the range of diversity that underlies what are typical traits of the human species. There is also evidence in the human genome of the continuing burden of detrimental variations that sometimes lead to disease.



Knowledge of the human genome provides an understanding of the origin of the human species, the relationships between subpopulations of humans, and the health tendencies or disease risks of individual humans. Indeed, in the past 20 years knowledge of the sequence and structure of the human genome has revolutionized many fields of study, including medicine, anthropology, and forensics. With technological advances that enable inexpensive and expanded access to genomic information, the amount of and the potential applications for the information that is extracted from the human genome is extraordinary.

Role of the Human Genome in Research

Since the 1980s there has been an explosion in genetic and genomic research. The combination of the discovery of the polymerase chain reaction, improvements in DNA sequencing technologies, advances in bioinformatics (mathematical biological analysis), and increased availability of faster, cheaper computing power has given scientists the ability to discern and interpret vast amounts of genetic information from tiny samples of biological material. Further, methodologies such as fluorescence in situ hybridization (FISH) and comparative genomic hybridization (CGH) have enabled the detection of the organization and copy number of specific sequences in a given genome.



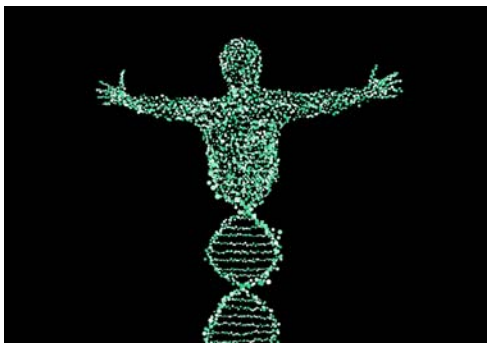
Understanding the origin of the human genome is of particular interest to many researchers since the genome is indicative of the evolution of humans. The public availability of full or almost full genomic sequence databases for humans and a multitude of other species has allowed researchers to compare and contrast genomic information between individuals, populations, and species. From the similarities and differences observed, it is possible to track the origins of the human genome and to see evidence of how the human species has expanded and migrated to occupy the planet.

Origins of the Human Genome

Comparisons of specific DNA sequences between humans and their closest living relative, the chimpanzee, reveal 99 percent identity, although the homology drops to 96 percent if insertions and deletions in the organization of those sequences are taken into account. This degree of sequence variation between humans and chimpanzees is only about 10-fold greater than that seen between two unrelated humans. From comparisons of the human genome with the genomes of other species, it is clear that the genome of modern humans shares common ancestry with the genomes of all other animals on the planet and that the modern human genome arose between 150,000 and 300,000 years ago.

Ongoing collaboration between archaeologists, anthropologists, and molecular geneticists at the Max Planck Institute in Germany and the Lawrence Berkeley National **Laboratory** and the Joint Genome Institute in the United States has enabled sequence comparisons between modern humans (*Homo sapiens*) and Neanderthals (*H. neanderthalensis*). The data obtained so far demonstrate that modern humans and Neanderthals share about 99.5 percent genome sequence identity; some scientists have claimed that sequence identity may actually be as high as 99.9 percent.

Laboratory is a facility that provides controlled conditions in which scientific or technological research, experiments, and measurement may be performed.



Research suggests that populations of *H. sapiens* split from *H. neanderthalensis* ancestral populations perhaps as recently as 370,000 years ago and likely shared a common ancestor some 500,000–700,000 years ago. Genomic studies have indicated that there was almost no interbreeding between *H. sapiens* and *H. neanderthalensis*. This suggests that when Neanderthals, the last of the *Homo* relatives of modern humans, became extinct about 30,000 years ago, only modern humans were left to populate Earth. However, other research has revealed that modern *H. sapiens* in Eurasia, specifically peoples in Europe, China, and Papua New Guinea, have genomes that are more similar to the Neanderthal genome than they are to the genomes of modern *H. sapiens* in Africa. Scientists estimate that 1 to 4 percent of DNA of modern Eurasians is shared with Neanderthals, a level of similarity that is not found between Neanderthals and modern Africans. These findings indicate that limited interbreeding and gene flow took place between Neanderthals and ancestral *H. sapiens* populations after the latter migrated out of Africa but before they dispersed to other parts of the world.

Comparing the DNA sequences of groups of modern humans from different continents also allows scientists to define the relationships and even the ages of these different populations. By combining these genetic data with archeological and linguistic information, anthropologists have been able to discern the origins of *Homo sapiens*in Africa and to track the timing and location of the waves of human migration out of Africa that led to the eventual spread of humans to other continents of the globe. For example, genetic evidence indicates that the first humans migrated out of Africa approximately 60,000 years ago, settling in southern Europe, the Middle East, southern Asia, and Australia. From there, subsequent and sequential migrations brought humans to northern Eurasia and across what was then a land bridge to North America and finally to South America.

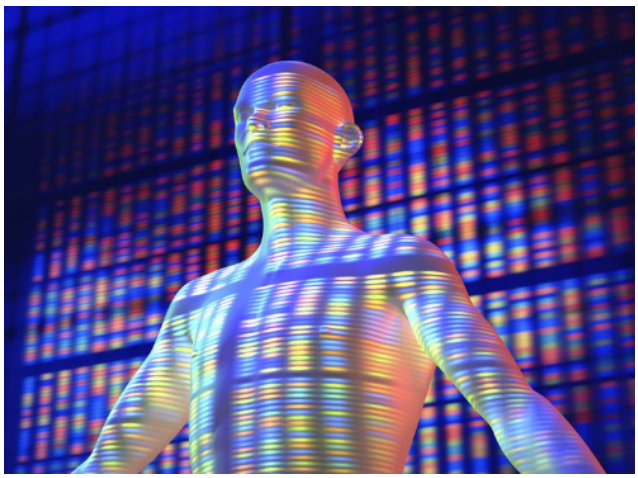


As humans migrated across the continents, sequence variations arose that became differentially fixed in different populations. Some variations likely reflect what are called founder effects, changes in gene frequency that occur in small populations. Founder effects are generally characterized by genes that are expressed with increasing frequency from one generation to the next and can be traced back to the original founders of the population. Other variations reflect differential selective pressures at work. For example, populations living in equatorial climates were under strong

selective pressure that favored dark skin color to protect against extreme sun exposure, thereby decreasing the deleterious health effects caused by sunburn and skin cancer. In contrast, populations migrating to more polar latitudes, where levels of sun exposure are relatively low, experienced strong selective pressure that favored light skin color, thereby facilitating the absorption of sunlight by the skin for the synthesis of vitamin D. In northern Europe and Scandinavia, therefore, individuals with genetic variations leading to lighter skin color were less likely to become vitamin D deficient and suffer from the bone disease known as rickets.

Social Impacts of Human Genome Research

Databases have been compiled that list and summarize specific DNA variations that are common in certain human populations but not in others. Because the underlying DNA sequences are passed from parent to child in a stable manner, these genetic variations provide a tool for distinguishing the members of one population from those of the other. Public genetic ancestry projects, in which small samples of DNA can be submitted and analyzed, have allowed individuals to trace the continental or even subcontinental origins of their most ancient ancestors.



Surveillance is the monitoring of behavior, activities, or other changing information for the purpose of influencing, managing, directing, or protecting people.

The role of genetics in defining traits and health risks for individuals has been recognized for generations. Long before DNA or genomes were understood, it was clear that many traits tended to run in families and that family history was one of the strongest predictors of health or disease. Knowledge of the human genome has advanced that realization, enabling studies that have identified the genes and even specific sequence variations that contribute to a multitude of traits and disease

risks. With this information in hand, health care professionals are able to practice predictive medicine, which translates in the best of scenarios to preventative medicine. Indeed, presymptomatic genetic diagnoses have enabled countless people to live longer and healthier lives. For example, mutations responsible for familial cancers of the breast and colon have been identified, enabling presymptomatic testing of individuals in at-risk families. Individuals who carry the mutant gene or genes are counseled to seek heightened **surveillance**. In this way, if and when cancer appears, these individuals can be diagnosed early, when the cancers are most effectively treated.

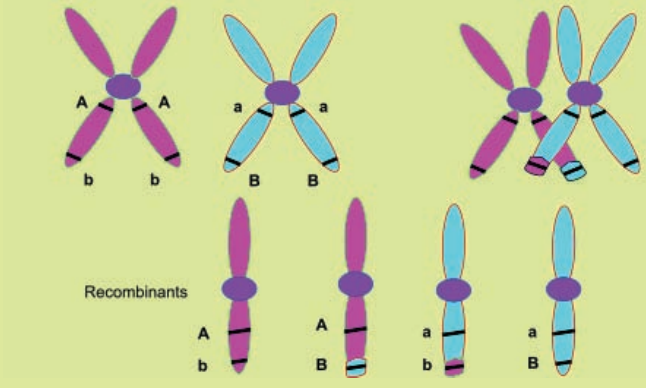
CONCEPT OF CHROMOSOMAL THEORY OF INHERITANCE

The chromosomal theory of inheritance is a fundamental concept in genetics that was proposed in the early 20th century by American geneticist Thomas Hunt Morgan and his colleagues. It revolutionized our understanding of how traits are inherited and laid the foundation for modern genetics. The theory postulates that genes, the units of heredity, are located on specific structures within cells called chromosomes. Chromosomes are thread-like structures made of DNA and proteins, found in the cell nucleus of eukaryotic organisms.

According to the chromosomal theory of inheritance, genes are arranged linearly on chromosomes, and their specific positions on the chromosome determine the inheritance patterns of traits. During cell division, chromosomes duplicate and are passed from parent cells to daughter cells. This process, known as meiosis, ensures that each gamete (sperm or egg) receives one set of chromosomes, resulting in genetic diversity in the offspring.

One of the groundbreaking experiments that supported the chromosomal theory was conducted by Morgan and his team using fruit flies (Drosophila melanogaster) as model organisms. They observed that certain traits, such as eye color, were linked to specific sex chromosomes. The inheritance of these traits could be explained by the genes residing on the sex chromosomes, providing strong evidence for the chromosomal basis of inheritance.

The chromosomal theory of inheritance also clarified the mechanism of sex determination. In many organisms, including humans, sex is determined by the presence of sex chromosomes. For instance, in humans, females have two X chromosomes (XX), while males have one X and one Y chromosome (XY). The presence or absence of the Y chromosome determines whether an individual develops as male or female.



Since the formulation of the chromosomal theory, significant advancements in genetics have occurred. The identification of genes and their specific locations on chromosomes has become feasible through molecular techniques like genetic mapping and DNA sequencing. Modern genetic research continues to uncover the genetic basis

of various traits, diseases, and biological processes, deepening our understanding of the chromosomal theory and its implications. The chromosomal theory of inheritance is a cornerstone concept in genetics, proposing that genes are located on chromosomes and their inheritance is dependent on the transmission of chromosomes during cell division. This theory provided a unifying framework to explain the inheritance patterns of traits, sex determination, and genetic diversity. Its legacy continues to shape our knowledge of genetics and remains integral to ongoing research in the field.

SUMMARY

- Genetics is a captivating and crucial branch of biology that explores the inheritance and variation of traits in living organisms. It delves into the mechanisms by which genetic information is passed from one generation to the next, shaping the characteristics and diversity of all living beings.
- Central to genetics is the concept of genes, the hereditary units that encode specific traits and characteristics within an organism's DNA. Genes are composed of DNA sequences and serve as the instructions for building proteins, which play vital roles in the structure and functioning of cells and organisms.
- Mendel's studies have provided scientists with the basis for mathematically predicting the probabilities of genotypes and phenotypes in the offspring of a genetic cross.
- Chromosomes are the structures that hold genes. Genes are the individual instructions that tell our bodies how to develop and function; they govern physical and medical characteristics, such as hair color, blood type and susceptibility to disease.
- Human genome, all of the approximately three billion base pairs of deoxyribonucleic acid (DNA) that make up the entire set of chromosomes of the human organism.

MULTIPLE CHOICE QUESTIONS

1. The tendency of an offspring to resemble its parent is known as

a. Variation
b. Resemblance
c. Heredity
d. Inheritance

2. Who is known as the "Father of Genetics"?

a. Mendel
b. Morgan
c. Bateson
d. Watson

3. The alternate form of a gene is

a. Recessive character
b. Alternate type
c. Allele
d. Dominant character

4. The genotypic ratio of a monohybrid cross is

a. 3:1
b. 1:2:1
c. 9:3:3:1
d. 2:1:1

5. The crossing of F1 to either of the parents is known as

a. Test cross
b. F1 cross
c. Back cross
d. All of the above

Answers

1. (c) 2. (a) 3. (c) 4. (b) 5. (c)

REVIEW QUESTIONS

1. State the three Mendel's laws of inheritance. Which one of these laws in universal?
2. What will be the blood group of the progeny of parents with AB and O groups?
3. State the chromosomal theory of inheritance.
4. Why is X-linked inheritance termed cris-cross inheritance?
5. Give an account of genetic disorders caused by abnormal chromosomal number.
6. What is amniocentesis? How and for what is it carried out?
7. What do you understand by human genome?
8. Write notes on :
    a. Pleiotropy
    b. Human karyotype
    c. Mitochondrial inheritance

REFERENCES

1. Aguillon, S. M., J. W. Fitzpatrick, R. Bowman, S. J. Schoech, A. G. Clark, G. Coop, and N. Chen, 2017, 08)Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. PLOS Genetics 13(8): 1–27.
2. Allen, G. E., 2011 Eugenics and modern biology: Critiques of eugenics, 1910–1945. Annals of human genetics 75(3): 314–325.
3. Alvarez, G., F. C. Ceballos, and C. Quinteiro, 2009 The role of inbreeding in the extinction of a European royal dynasty. PLoS One 4(4): e5174.
4. Andolfatto, P. and M. Przeworski, 2001 Regions of lower crossing over harbor more rare variants in African populations of Drosophila melanogaster. Genetics 158: 657–665.
5. Andolfatto, P., 2007 Hitchhiking effects of recurrent beneficial amino acid substitutions in the Drosophila melanogaster genome. Genome Res. 17: 1755–1762.
6. Cystic fibrosis Genetics Home Reference, NIH, Accessed 16 May 2008
7. Genetic engineering: Bacterial arsenal to combat chewing insects Archived 15 May 2011 at the Wayback Machine GMO Safety, Jul 2010.
8. Peto J (June 2002). “Breast cancer susceptibility – A new look at an old model”. Cancer Cell. 1 (5): 411–2.
9. Staff (18 November 2005). “Gene Therapy” (FAQ). Human Genome Project Information. Oak Ridge National Laboratory. Retrieved 28 May 2006.



CHAPTER 2

PRINCIPLES OF GENETIC TRANSMISSION



Genetic transmission is the transfer of genetic information from genes to another generation (from parent to offspring), almost synonymous with heredity, or from one location in a cell to another. It should not be confused with chromosomal translocation, which is rearrangement of parts between non-homologous chromosomes.

After studying this chapter, you will understand the core concepts of:

- Mendel's Experiments
- Segregation
- Dominance
- Nomenclature
- Multiple Alleles
- Biochemical Genetics

INTRODUCTION

Genetic transmission, also known as heredity, is a fundamental biological process that governs the passing of genetic information from one generation to the next. It plays a central role in shaping the characteristics and traits of living organisms, including humans. The concept of genetic transmission has intrigued scientists and scholars for centuries, leading to remarkable discoveries and advancements in the field of genetics.

The foundation of genetic transmission lies in the hereditary units called genes, which are segments of DNA that encode specific traits and characteristics within an organism. Genes act as the instructions for building proteins, molecules essential for the structure and functioning of cells and the entire organism. The arrangement and sequence of genes along the DNA molecule form the unique genetic code of each individual, determining traits such as eye color, height, and susceptibility to certain diseases.

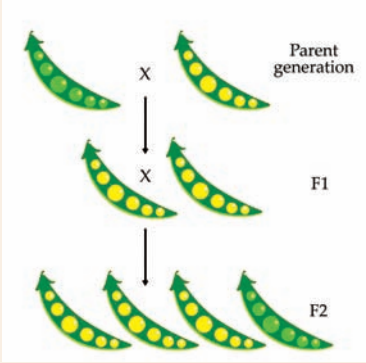
The process of genetic transmission occurs during sexual reproduction, where each parent contributes half of their genetic material to form the offspring's complete genetic makeup. The fusion of the mother's egg and the father's sperm results in a new individual with a combination of genetic traits from both parents. This blending of genetic material creates offspring with genetic diversity, leading to variation among individuals within a population.

During sexual reproduction, meiosis is a crucial step in the genetic transmission process. Meiosis is a specialized cell division process that reduces the chromosome number in the sex cells (gametes) to half. This reduction ensures that the offspring receive the appropriate number of chromosomes, one set from each parent, and that the genetic diversity is maintained.

Genetic transmission is not limited to humans but applies to all living organisms. In plants, animals, and other organisms, genes are transmitted from one generation to the next, shaping their characteristics and traits. This transmission is vital for the continuity and evolution of species, as genetic variation allows organisms to adapt to changing environments and survive over time.

The study of genetic transmission has been instrumental in the field of genetics and has led to groundbreaking discoveries. One of the most famous experiments that laid the foundation for understanding genetic transmission was conducted by Gregor Mendel, an Austrian monk, in the 19th century. Mendel's work with pea plants revealed the principles of inheritance, including the concepts of dominant and recessive traits and the segregation of alleles during genetic transmission. Genetic transmission also has profound implications for human health and disease. Inherited genetic mutations can result in genetic disorders and conditions, while understanding genetic transmission allows for genetic counseling and screening to assess the risk of passing on certain genetic traits.





CORE CONCEPT: MENDEL'S EXPERIMENTS

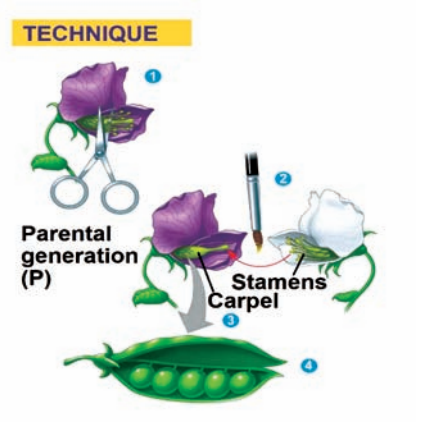
Gregor Mendel developed the model of heredity that now bears his name by experiments on various charactersitics of pea plants: height (tall vs. Short); seed color (yellow vs. Green); seat coat (smooth vs. wrinkled), etc.

Gregor Mendel, through his work on pea plants, discovered the fundamental laws of inheritance. He deduced that genes come in pairs and are inherited as distinct units, one from each parent. Mendel tracked the segregation of parental genes and their appearance in the offspring as dominant or recessive traits. He recognized the mathematical patterns of inheritance from one generation to the next.

Gregor Mendel, in his experiments, tried to **crossbreed** plants that had discrete, non-overlapping characteristics and then to observe the distribution of these characteristics over the next several generations. Mendel worked with the common garden pea plant, *Pisum sativum*.

He chose the pea plant for at least three reasons:

- The garden pea was easy to cultivate and had a relatively short life cycle.
- The plant had discontinuous characteristics such as flower color and pea texture.
- In part because of its anatomy, pollination of the plant was easy to control.



Crossbreed is an organism with purebred parents of two different breeds, varieties, or populations.

Foreign pollen could be kept out, and cross-fertilization could be accomplished artificially. Figure 1 shows a cross section of the pea flower that indicates the keel, in which the male and female parts develop. Normally, self-fertilization occurs when pollen falls onto the stigma before the bud opens. Mendel cross-fertilized the plants by opening the keel of a flower before the anthers matured and placing pollen from another plant on the stigma. In the more than ten thousand plants Mendel examined, only a few were fertilized other than the way he had intended (either self- or cross-pollinated).

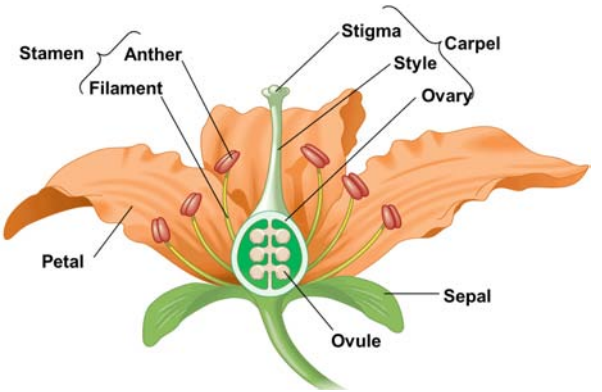


Figure 1: Anatomy of the garden pea plant flower.

Mendel used plants obtained from suppliers and grew them for two years to ascertain that they were homogeneous, or true-breeding, for the particular characteristic under study. He chose for study the seven characteristics shown in figure 2. Take as an example the characteristic of plant height. Although height is often continuously distributed, Mendel used plants that displayed only two alternatives: tall or dwarf. He made the crosses shown in figure 3.

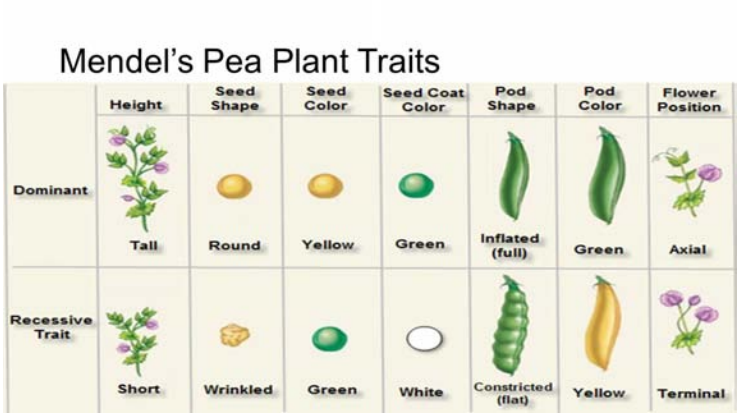


Figure 2: Seven characteristics that Mendel observed in peas.

In the parental, or P_1, generation, dwarf plants pollinated tall plants, and, in a reciprocal cross, tall plants pollinated dwarf plants, to determine whether the results were independent of the parents' sex. As we will see later on, some traits follow inheritance patterns related to the sex of the parent carrying the traits. In those cases, reciprocal crosses give different results; with Mendel's tall and dwarf pea plants, the results were the same.

Offspring of the cross of P1 individuals are referred to as the first filial generation, or F_1. Mendel also referred to them as hybrids because they were the offspring of unlike parents (tall and dwarf). We will specifically refer to the offspring of tall and dwarf peas as monohybrids because they are hybrid for only one characteristic (height). Since all the F_1 **offspring** plants were tall, Mendel referred to tallness as the dominant trait.

Offspring are the young born of living organisms, produced either by a single organism or, in the case of sexual reproduction, two organisms.

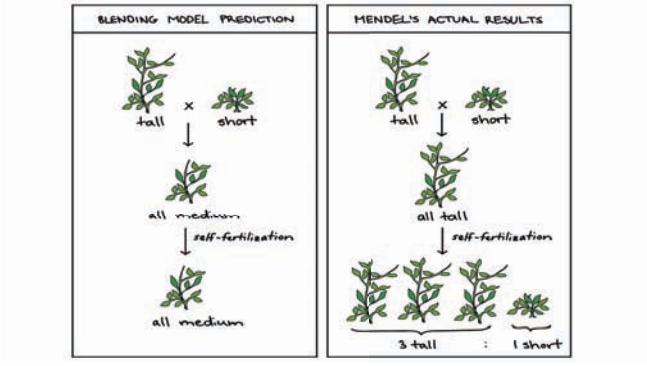


Figure 3: First two offspring generations from the cross of tall plants with dwarf plants.

The alternative, dwarfness, he referred to as recessive. Different forms of a gene that exist within a population are

termed alleles. The terms dominant and recessive are used to describe both the relationship between the alleles and the traits they control. Thus, we say that both the allele for tallness and the trait, tall, are dominant. Dominance applies to the appearance of the trait when both a dominant and a recessive allele are present. It does not imply that the dominant trait is better, is more abundant, or will increase over time in a population.

When the F_1 offspring of figure 3 were self-fertilized to produce the F_2 generation, both tall and dwarf offspring occurred; the dwarf characteristic reappeared. Among the F_2 offspring, Mendel observed 787 tall and 277 dwarf plants for a ratio of 2.84:1. It is an indication of Mendel's insight that he recognized in these numbers an approximation to a 3:1 ratio, a ratio that suggested to him the mechanism of inheritance at work in pea plant height.

Mendel's Crossing Experiments with Peas

Austrian monk Gregor Mendel began his series of crossing experiments with garden peas. By any standard, Mendel was a very gifted scientist who appreciated the importance of carefully designing and interpreting his experiments. Mendel sought to understand the patterns of inheritance of seven different characters in peas by crossing plants with identifiably different phenotypes of a particular character, and then following the presence or absence of the character in successive generations. The traits Mendel studied were morphological attributes of the seed, pod, flower and stem. The types of crosses Mendel made are still widely used in genetics today. Unfortunately, Mendel's discoveries were largely ignored during his lifetime, and it was not until 1900 that his work was rediscovered and its importance finally appreciated.

Monohybrid Cross

The first crosses Mendel made were between inbred lines of peas that were true breeding for opposite types of a character, for example, tall x dwarf, yellow x green seed, round x wrinkled seed coat, etc.

These are called monohybrid crosses because the peas were inbred lines that differed for a single character (Figure 4). In the progeny from the hybrid cross, called the F_1 generation, he observed that all offspring were always just one of the two alternate types, e.g. tall. Never were both types observed or were offspring of some intermediate type.

Mendel then self-pollinated some F_1 offspring to create an F_2 generation, whose plants segregated for both characters found in the inbred parents (Figure 4). It was at this point that Mendel was at his scientific best. Not only did he observe the character types in the offspring, but he also counted them and calculated ratios of each type. In the F_2 for all seven of the characters he was studying, he observed approximately three times as many plants of one-character type as the other. Furthermore, the type observed in the $F_1$was the one in greater number in the F_2.

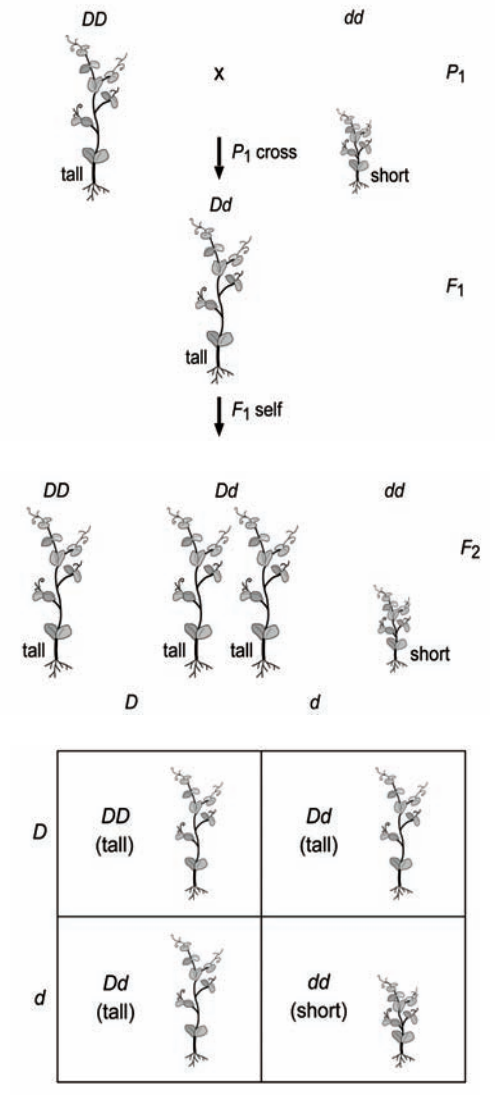


Figure 4: Example of one of Mendel's monohybrid crosses with peas.

The experiments involved three generations of plants. First, the true-breeding tall and short parental plants (P1) were crossed to form the F1 generation, which were uniformly tall plants. A single F1 generation plant was then self-pollinated to form the F2 generation. The Punnett square shows the F2 offspring resulting from selfing the F1. There are three possible genotypes (DD, Dd and dd), but only two phenotypes (tall and short) which segregate in a 3:1 ratio.

Mendel interpreted the results of his monohybrid crossing experiments to develop his first three laws of inheritance. He surmised that there must be one factor for tall and one factor for dwarf, and that these factors are somehow paired (Law 1, Table 1). Later this became the basis of the concept of pairs of homologous chromosomes with the gene determining a particular character located at the same position (i.e. locus) in each homologue. A locus can have different forms of the gene, which are called alleles. For example, in Mendel's peas there was one allele coding for tall plants

(D) and one allele coding for short plants (d) and these alleles segregated among the offspring. Plants that have the same allele at a locus on each of the homologous chromosomes are homozygous (e.g. DD and dd), whereas those with a different allele on each homologous chromosome are heterozygous (e.g. Dd).

Table 1: Mendel's laws of segregation and independent assortment

Law 1. Each genetic character is controlled by unit factors that come in pairs in individual organisms; these unit factors are now known as alleles.
Law 2. Paired unit factors for a given character are dominant and recessive to one another.
Law 3. Unit factors for the same character segregate independently from one another during the formation of gametes.
Law 4. Pairs of unit factors for two different traits also segregate independently of one another during the formation of gametes.

Mendel further hypothesized that one-unit factor (i.e. allele) is dominant to the other recessive factor (i.e. the dominant allele masks the effect of the recessive allele), based on the phenotypes he found in the F_1 and F_2 (Law 2, Table 1). For example, the allele for tall (D) is dominant and the allele for dwarf (d) is recessive. This leads to the important distinction between genotype and phenotype. The three possible genotypes are DD, Dd, and dd, although with D being dominant to d, there are only two phenotypes. The DD and Dd genotypes are both tall phenotypes and the dd genotype is the dwarf phenotype.

A simple method to interpret the results of the monohybrid cross is through the Punnett Square, named after its inventor Reginald Punnett (Figure 4). It can be seen that when a heterozygous F_1 is mated to itself (or crossed to an identical F_1), three genotypes (DD, Dd and dd) are found in the 1:2:1 ratio, respectively. However, only two phenotypes are found, tall and dwarf, in the ratio of 3:1, respectively.

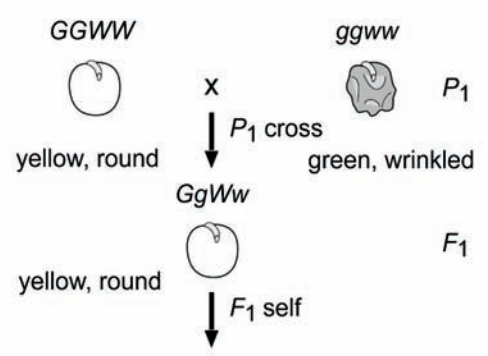
Finally, Mendel proposed that the 3:1 ratio observed in the progeny of selfed F1 plants is expected if the D and d alleles are transmitted from each parent to the offspring at random (Law 3, Table 1).

The frequencies of the offspring genotypes are the products of the frequencies of the alleles transmitted from each parent. The frequency of the DD genotype is 1/2 x 1/2 = 1/4, the frequency of the heterozygous Dd class is 2(1/2 x 1/2) = 1/2, and the frequency of dd is 1/2 x 1/2 = 1/4. Because D is dominant to d and therefore, DD and Dd have the same phenotype, the expected proportion of tall offspring is 3/4 and the dwarf offspring is 1/4 (or a 3:1 ratio). Mendel confirmed the expectations of random segregation with another type of cross, called a test cross. Here he crossed a Dd plant with a dd plant and found that approximately half of the offspring were tall and half were dwarf.

Dihybrid Cross

A second type of experiment Mendel performed was the dihybrid cross, which was

an extension of the monohybrid cross to two characters (Figure 5). Mendel crossed a yellow and round seeded type to a green and wrinkled seeded type. The F_1 offspring were all yellow and round and the alternate types of each character were found in 3:1 ratios in the F_2, as observed in monohybrid crosses involving the same traits.



	GW	Gw	gW	gw
GW	GGWW (yellow, round)	GGWw (yellow, round)	GgWW (yellow, round)	GgWw (yellow, round)
Gw	GGWw (yellow, round)	GGww (yellow, wrinkled)	GgWw (yellow, round)	Ggww (yellow, wrinkled)
gW	GgWW (yellow, round)	GgWw (yellow, round)	ggWW (green, round)	ggWw (green, round)
gw	GgWw (yellow, round)	Ggww (yellow, wrinkled)	ggWw (green, round)	ggww (green, wrinkled)

Figure 5: Example of one of Mendel's dihybrid crosses with peas. The Punnett square shows the offspring resulting from selfing the F_1. There are nine possible genotypes but only four phenotypes (yellow and round, yellow and wrinkled, green and round, green and wrinkled) which segregate in a 9:3:3:1 ratio.

Mendel also counted the two-character phenotypes and these were in a 9:3:3:1 ratio in the F_2. This is the expected ratio of two-character phenotypes given that both pairs of traits segregate randomly and completely independently of one another (Law 4, Table 1).

Under independent assortment, the expected frequency of any two-character phenotype is the product of the frequencies of each component character. Therefore, the frequency of yellow round offspring is f(yellow) x f(round) = 3/4 x 3/4 = 9/16 and the frequency of yellow wrinkled is 3/4 x 1/4 = 3/16. In this case, the genes controlling color and smoothness of seed coat reside on different chromosomes. It is shown that when genes controlling different characters reside close to each other on the same chromosome, they do not assort independently (called linkage).

Mendelian Inheritance of Traits in Forest Trees

A dihybrid cross describes a mating experiment between two organisms that are identically hybrid for two traits. A hybrid organism is one that is heterozygous, which means that is carries two different alleles at a particular genetic position, or locus

Simple Mendelian inheritance of morphological traits in forest trees is rarely observed. The reason for this is that forest geneticists and tree breeders have primarily concerned themselves with the genetic improvement of economic traits, such as stem volume and wood quality, which are under the control of many genes. Nevertheless, observant geneticists have identified a small number of traits showing Mendelian inheritance in the progeny of controlled crosses. Some interesting examples of Mendelian inheritance of morphological traits in trees include cone color in Pinus monticola, seedling foliage color in Picea abies, chlorophyll deficiencies and other morphological variants in Pinus taeda, diameter growth in Pinus patula, blister rust resistance in Pinus lambertiana, and the narrow-crown phenotype in P. abies. Almost certainly, many other morphological traits under simple genetic control could be identified if geneticists were to invest time in observing such traits in segregating populations.

Statistical Tests for Mendelian Inheritance

What if Mendel had observed only a small number of F_2's instead of several hundred or even thousands of offspring? Would he have observed the 3:1 and 9:3:3:1 ratio that led him to formulate the laws of inheritance? Genetic segregation is a random process subject to chance fluctuations. Small numbers of offspring can lead to observed ratios quite different than expected. If Mendel had looked at only ten F_2 plants following the crossing of a tall and a dwarf variety of peas, he might very easily have observed five tall plants and five dwarf plants just due to chance.

Since the early 1900s, geneticists have used statistical methods to test whether genetic segregation ratios conform to expected Mendelian ratios. The test most commonly used is a goodness-of-fit test to determine if observed ratios are significantly different from expected ratios. The goodness-of-fit test is performed by calculating a chi-square (χ^2) statistic:

$$\chi^2 = \Sigma\ (o - e)^2/e$$

where o equals the observed number and e equals the expected number in each class and the summation is over all offspring classes. The larger the deviation of the observed number from the expected number, the larger the magnitude of χ_2. A chi-square table must then be consulted to determine the statistical probability of the calculated value. The probability depends on the degrees of freedom (df) in the data, which when testing Mendelian segregation ratios is just the number of offspring classes minus one. The convention is to reject the hypothesis only if the probability of the calculated χ^2 is 0.05 or less.

Extensions to Mendel's Laws

Mendel's laws elegantly described the inheritance of a variety of characters in peas and provided the fundamental framework for the laws of inheritance for all characters in eukaryotic organisms. As with most things in biology, however, there are exceptions to the basic rules. In this section, we describe these exceptions and show how all are merely extensions to Mendel's laws.

Partial Dominance

Mendel's experiments established that at every locus there are two alleles, one dominant and one recessive, and that the ratio of dominant to recessive phenotypes in an F_2 is 3:1 when one locus is considered. Subsequent to Mendel's experiments with peas, the inheritance of characters was studied in crosses of many different organisms, and it was found that F_2's do not always segregate in a 3:1 ratio. One explanation for departure from the 3:1 ratio is incomplete or partial dominance

Consider the example of a gene controlling needle color in P. abies in a study conducted by Langner. In P. abies there is a mutant phenotype, known as the aureaform, that produces white needles in seedlings homozygous gg for the aurea allele. Trees heterozygous (Gg) for the aurea allele and the wild-type allele have light-green or golden needles, whereas wild-type trees (GG) have normal dark green needles. In this case, it appears that the wild-type allele is not dominant to the aurea allele or else Gg trees would have normal green needles. Instead they have a golden needle color that is intermediate between green and white, a so-called blended expression.

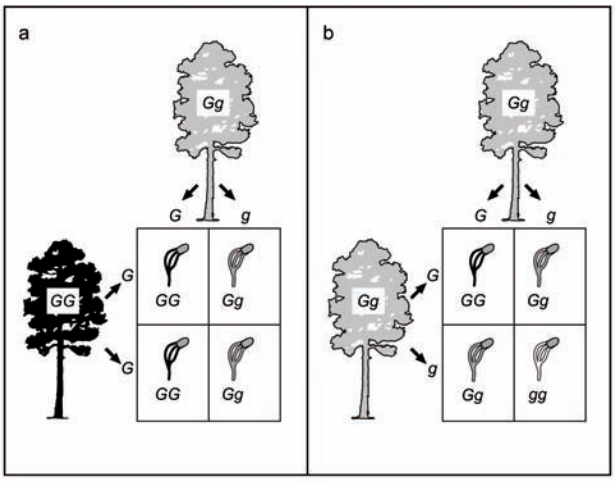


Figure 6: The partial dominance of needle color in Picea abies.

Langner established the inheritance of the locus-controlling needle color in P. abies by making three different crosses and calculating ratios of needle color phenotypes. The three crosses he made were GG x GG, GG x Gg, and Gg x Gg. As expected, he observed all green needle progeny in the GG x GG cross. In the GG x Gg cross and the Gg x Gg cross, however, he observed 1:1 and 1:2:1 ratios of green to golden and

green to golden to white needles, respectively (Figure 6). These phenotypic ratios are consistent only with partial dominance and not complete dominance.

Trees homozygous for the wild-type allele G are green and trees homozygous for the mutant aurea allele g are white. However, heterozygous trees (Gg) are not green as if G were dominant to g, but rather they are an intermediate color of light green or golden: (a) Crossing GG x Gg results in an expected 1:1 ratio of green: golden progeny; and (b) Crossing two heterozygotes results in an expected 1:2:1 ratio of normal: golden: white foliage in the progeny.

Codominance

In the previous example, which described the partial dominant inheritance of the needle color locus in P. abies, the wild-type (G) and aurea (g) alleles are said to act additively, because the phenotype of the heterozygote (Gg) is a blended result of their individual expression. Alternatively, it is possible to observe a separate phenotype of each of the two alleles at a heterozygous locus, which is known as codominance. A good example of codominance occurs with allozyme genetic markers. At a heterozygous allozyme locus, two distinct bands are observed on gels resulting from the visible and distinct expression of both alleles.

Epistasis

Deviation from Mendelian ratios can also occur due to the interaction between alleles at different loci. This is known as epistasis. Alleles at two loci can interact in antagonistic or complementary ways to modify the expression of phenotypes. When epistasis occurs, twolocus phenotypic ratios do not conform to the standard ratios expected under independent assortment.

Genetic Linkage

Mendel observed independent assortment among the seven characters he studied in peas, which suggests that all genes assort independently when transmitted from parent to offspring. However, as we learned in the previous discussion about meiosis, if the chromosome is the unit of heredity, then it would seem that all genes on a chromosome would be inherited as a unit and would not assort independently. This is known as genetic linkage, which means that there is a greater association of two or more non-allelic genes than would be expected from independent assortment. Linkage is complete if crossing over does not occur during meiosis and there is no recombination among genes on homologous chromosomes. Linkage is incomplete if crossing over and recombination occur, and the strength of the linkage is a function of the probability that a crossover will occur between two genes. This probability increases with increasing physical (i.e. map) distance between genes on a chromosome.

In the early 1900s, British geneticists William Bateson and Reginald Punnett were the first to observe genetic linkage between two flower color genes in peas; however, the understanding of genetic linkage was most developed by Thomas Morgan and

Alfred Sturtevant who studied the inheritance of characters in the common fruit fly, Drosophila melanogaster

The standard method for quantifying the amount of linkage between two genes is to estimate the recombination fraction (r):

$$r = R/N$$

where R equals the number of recombinant gametes and N is the total number of gametes in a sample of meioses. Two types of gametes can be distinguished: parental and recombinant gametes. Before gametes can be classified, however, the linkage phase of the two genes on a chromosome must be determined. For example, for two loci with alleles A and a and B and b, respectively, coupling linkage phase is when A and B are on one homologous chromosome and a and b are on the other homologous chromosome. Alternatively, repulsion linkage phase is when A and b and a and B are on opposite homologous chromosomes. Under coupling linkage phase, AB and ab are parental gametes and Ab and aB are recombinant gametes. Under repulsion linkage phase, the opposite is true; Ab and aB are parental gametes and AB and ab are recombinant gametes.

Organelle Genome Inheritance

Most of the genetic material is found in chromosomes in the nucleus of the cell. However, both chloroplasts and mitochondria have small genomes that encode genes related to their functions. The chloroplast DNA (cpDNA) and mitochondrial DNA (**mtDNA**) in plants are circular DNA molecules just like in prokaryotic genomes, such as bacteria. So how are these genomes inherited? Organelle DNA inheritance studies in plants were traditionally conducted by following the transmission of some mutant phenotype that was assumed to be cpDNA- or mtDNA-encoded, such as chlorophyll-deficient mutants. Studies of this type showed that cpDNA and mtDNA in plants are generally inherited from only a single parent, unlike nuclear chromosomes that are inherited from both parents. In angiosperms, it is nearly always true that both the cpDNA and the mtDNA are inherited from the maternal parent (Figure 7), although there are several cases where cpDNA is inherited from both parents. Both cpDNA and mtDNA are inherited through the mother in angiosperms. In all conifers and in Taxodiaceae and Cupressaceae, however, cpDNA is paternally inherited. The mtDNA is also paternally inherited in Taxodiaceae and Cupressaceae, but is maternally inherited in conifers.

mtDNA is the DNA located in mitochondria, cellular organelles within eukaryotic cells that convert chemical energy from food into a form that cells can use, adenosine triphosphate (ATP).

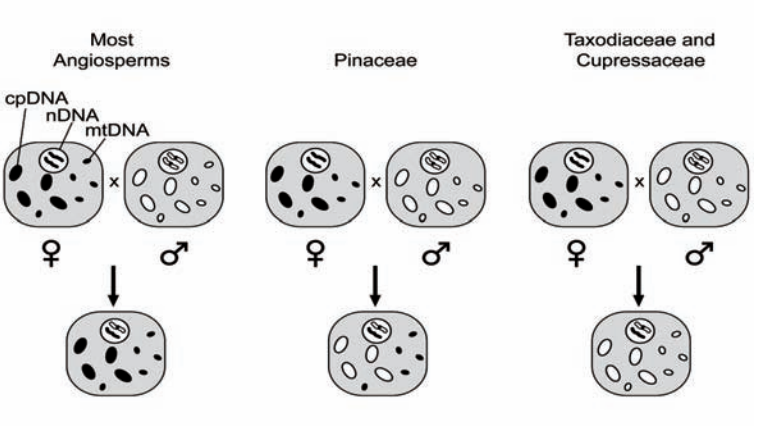
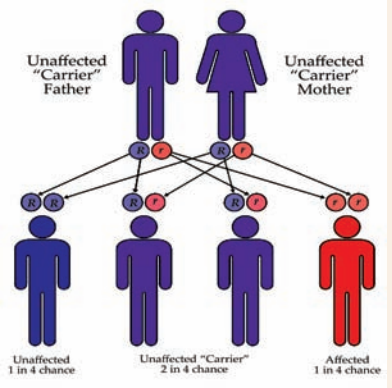


Figure 7: Chloroplast (cpDNA) and mitochondrial (mtDNA) genomes are inherited from one parent, rather than both parents as with the nuclear genome (nDNA).

Deoxyribonucleic acid-based genetic markers also have been used to study the inheritance of mtDNA in conifers. In the Pinaceae, mtDNA appears to be strictly inherited from the maternal parent as in angiosperms, although there is some evidence that mtDNA is occasionally inherited through the paternal parent. A most surprising result, however, is that mtDNA is strictly inherited from the paternal parent in Taxodiaceae and Cupressaceae. In no other plant or animal species has it been shown that the mtDNA is inherited strictly from the paternal parent.



CORE CONCEPT: SEGREGATION

Mendel's Law of Segregation states that a diploid organism passes a randomly selected allele for a trait to its offspring, such that the offspring receives one allele from each parent.

The essential feature of Mendelian genetics is the separation, technically called segregation, in unaltered form, of the two hereditary determinants in a hybrid plant in the formation of its reproductive cells.

Rule of Segregation

Mendel assumed that each plant contained two determinants (which we now call genes) for the characteristic of height. For example, a hybrid F_1 pea plant possesses the dominant allele for tallness and the recessive allele for dwarfness for the gene that determines plant height. A pair of alleles for dwarfness is required to develop the recessive phenotype. Only one of these alleles is passed into a single gamete, and the union of two gametes to form a zygote restores the double complement of alleles. The fact that the recessive trait reappears in the F_2 generation shows that the allele controlling it was hidden in the F_1 individual and passed on unaffected. This explanation of the passage of discrete trait determinants, or genes, comprises Mendel's first principle, the rule of segregation. The rule of segregation can be summarized as follows: A gamete receives only one allele from the pair of alleles an organism possesses; fertilization (the union of two gametes) reestablishes the double number. We can visualize this process by redrawing figure 8 using letters to denote the alleles. Mendel used capital letters to denote alleles that control dominant traits and lowercase letters for alleles that control recessive traits. Following this notation, T refers to the allele controlling tallness and t refers to the **allele** controlling shortness (dwarf stature).

Allele is the variant form of a given gene. Sometimes, different alleles can result in different observable phenotypic traits, such as different pigmentation. However, most genetic variations result in little or no observable variation.

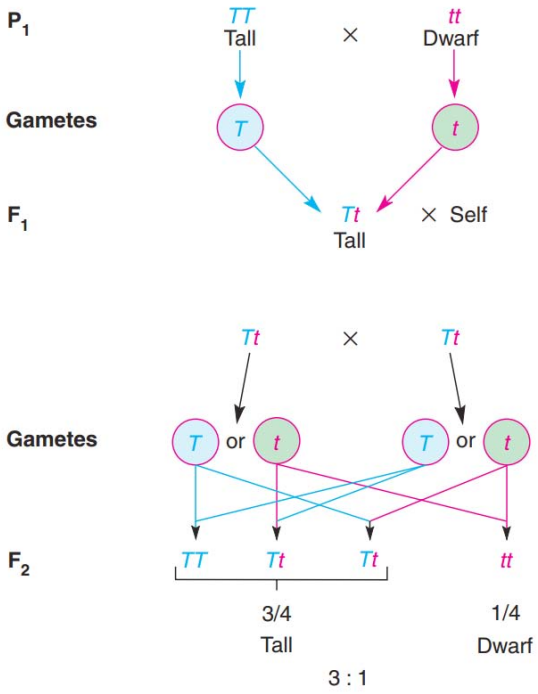


Figure 8: Assigning genotypes to the cross in figure 3.

From figure 3, we can see that Mendel's rule of segregation explains the homogeneity of the F_1 generation (all tall) and the 3:1 ratio of tall-to-dwarf offspring in the F_2 generation. Let us define some terms. The genotype of an organism is the gene combination it possesses. In figure 8, the genotype of the parental tall plant is TT; that of the F1 tall plant is Tt. Phenotype refers to the observable attributes of an organism. Plants with either of the two genotypes TT or Tt are phenotypically tall. Genotypes come in two general classes: homozygotes, in which both alleles are the same, as in TT or tt, and heterozygotes, in which the two alleles are different, as in Tt. William Bateson coined these last two terms in 1901. Danish botanist Wilhelm Johannsen first used the word gene in 1909.

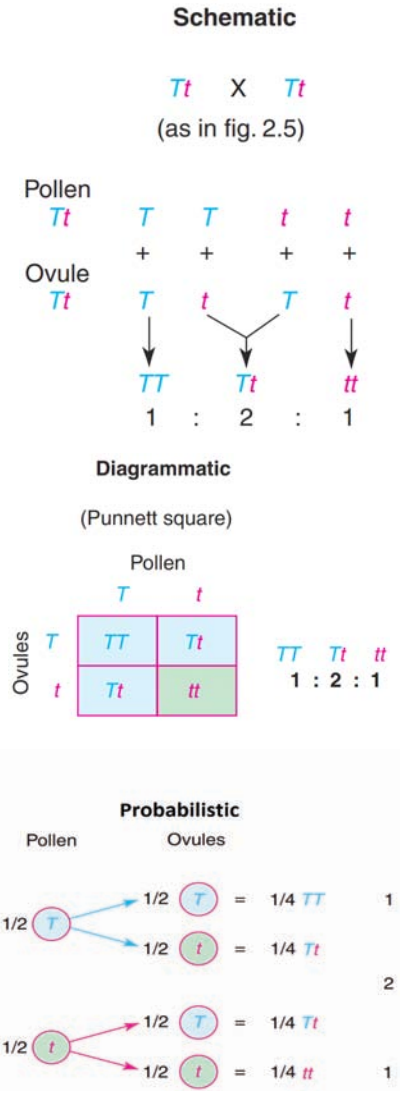


Figure 9: Methods of determining F_2 genotypic combinations in a self-fertilized monohybrid. The Punnett square diagram is named after the geneticist Reginald C. Punnett.

If we look at figure 8, we can see that the TT homozygote can produce only one type of gamete, the T-bearing kind, and the tt homozygote can similarly produce only

t-bearing gametes. Thus, the F_1 individuals are uniformly heterozygous Tt, and each F1 individual can produce two kinds of gametes in equal frequencies, T- or t-bearing. In the F_2 generation, these two types of gametes randomly pair during fertilization. Figure 9 shows three ways of picturing this process.

Testing the Rule of Segregation

We can see from figure 9 that the F_2 generation has a phenotypic ratio of 3:1, the classic Mendelian ratio. However, we also see a genotypic ratio of 1:2:1 for dominant homozygote: heterozygote: recessive homozygote. Demonstrating this genotypic ratio provides a good test of Mendel's rule of segregation.

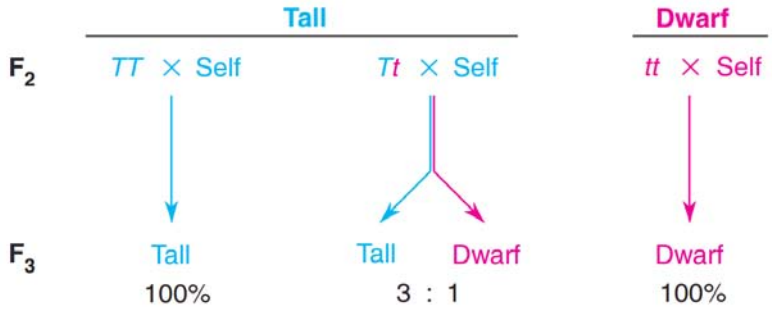


Figure 10: Mendel self-fertilized F_2 tall and dwarf plants. He found that all the dwarf plants produced only dwarf progeny. Among the tall plants, 72% produced both tall and dwarf progeny in a 3:1 ratio.

The simplest way to test the hypothesis is by progeny testing, that is, by self-fertilizing F_2 individuals to produce an F3 generation, which Mendel did (figure 10). Treating the rule of segregation as a hypothesis, it is possible to predict the frequencies of the phenotypic classes that would result. The dwarf F_2 plants should be recessive homozygotes, and so, when selfed (self-fertilized), they should produce only t-bearing gametes and only dwarf offspring in the F_3 generation. The tall F2 plants, however, should be a heterogeneous group, one-third of which should be homozygous TT and two-thirds heterozygous Tt. The tall homozygotes, when selfed, should produce only tall F_3 offspring (genotypically TT). However, the F_2 heterozygotes, when selfed, should produce tall and dwarf offspring in a ratio identical to that the selfed F_1 plants produced: three tall to one dwarf offspring. Mendel found that all the dwarf (homozygous) F_2 plants bred true as predicted. Among the tall, 28% (28/100) bred true (produced only tall offspring) and 72% (72/100) produced both tall and dwarf offspring. Since the prediction was one-third (33.3%) and two-thirds (66.7%), respectively, Mendel's observed values were very close to those predicted. We thus conclude that Mendel's progeny-testing experiment confirmed his hypothesis of segregation. Another way to test the segregation rule is to use the extremely useful method of the testcross, that is, a cross of any organism with a recessive homozygote. (Another type of cross, a backcross, is the cross of a progeny with an individual that has a parental genotype.

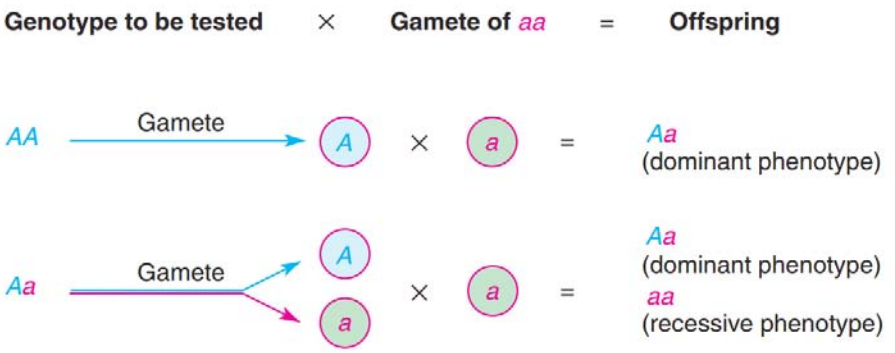


Figure 11: Testcross. In a testcross, the phenotype of an offspring is determined by the allele the offspring inherits from the parent with the genotype being tested.

Hence, a testcross can often be a backcross.) Since the gametes of the recessive homozygote contain only recessive alleles, the alleles that the gametes of the other parent carry will determine the phenotypes of the offspring. If a gamete from the organism being tested contains a recessive allele, the resulting F_1 organism will have a recessive phenotype; if it contains a dominant allele, the F_1 organism will have a dominant phenotype. Thus, in a testcross, the genotypes of the gametes from the organism being tested determine the phenotypes of the offspring (figure 11). A testcross of the tall F_2 plants in figure 8 would produce the results shown in figure 8. These results further confirm Mendel's rule of segregation.

Tall (two classes)

$TT \times tt$ = all Tt

$Tt \times tt$ = $Tt : tt$

1 : 1

Figure 12: Testcrossing the dominant phenotype of the F_2 generation from figure 8.



CORE CONCEPT: DOMINANCE

Mendel's Law of Dominance states that recessive alleles will always be masked by dominant alleles. Therefore, a cross between a homozygous dominant and a homozygous recessive will always express the dominant phenotype, while still having a heterozygous genotype.

If dominance were universal, the heterozygote would always have the same phenotype as the dominant homozygote, and we would always see the 3:1 ratio when heterozygotes are crossed. If, however, the heterozygote were distinctly different from both homozygotes, we would see a 1:2:1 ratio of phenotypes when heterozygotes are crossed. In partial dominance (or incomplete dominance), the phenotype of the heterozygote falls between those of the two homozygotes. An example occurs in flower petal color in some plants.

Using four-o'clock plants (**Mirabilis jalapa**), we can cross a plant that has red flower petals with another that has white flower petals; the offspring will have pink flower petals. If these pink-flowered F1 plants are crossed, the F_2 plants appear in a ratio of 1:2:1, having red, pink, or white flower petals, respectively (figure 13). The pink-flowered plants are heterozygotes that have a petal color intermediate between the red and white colors of the homozygotes. In this case, one allele (R_1) specifies red pigment color, and another allele specifies no color (R_2; the flower petals have a white background color). Flowers in heterozygotes (R_1R_2) have about half the red pigment of the flowers in red homozygotes (R_1R_1) because the heterozygotes have only one copy of the allele that produces color, whereas the homozygotes have two copies.

Mirabilis jalapa is the most commonly grown ornamental species of Mirabilis plant, and is available in a range of colours.

As technology has improved, we have found more and more cases in which we can differentiate the heterozygote. It is now clear that dominance and recessiveness are phenomena dependent on which alleles are interacting and on what phenotypic level we are studying. For example, in Tay-Sachs disease, homozygous recessive children usually die before the age of three after suffering severe nervous system degeneration; heterozygotes seem to be normal. As biologists have discovered how the disease works, they have made the detection of the heterozygotes possible.

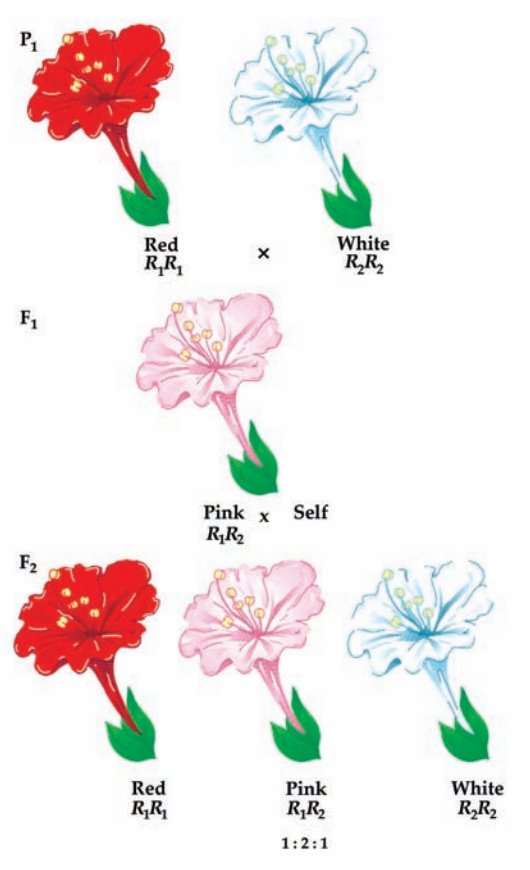
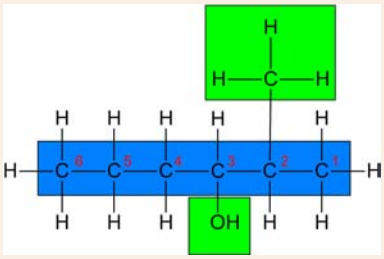


Figure 13: Flower color inheritance in the four-o'clock plant: an example of partial, or incomplete, dominance.

As with many genetic diseases, the culprit is a defective enzyme (protein catalyst). Afflicted homozygotes have no enzyme activity, heterozygotes have about half the normal level, and, of course, homozygous normal individuals have the full level. In the case of Tay-Sachs disease, the defective enzyme is hexoseaminidase-A, needed for proper lipid metabolism. Modern techniques allow technicians to assay the blood for this enzyme and to identify heterozygotes by their intermediate level of enzyme activity. Two heterozygotes can now know that there is a 25% chance that any child they bear will have the disease. They can make an educated decision as to whether or not to have children.

The other category in which the heterozygote is discernible occurs when the heterozygous phenotype is not on a scale somewhere between the two homozygotes, but actually expresses both phenotypes simultaneously. We refer to this situation as codominance. For example, people with blood type AB are heterozygotes who express both the A and B alleles for blood type. Electrophoresis lets us see proteins directly and also gives us many examples of codominance when we can see the protein products of both alleles.



CORE CONCEPT: NOMENCLATURE

Nomenclature is a system of names or terms, or the rules for forming these terms in a particular field of arts or sciences. The principles of naming vary from the relatively informal conventions of everyday speech to the internationally agreed principles, rules and recommendations that govern the formation and use of the specialist terms used in scientific and any other disciplines.

Throughout the last century, botanists, zoologists, and microbiologists have adopted different methods for naming alleles. Botanists and mammalian geneticists tend to prefer the capital-lowercase scheme. Drosophila geneticists and microbiologists have adopted schemes that relate to the wild-type. The wild-type is the phenotype of the organism commonly found in nature. Though other naturally occurring phenotypes of the same species may also be present, there is usually an agreed-upon common phenotype that is referred to as the wild-type. For fruit flies (*Drosophila*), organisms commonly used in genetic studies, the wild-type has red eyes and round wings (figure 14).



Figure 14: Wild-type fruit fly, *Drosophila melanogaster*.

Alternatives to the wild-type are referred to as mutants (figure 15). Thus, red eyes are wild-type, and white eyes are mutant. Fruit fly genes are named after the mutant, beginning with a capital letter if the mutation is dominant and a lowercase letter if it is recessive.

Table 2 gives some examples. The wild-type allele often carries the symbol of the mutant with a + added as a superscript; by definition, every mutant has a wild-type allele as an alternative. For example, w stands for the white-eye allele, a recessive mutation. The wild-type (red eyes) is thus assigned the symbol w+. Hairless is a dominant allele with the symbol H. Its wild-type allele is denoted as H+. Sometimes geneticists use the + symbol alone for the wild-type, but only when there will be no confusion about its use.

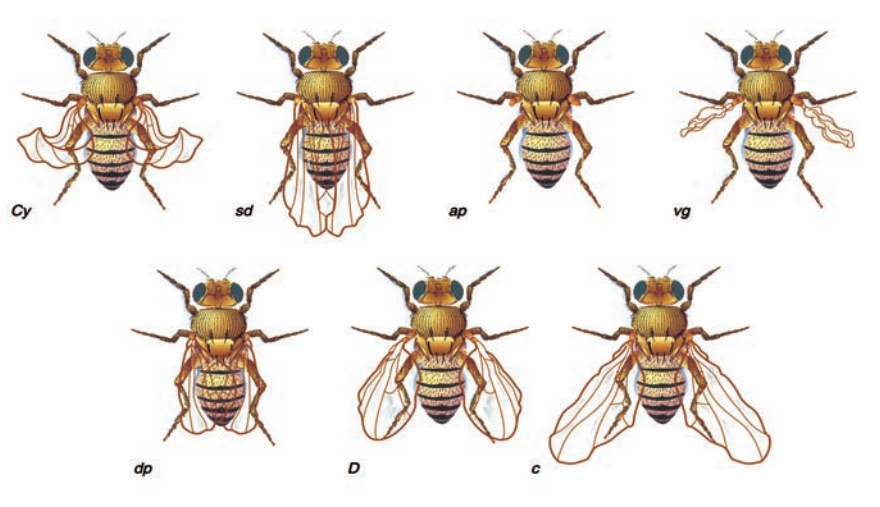


Figure 15: Wing mutants of *Drosophila melanogaster* and their allelic designations: Cy, curly; sd, scalloped; ap, apterous; vg, vestigial; dp, dumpy; D, Dichaete; c, curved.

If we are discussing eye color only, then + is clearly the same as w+: both mean red eyes. However, if we are discussing both eye color and bristle morphology, the + alone could refer to either of the two aspects of the phenotype and should be avoided.

Table 2: Some Mutants of Drosophila

Mutant Designation	Description	Dominance Relationship to Wild-Type
abrupt (*ab*)	Shortened, longitudinal, median wing vein	Recessive
amber (*amb*)	Pale yellow body	Recessive
black (*b*)	Black body	Recessive
Bar (*B*)	Narrow, vertical eye	Dominant
dumpy (*dp*)	Reduced wings	Recessive
Hairless (*H*)	Various bristles absent	Dominant
white (*w*)	White eye	Recessive
white-apricot (w^a)	Apricot-colored eye (allele of white eye)	Recessive

Classification and Nomenclature of Organism

Classification determines methods for organizing the diversity of life on Earth.

Since life first appeared on Earth 3.5 billion years ago, many new types of organisms have evolved.

Many of these organisms have become extinct, while some have developed into the present flora and fauna of the world. It is believed that the extinct species may outnumber the living ones by 50 to 100 times. At the same time a large number of species are still unknown and yet to be identified and described. Every year about

15,000 new species are discovered. Thus, extinction and diversification continue nonstop on earth, and some of these we know by the fossils (impressions or dead remains of past organisms) they left behind in the rocks. In order to study the vast diversified world of living and extinct organisms it is essential to classify them into groups (categories) so that they could be named, remembered and studies properly.

Basics in Biological Classification

Biological Classification involves the techniques of characterization, identification, nomenclature and grouping of organisms. Sometimes, the terms such as biological classification, taxonomy and systematic are used as synonyms. However, G. Simpson regards them as separate fields.

He explained classification as the ordering of organisms into groups and taxonomy as the study of principles and procedures of classification. He considered systematics as the study of organisms and their diversities and the evolutionary relationships among them. The scope of systematics includes identification, nomenclature and classification.

Need of Classification

Classification is the arrangement of organisms into groups and sub-groups on the basis of similarities and dissimilarities and placing them in a hierarchy that brings out their relationships.

The organisms should be classified due to following reasons:

(i) Classification makes identification and the study of a wide variety of organisms easy.

(ii) It reveals inter-relationships among different groups of organisms.

(iii) It gives information about the organisms and fossils of other localities.

Importance of Classification

The contribution of classification is complementary to all branches of basic and applied biology directly or indirectly.

1. Understanding the diversity:

   Classification helps in knowing about bio-resources, their diversity and evolution.

2. Understanding phylogeny:

Archaefructus is an extinct genus of herbaceous aquatic seed plants with 3 known species.

Classification based on the studies of fossil reveal evolutionary relationships (phylogeny) among the organisms. For example, the fossil of Cooksonia that appeared about 420 million years ago was the first known vascular land plant from which other vascular plan is have descended. Similarly the **Archaefructus** and Amborella are the oldest known angiosperms in the fossil record.

3. Inter-relations:

   Classification is essential to understand the inter-relationships among different categories of organisms.

4. Development of other Biological Science:

   Classification provides information for the development of all branches of biology, e.g., biogeography, ecology, ethology, forestry etc.

5. In agriculture and forestry:

   Taxonomic knowledge provides vital information about the pests and hosts that is essential for a cultivator before undertaking any control measures.

6. Mineral Prospecting:

   Taxonomy gives the correct sequence of geological events in an area. It is basic to any search for fossil fuels and mineral deposits.

7. Biological Control:

   Taxonomic knowledge is essential for accurate identification of pests and its natural enemies. Now a days instead of insecticides (chemical control), natural enemies of pest are introduced for biological control. Thus, correct identification of target species ensures a effective control at minimum costs. It would otherwise confusing, expensive and even chances of introducing wrong control agents.

8. Conservation of wild life:

   Due to human greed and ignorance, many species of flora and fauna have become extinct and many others are following suit. Taxonomists help to identify these endangered species and help in initiating the wild life conservation programmes.

9. Quarantine:

   The spreading of new pests and diseases from one country to the other through the agency of human is called quarantine. It can be checked by establishing quarantine laboratories at airports, seaports, railway stations etc. Here, taxonomists provide correct and prompt identification of the pest of disease.

Nomenclature (L. Nome = Name; Calere = To Call):

The term nomenclature means the scientific naming of organisms according to an established system. The naming of plants on a

scientific basis is called botanical or plant nomenclature. In earlier days common or vernacular names were in use which generally changes with change of language. Later in the haunt of one common internationally accepted name for a species, scientific names (Technical names) have been introduced in form of polynomial, binomial and trinomial systems of nomenclature.

At least, nomenclature is regarded as a part of (though distinct from) taxonomy. Taxonomy can be defined as the study of classification including its principles, procedures and rules, while classification itself is the ordering of taxa (the objects of classification) into groups based on similarities or differences.

(a) Polynomial Nomenclature:

Before 1750, taxonomists started using a string of descriptive Latin words to designate a species. For example, in the herbal of Clusius a species of willow is named Salix pumila angustifolia altera.

However polynomial nomenclature was discarded for two reasons:

(i) It was lengthy and difficult to remember

(ii) It often differs from scholar to scholar based on the characters chosen by them.

(b) Trinomial Nomenclature:

Mayer introduced the concept of sub-species which meant the geographically defined aggregates of local populations. The naming of a species, especially the animals, up to sub-species level is called trinomial nomenclature, e.g. Homo sapiens sapiens, Homo sapiens Neanderthals. Botanists don't consider trinomial nomenclature.

Binomial Nomenclature

Carolus Linnaeus the great Swedish naturalist, devised binomial system of nomenclature in his book Philosophia Botanica. Now the binomial system of scientific naming of organisms becomes a common and established practice.

According to binomial nomenclature, the scientific name of an organism composed of two Latin or Latinized words i.e. the first word is called genus (= generic name or generic epithet) followed by the second word called species (= specific name or specific epithet). For example, the botanical name of sugarcane is Saccharum officinarum. Very rarely the generic and specific names are same. They are called tautonymes, e.g. Gorilla gorilla, Catla catla, Naja naja, Rattus rattus etc. Some organisms (fossils and most microorganisms) are known by their technical names only.

Binomial nomenclature is the system used to identify all organisms on Earth, from elephants to algae. A binomial or scientific name identifies an organism by its genus and species, ensuring that everyone understands which organism is under discussion. Binomial nomenclature fits within the larger framework of taxonomy, the science of categorizing living organisms and assigning traits to them to understand the links and differences between them. The scientific name of an organism could be considered

its definitive name, with scientific names being understood by scientists all over the world.

You may also hear scientific names being referred to as "Latin names," in a reference to the heavy use of Latin in taxonomy. However, it is also common to see Latinized names, typically honoring the person who discovered the organization, or the region in which it was discovered; for example, Branta canadensis is the Canadian Goose. Greek is also used in scientific names, often in a jumble with Latin which brings some classical scholars to tears.

The system of binomial nomenclature was developed by Carolus Linnaeus, an 18th century scientist who attempted to codify the natural world with a taxonomic system. Various taxonomic systems had been used before this point, but Linnaeus established a flexible, easy to use system which caught on rapidly. Taxonomy was actually largely undisciplined until the 19th century, when people began to establish codes and organizations to oversee the field of taxonomy. When new organisms are discovered, they are reported to these organizations to ensure that the discovery is, in fact new, allowing a new name to be generated.

It can be helpful to know about some of the conventions used in regards to binomial nomenclature. For example, scientific names are always given with the genus capitalized, in italics, like this: Genus species. In scientific journals, credit is given to the person who discovered the organism in parentheses after the first listing of the scientific name, like this: Example animal(Jones, 1997). When the common name of an organism is given, the scientific name follows in parentheses, as in this example: "The Common Wombat (Vombatus ursinus) lives in Australia."

The name of the genus is always spelled out, unless you mention an organism's scientific name more than once in a written document, in which case you may turn it into an initial, like this: "The biology of the Atlantic giant squid Architeuthis dux is not fully understood, but scientists hope that further studies of A. dux and its cousin, the Southern giant squid (A. sanctipauli) will yield more information on these fascinating creatures." Common usages like "E. coli" are frowned upon by the conventions of binomial nomenclature, with scientists preferring to see Escherichia coli written out in any discussion of this fascinating bacterium.

In zoology, taxonomy is overseen by the International Commission on Zoological Nomenclature (ICZN), with equivalent bodies for botany, bacteria, and viruses. These groups all apply specific rules and codes to the scientific names they oversee, ensuring uniformity within their fields. Taxonomy is also by no means set in stone; organisms may move between genera, for example, as more information is gathered about them.

Rules of Binomial nomenclature

The rules of nomenclature are framed and standardized by five separate codes such as:

(i) International Code of Botanical Nomenclature (ICBN),

(ii) International code of Zoological Nomenclature (ICZN),

(iii) International Code of Bacteriological Nomenclature (ICBacN),

(iv) International Code of Viral Nomenclature (ICVN) and

(v) International Code of Nomenclature for Cultivated Plants (1CNCP).

These codes help in avoiding errors, duplication, confusion and ambiguity in scientific names.

Some rules framed under these codes as well as the rules set by Linnaeus are as follows:

1. The scientific names of plants and animals should be in Latin or Greek because it is officially dead language. The use of Latin for naming also means that no-one can be offended by being forced 10 use someone else's language.
2. The scientific names prior to the 1.8.1758 for animals of Systema Naturae (10th Edition) and 1.5.1753 for plants of Species Plantarum by Linnaeus are not recognized.
3. These names should be in italics when printed or separately underlined when hand written to indicate their Latin origin.
4. The genus starts with capital letter, while species in small letter.
5. The name of the author, first reporting it should remain in abbreviated form at the end of the scientific name and it is printed in Roman, e.g. Oriza sativa Linn. The scientific name with name of the author at the end is called complete scientific name.
6. Only one valid name for one species is permitted and it is based on the rule of priority that is the author first effectively and validly publishing the name will be considered.
7. In case of changing a scientific name that is double citation, the name of the second author is placed in bracket after the scientific name and the first author's name in abbreviated form comes after that, the new name is always based on the older name and it is called the **basionym**.
8. To avoid confusion no two generic names in any kingdom can be same. Specific names

Basionym means the original name on which a new name is based.

can however be repeated as they often qualify the generic name. For example, the specific name of both mango (Mangifera indica) and tamarind (Tamarindus indicus) are the same which mean Indian.

9. The reporting of a new species of plant should be accompanied by a Latin description or Diagnosis.

10. When a plant species is reported, the author should submit a herbarium sheet of the specimen (Dried plant with reproductive part placed on a sheet of paper). This is designated as type specimen (holotype, isotype, paratype, topotype, lectotype etc.). Holotype is the type specimen submitted by the original author at the time of publication. All other specimens of the same species collected at the same time are called isotypes.

The specimen cited with the original description other than holotype or isotype(s) is called paratype. When the same specimen is collected from the same locality from where the holotype was collected is called topotype. In absence of a holotype, the type specimen selected from the original material by a subsequent author is called lectotype. In the absence of original type specimen, a new specimen selected from a new location by a new author is called neotype. This concept of permanent naming to a type specimen is called typification and the type specimens should be preserved in the Herbaria of all international Botanic Gardens.

Hierarchy of Classification

Classification is not a single step process but solves hierarchy of steps in which each step represents a rank or category. A category is an abstract term which basically represents a rank or taxon (plural: taxa). A taxon represents a group of real biological objects having a set of common features called correlated characters. For example, Monera is a taxon of bacteria belonging to category kingdom.

All the categories arranged in ascending order constitute a taxonomic hierarchy. There are seven obligate categories used in taxonomical studies of all known organisms which are: Kingdom, Division (for plants) or Phylum (for animals), class, order, family, genus and species. These broad categories have been further divided into sub-categories by using prefix super-sub-or infra-to facilitate more sound and scientific placement of various taxa. As we go from the species to kingdom, the number of correlated characters goes on decreasing.

Linnaeus first introduced the hierarchical system of classification and is called Linnaean hierarchy. In his classification two categories i.e. phylum and family were not included. In taxonomic hierarchy, species is the basic unit of classification and is the lowest category while kingdom is the top most category.



CORE CONCEPT: MULTIPLE ALLELES

A population or species of organisms typically includes multiple alleles at each locus among various individuals. Allelic variation at a locus is measurable as the number of alleles (polymorphism) present, or the proportion of heterozygotes in the population.

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.

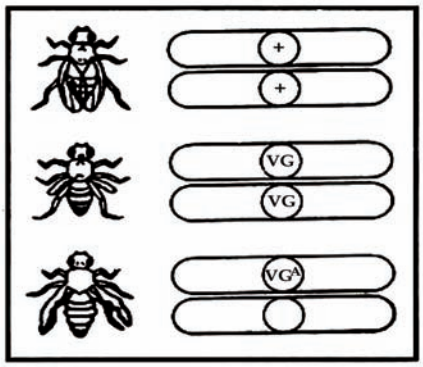


Figure 16: Multiple alleles Different from of wing in Drosophila.

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 +.

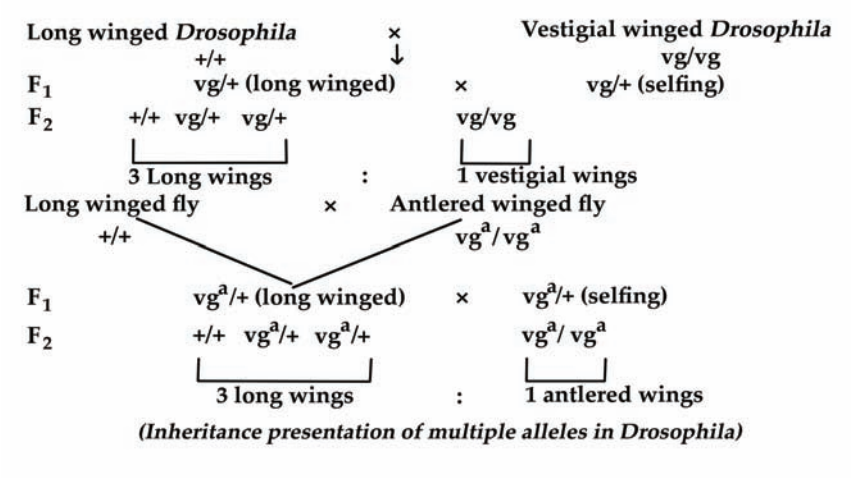
Thus, there are three races of Drosophila:

(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 sometimes 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
	(vestigal)	(intermediate)	(antlered)
	1 :	2 :	1

(Cross representation between a vestigal 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.

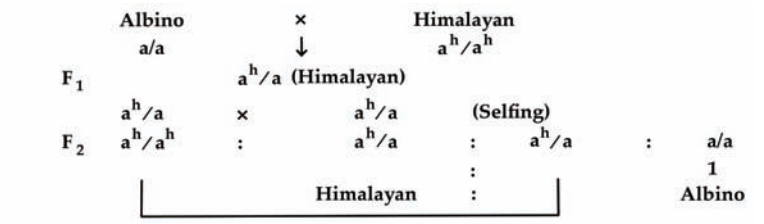
Another example of multiple alleles is the eye colour in Drosophila. The normal colour of the eye is red. Mutation changed this red eye colour to white. Other mutations at white locus took place changing the red eye colour to various lighter shades like cherry, apricot, eosin, creamy, ivory, blood etc.,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



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

Self-Sterility in Plants

Nicotiana longiflora is a species of tobacco native to South America that is sometimes cultivated for its tubular flowers that emit a very sweet odour at night.

Kolreuter (1764) described self- sterility in tobacco (**Nicotiana longiflora**). The reason was done by East. He described that self-sterility is due to series of alleles designated as s_1, s_2, s_3 and s_4 etc. The hybrids S_1/S_2 or S_1/S_3 or S_3/S_4 are self-sterile because pollen grains from these varieties did not develop, but pollens of S_1/S_2 were effective and capable of fertilization with S_3/S_4.

The genes causing self-sterility in plants probably produce their effects by controlling the growth rate of the pollen tubes. In compatible combinations, the pollen tube grows more and more rapidly as it approaches the ovule, but in non-suitable ones, the growth of the pollen tube slows down considerably, so that the flower withers away before fertilization can take place.

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 peculiear 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^B i
A	I^AI^A or I^A i
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.

Table 3: Table showing possible blood types of children from parents of various blood groups

PARENTS		CHILDREN	
Phenotypes	Genotypes	Phenotypes	Genotypes
0 X 0	ii X ii	0	ii
0 X A	ii X I^AI^A or I^Ai	0, A	ii, l^Ai
0 X B	ii X I^BI^B or I^Bi	0, B	i, 1"i
0 X AB	ii X 1^AI^B	A. R	I^Ai, I^Bi
A X A	I^AI^A or I^Ai X I^AI^A or I^Ai	A, 0	I^AI^A, ii

A X B	I^AI^A or I^Ai X I^BI^B or I^Bi	A, AB, 0, B	l^Ai, I^A1, ii, l^Bi
A X AB	I^AI^A or l^Ai X I^AI^B	A, B, AB	IAIA 115, op
B X B	I^BI^Bor IBI X I^BI^B or I^Bi	B, 0	I^BI^B, ii
B X AB	1^B1^Bor 1^Bi X I^AI^B	A, B, AB	I^AI^A,I^BI^B,I^AI^B
AB X AB	1^A1^B X I^AI^B	A, B, AB	I^AI^A,I^BI^B,I^AI^B

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 anaemia, known as Erythroblastosis foetalis, which occurs occasionally in new born infants.

It was found that the infants suffering from this anaemia 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 anaemia; this may be serious enough to cause the death of the newborn infant or abortion of the foetus.

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.

Pleiotropism

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.

Theories of Allelism

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

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.

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.

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.

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.

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 % penetrance. If its phenotypic effect is produced only 60 % of the individuals that contains it then it is said to show 60 % 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 venture 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.

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.

Phenocopy

Characters are the result of interaction between the genotype and the environment. When a gene mutates, its phenotypic effect also changes. Sometimes, 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**.

Genocopy is a trait that is a phenotypic copy of a genetic trait but is caused by a different genotype.

Xenia and Metaxenia

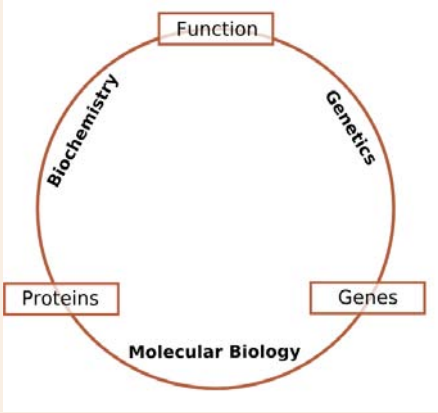
The immediate effect of foreign pollen on visible characters of the endosperm is called xenia. The 'xenia' term was given by Focke. This has been studied in maize plant. If a white endosperm variety is open pollinated in the field where there are also plants of the yellow endosperm variety then the cobs 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.

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 flavour 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.



CORE CONCEPT: BIOCHEMICAL GENETICS

A biochemical genetics laboratory differs from the clinical chemistry laboratory in the extent of interpretation that is necessary to make its results meaningful to the clinician. Testing for hereditary metabolic disorders has developed from a highly specialized and fragmented activity, provided mostly by research-oriented scientists, to a critical component at the forefront of the laboratory work-up of patients of all ages.

Biochemical genetics involves diagnosing and treating metabolic diseases. These are problems with how the body makes, breaks down or uses proteins, fats or carbohydrates. They are caused by genes that are not normal and cannot make the enzymes the body needs. (Enzymes are found in the blood, organs, saliva and other parts of the body. They are important to all bodily functions.) These diseases are called inborn errors of metabolism.

Inborn Errors of Metabolism

The examples of mouse coat color, corn kernel color, and snapdragon flower petal color demonstrate that genes control the formation of enzymes, proteins that control the steps in biochemical pathways. For the most part, dominant alleles control functioning enzymes that catalyze biochemical steps. Recessive alleles often produce nonfunctioning enzymes that cannot catalyze specific steps. Often a heterozygote is normal because one allele produces a functional enzyme; usually only half the enzyme quantity of the dominant homozygote is enough. The study of the relationship between genes and enzymes is generally called biochemical genetics because it involves the genetic control of biochemical pathways. A. E. Garrod, a British physician, pointed out this general concept of human gene action in Inborn Errors of Metabolism, published in 1909. Only nine years after Mendel was rediscovered, Garrod described several

human conditions, such as albinism and alkaptonuria, that occur in individuals who are homozygous for recessive alleles (see figure 17).

Table 4: Some examples of epistatic interactions among alleles of two genes

Characteristic	Phenotype of F_1 Dihybrid (*AaBb*)	Phenotypic F_2 Ratio
Corn and sweet pea color	Purple	Purple:white 9:7
Mouse coat color	Agouti	Agouti:black:albino 9:3:4
Shepherd's purse seed capsule shape	Triangular	Triangular:oval 15:1
Summer squash shape	Disk	Disk:sphere:elongate 9:6:1
Fowl color	White	White:colored 13:3

For example, people normally degrade homogentisic acid (alkapton) into maleylacetoacetic acid. Persons with the disease alkaptonuria are homozygous for a nonfunctional form of the enzyme essential to the process: homogentisic acid oxidase, found in the liver. Absence of this enzyme blocks the degradation reaction so that homogentisic acid builds up. This acid darkens upon oxidation. Thus, affected persons can be identified by the black color of their urine after its exposure to air. Eventually, alkaptonuria causes problems in the joints and a darkening of cartilage that is visible in the ears and the eye sclera.

One-Gene-One-Enzyme Hypothesis

Pioneering work in the concept that genes control the production of enzymes, which in turn control the steps in biochemical pathways, was done by George Beadle and Edward Tatum, who eventually shared the Nobel Prize for their work. They not only put forth the one-gene-one enzyme hypothesis, but also used mutants to work out the details of biochemical pathways. In 1941, Beadle and Tatum were the first scientists to isolate mutants with nutritional requirements that defined steps in a biochemical pathway. In the early 1940s, they united the fields of biochemistry and genetics by using strains of a bread mold with specific nutritional requirements to discover the steps in biochemical pathways in that organism.

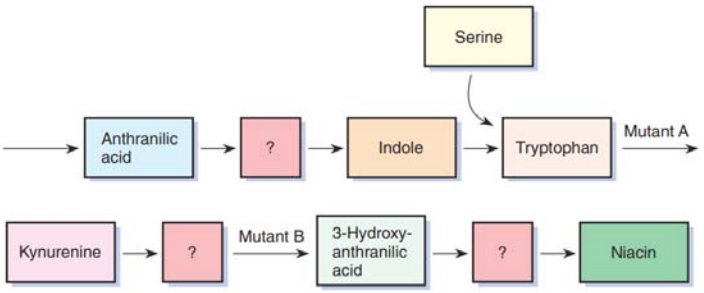


Figure 17: Pathway of niacin synthesis in Neurospora. Each arrow represents an enzyme-mediated step. Each question mark represents a presumed but (at the time Beadle and Tatum were working) unknown compound.

Through this century, the study of mutations has been the driving force in genetics. The process of mutation produces alleles that differ from the wild-type and shows us that a particular aspect of the phenotype is under genetic control. Beadle and Tatum used mutants to work out the steps in the biosynthesis of niacin (vitamin B3) in pink bread mold, Neurospora crassa.

Normally, Neurospora synthesizes niacin via the pathway shown in figure 17. Beadle and Tatum isolated mutants that could not grow unless niacin was provided in the culture medium; these mutants had enzyme deficiencies in the synthesizing pathway that ends with niacin. Thus, although wild-type Neurospora could grow on a medium without additives, the mutants could not. Beadle and Tatum had a general idea, based on the structure of niacin, as to what substances would be in the niacin biosynthesis pathway. They could thus make educated guesses as to what substances they might add to the culture medium to enable the mutants to grow. Mutant B (table 5), for example, could grow if given niacin or, alternatively, 3-hydroxyanthranilic acid. It could not grow if given only kynurenine. Thus, Beadle and Tatum knew that the B mutation affected the pathway between kynurenine and 3-hydroxyanthranilic acid. Similarly, mutant A could grow if given 3-hydroxyanthranilic acid or kynurenine instead of niacin. Therefore, these two products must be in the pathway after the step interrupted in mutant A. Conversely, since neither of these mutant organisms could grow when given only tryptophan, Beadle and Tatum knew that tryptophan occurred in the pathway before the steps with the deficient enzymes. By this type of analysis, they discovered the steps in several biochemical pathways of Neurospora. Many biochemical pathways are similar in a huge range of organisms, and thus Beadle and Tatum's work was of general importance.

Table 5: Growth performance of *Neurospora* mutants (plus sign indicates growth; minus sign indicates no growth)

	Additive			
	Tryptophan	Kynurenine	3-Hydroxyanthranilic Acid	Niacin
Wild-type	+	+	+	+
Mutant A	–	+	+	+
Mutant B	–	–	+	+

Beadle and Tatum could further verify their work by observing which substances accumulated in the mutant organisms. If a biochemical pathway is blocked at a certain point, then the substrate at that point cannot convert into the next product, and it builds up in the cell. For example, in the niacin pathway (figure 17), if a block occurs just after 3-hydroxyanthranilic acid, that substance will build up in the cell because it cannot convert into the next substance on the way to niacin.

This analysis could be misleading, however, if the built-up substance is being "siphoned off" into other biochemical pathways in the cell. Also, the cell might attempt to break down or sequester toxic substances. This would mean there might not be an obvious buildup of the substance just before the blocked step. Beadle and Tatum concluded from their studies that one gene controls the production of one enzyme.

The one-gene-one-enzyme hypothesis is an oversimplification that we will clarify later. As a rule of thumb, however, the hypothesis is valid, and it has served to direct attention to the functional relationship between genes and enzymes in biochemical pathways.

Although a change in a single enzyme usually disrupts a single biochemical pathway, it frequently has more than one effect on phenotype. Multiple effects are referred to as pleiotropy. A well-known example is sickle-cell anemia, caused by a mutation in the gene for the β chain of the hemoglobin molecule. In a homozygote, this mutation causes a sickling of red blood cells (figure 18). The sickling of these cells has two major ramifications.

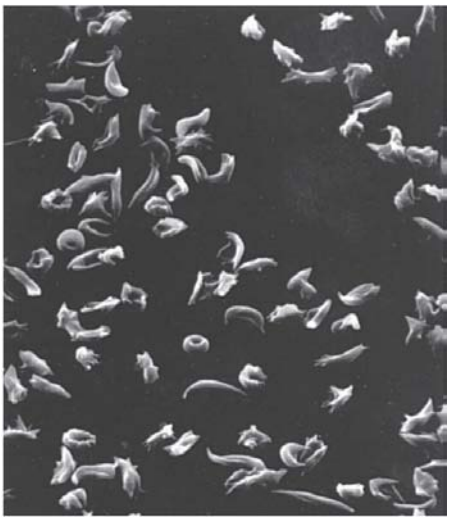


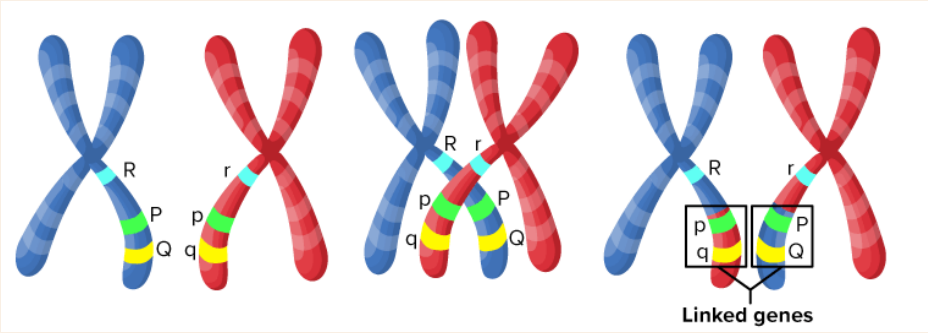
Figure 18: Sickle-shaped red blood cells from a person with sickle-cell anemia. Red blood cells are about 7 to 8 μm in diameter.

First, the liver destroys the sickled cells, causing anemia. The phenotypic effects of this anemia include physical weakness, slow development, and hypertrophy of the bone marrow, resulting in the "tower skull" seen in some of those afflicted with the disease. The second major effect of sickle-cell anemia is that the sickled cells interfere with capillary blood flow, clumping together and resulting in damage to every major organ. The individual can suffer pain, heart failure, rheumatism, and other ill effects. Hence, a single mutation shows itself in many aspects of the phenotype.

CONCEPT OF GENETIC LINKAGE

Genetic linkage is a fundamental concept in genetics that describes the tendency of certain genes to be inherited together because they are located on the same chromosome. It was first observed by British geneticist William Bateson and his colleagues in the early 20th century during their studies of inheritance patterns in pea plants. Genetic linkage occurs due to the physical proximity of genes on a chromosome, which affects their likelihood of being inherited as a unit.

During sexual reproduction, each parent contributes one set of chromosomes to the offspring. The chromosomes that carry the genes responsible for specific traits are transmitted as a whole unit, and the genes located close together on the same chromosome tend to be inherited together more often than not. This phenomenon is called genetic linkage.



The degree of genetic linkage between two genes is measured by the recombination frequency, which is the proportion of offspring that show a recombination (exchange of genetic material) between the two genes compared to the total number of offspring. Genes that are closely linked have a low recombination frequency, while genes that are farther apart on the chromosome have a higher recombination frequency.

Genetic linkage has significant implications for genetics and is a critical tool for gene mapping and understanding the inheritance of traits. By analyzing the patterns of genetic linkage, scientists can determine the relative positions of genes on a chromosome and create genetic maps that show the order and distances between genes.

One of the important consequences of genetic linkage is the phenomenon of genetic inheritance patterns called “linkage disequilibrium.” This occurs when certain combinations of alleles at different linked genes are found together more often than would be expected by chance. Linkage disequilibrium is a valuable tool in population genetics and can provide insights into the history and evolutionary relationships of populations.

Genetic linkage also plays a crucial role in human genetics and the study of inherited diseases. By identifying genes that are linked to specific diseases, researchers can better understand the genetic basis of these conditions and develop targeted therapies and treatments.

SUMMARY

- The process of genetic transmission occurs during sexual reproduction, where each parent contributes half of their genetic material to form the offspring's complete genetic makeup.
- Gregor Mendel, in his experiments, tried to crossbreed plants that had discrete, non-overlapping characteristics and then to observe the distribution of these characteristics over the next several generations. Mendel worked with the common garden pea plant, Pisum sativum.
- The chloroplast DNA (cpDNA) and mitochondrial DNA (mtDNA) in plants are circular DNA molecules just like in prokaryotic genomes, such as bacteria.
- 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.
- The genes causing self-sterility in plants probably produce their effects by controlling the growth rate of the pollen tubes.
- 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.
- Biochemical genetics involves diagnosing and treating metabolic diseases. These are problems with how the body makes, breaks down or uses proteins, fats or carbohydrates. They are caused by genes that are not normal and cannot make the enzymes the body needs.

MULTIPLE CHOICE QUESTIONS

1. Which of the following best expresses the concept of the word 'allele'?

a. Genes for wrinkled and yellow
b. Genes for wrinkled and round
c. phenotypes
d. the expression of agene

2. The inheritance of skin colour in humans is an example of which of the following?

a. Epistasis
b. Gene linkage
c. polygeneic inheritance
d. codominance

3. Which of the following is the most likely explanation for a high rate of crossing-over between two genes?

a. The two genes are far apart on the same chromosome.
b. The two genes are both located near the centromere.
c. The two genes are sex-linked.
d. The two genes code for the same protein.
e. The two genes are on different chromosomes.

4. If in the F1 and F2 generations the same characteristics appeared in both male and females, it would be safe to assume that these traits for eye color and wing length…

a. are sex-linked
b. vary in dominance according to sex
c. are sex-influenced characteristics
d. are autosomal characteristics

5. In the F2 generation, the results are best explained by the fact that…

a. the test cross with the F1 flies resulted in sterile offspring
b. these genes for eye color and wing shape do not pass through the F1 generation
c. these genes for eye color and wing shape are found on the same chromosome
d. crossing over decreases variability
e. the genes are sex-linked

Answer

1. (b) 2. (c) 3. (a) 4. (d) 5. (c)

REVIEW QUESTIONS

1. Write and illustrate the diagram for statistical tests of Mendelian inheritance.
2. What is the rule of segregation and define how to test it?
3. What do you understand by binomial nomenclature?
4. Explain the various characteristics of multiple alleles.
5. How to find inborn errors in metabolism? Explain.
6. What is the importance of genetic transmission?
7. Why did Mendel self-pollinate the tall F1 plants to get the F2 generation and crossed a pure breeding tall plant with a pure breeding dwarf plant to obtain the F1 generation?

REFERENCES

1. Andrews, Lori, Maxwell Mehlman, and Mark Rothstein. Genetics: Ethics, Law and Policy. 4th ed. Eagan, MN: West Academic, 215.
2. Atkinson, Paul, Peter Glasner, and Helen Greenslade, eds. New Genetics, New Identities. New York: Routledge, 2007.
3. Barash, Carol Isaacson. Just Genes: The Ethics of Genetic Technologies. Westport, CT: Praeger, 2007.
4. Berliner, Janice. Ethical Dilemmas in Genetics and Genetic Counseling: Principles through Case Scenarios. Oxford University Press, 2014.
5. Blackford, Russell. Humanity Enhanced: Genetic Choice and the Challenge for Liberal Democracies. Basic Bioethics. Cambridge, MA: MIT Press, 2014.
6. Boesky, Amy, ed. The Story Within: Personal Essays on Genetics and Identity. Baltimore: The Johns Hopkins University Press, 2013.
7. Bunton, Robin and Alan Petersen, eds. Genetic Governance: Health, Risk and Ethics in a Biotech Era. New York: Routledge, 2005.
8. Burley, Justine, and John Harris, eds. A Companion to Genethics. Malden, MA: Blackwell, 2002.
9. Corrigan, Oonagh. Genetic Databases: Socio-Ethical Issues in the Collection and Use of DNA. New York: Routledge, 2004.
10. Deane-Drummond, Celia, Robin Grove-White, and Bronislaw Szerszynski. Reordering Nature: Theology, Society and the New Genetics. New York: T&T Clark, 2003.
11. Deane-Drummond, Celia. Genetics and Christian Ethics. New York: Cambridge University Press, 2006.
12. DeGrazia, David. Creation Ethics: Reproduction, Genetics, and Quality of Life. Reprint ed. Oxford University Press, 2014.
13. Duster, Troy. Backdoor to Eugenics. 2nd ed. New York: Routledge, 2003.
14. Gehring, Verna V., ed. Genetic Prospects: Essays on Biotechnology, Ethics, and Public Policy. Lanham, MD: Rowman & Littlefield, 2004.
15. Hashiloni-Dolev, Yael. A Life (Un)worthy of Living: Reproductive Genetics in Israel and Germany. Dordrecht, Netherlands: Springer, 2007.

CHAPTER 3

POPULATION GENETICS



Population genetics is a subfield of genetics that deals with genetic differences within and between populations, and is a part of evolutionary biology. Studies in this branch of biology examine such phenomena as adaptation, speciation, and population structure.

After studying this chapter, you will understand the core concepts of:

- Factor of Population Genetics
- Problems of Population Genetics

INTRODUCTION

Population genetics is a field of biology that studies the genetic composition of biological populations, and the changes in genetic composition that result from the operation of various factors, including natural selection. Population geneticists develop abstract mathematical models of gene frequency dynamics, extract predictions about the likely patterns of genetic variation in actual populations, and test the predictions against empirical data.

Population genetics is intimately bound up with the study of evolution and natural selection, and is often regarded as the theoretical cornerstone of evolutionary biology. This is because "evolution" has traditionally been defined as any change in a population's genetic composition and natural selection is one important factor that can lead to such a change. Natural selection occurs when some variants in a population out-reproduce other variants as a result of being better adapted to the environment, or 'fitter'. Presuming the fitness differences are at least partly due to genetic differences, this will cause the population's genetic makeup to change, as the genetic variants associated with higher fitness increase in frequency. By devising models of gene frequency change, population geneticists are thus able to shed light on the genetic basis of evolutionary change, and to permit the consequences of different evolutionary hypotheses to be explored in a quantitatively precise way.

The foundation of population genetics lies in the principles of genetics and evolution. Genes are the units of heredity that encode specific traits and characteristics within an organism's DNA. In sexually reproducing populations, genes are passed from one generation to the next through the process of genetic transmission. Over time, the genetic makeup of a population can change due to various factors, such as mutation, natural selection, genetic drift, and gene flow.

Mutation, the ultimate source of genetic variation, introduces new genetic variants into a population. Mutations are changes in the DNA sequence of genes, and some mutations may result in new traits that provide a selective advantage or disadvantage to an individual.

Natural selection is a key driving force in population genetics, determining which genetic variants become more or less prevalent within a population. Individuals with advantageous traits that enhance survival and reproduction are more likely to pass on their genes to the next generation, increasing the frequency of those beneficial alleles in the population. Conversely, individuals with less favorable traits may have reduced reproductive success, leading to a decrease in the frequency of those alleles.

Genetic drift is a random process that can significantly impact small populations. It occurs when genetic variation is influenced by chance events rather than natural selection. Genetic drift is particularly influential in small populations, where chance fluctuations can lead to the fixation or loss of certain alleles over time.

Gene flow, also known as migration, is the movement of genes between populations due to the migration of individuals. Gene flow can introduce new genetic variants

into a population, increasing genetic diversity and potentially affecting the genetic makeup of both the source and recipient populations.

Population genetics provides various methods and tools for studying genetic variation and understanding the evolutionary history of species. It uses mathematical models and statistical analyses to quantify genetic diversity, assess genetic relatedness among populations, and estimate important parameters such as mutation rates and effective population size.

The study of population genetics has profound implications for diverse fields, including conservation biology, medicine, agriculture, and evolutionary biology. In conservation biology, population genetics is used to assess the genetic health of endangered species and design effective conservation strategies to preserve genetic diversity. In medicine, population genetics investigates the genetic basis of diseases and the role of genetic factors in drug responses and treatment outcomes.

Furthermore, population genetics has helped illuminate our understanding of human evolution and migration patterns. By analyzing the genetic variation among human populations, scientists have gained insights into the migratory history and interactions of ancient human populations.

Population genetics is a fascinating and essential field of biology that explores the genetic variation and dynamics within populations. It serves as a powerful tool for understanding the evolutionary processes and mechanisms that shape the genetic makeup of species over time. The study of population genetics has profound implications across various disciplines, contributing to our knowledge of genetics, evolution, and the complex interactions within living organisms and their environments.



CORE CONCEPT: FACTOR OF POPULATION GENETICS

Population genetics is the study of the genetic composition of populations, including distributions and changes in genotype and phenotype frequency in response to the processes of natural selection, genetic drift, mutation and gene flow. Population genetics was a vital ingredient in the emergence of the modern evolutionary synthesis.

Population genetics is the study of genetic variation within populations, and involves the examination and modelling of changes in the frequencies of genes and alleles in populations over space and time. Many of the genes found within a population will be polymorphic - that is, they will occur in a number of different forms (or

alleles). Mathematical models are used to investigate and predict the occurrence of specific alleles or combinations of alleles in populations, based on developments in the molecular understanding of genetics, Mendel's laws of inheritance and modern evolutionary theory. The focus is the population or the species - not the individual. The collection of all the alleles of all of the genes found within a freely interbreeding population is known as the gene pool of the population. Each member of the population receives its alleles from other members of the gene pool (its parents) and passes them on to other members of the gene pool (its offspring). Population genetics is the study of the variation in alleles and genotypes within the gene pool, and how this variation changes from one generation to the next.

Factors influencing the genetic diversity within a gene pool include population size, mutation, genetic drift, natural selection, environmental diversity, migration and non-random mating patterns. The Hardy-Weinberg model describes and predicts a balanced equilibrium in the frequencies of alleles and genotypes within a freely interbreeding population, assuming a large population size, no mutation, no genetic drift, no natural selection, no gene flow between populations, and random mating patterns.

In natural populations, however, the genetic composition of a population's gene pool may change over time. Mutation is the primary source of new alleles in a gene pool, but the other factors act to increase or decrease the occurrence of alleles. Genetic drift occurs as the result of random fluctuations in the transfer of alleles from one generation to the next, especially in small populations formed, say, as the result adverse environmental conditions (the bottleneck effect) or the geographical separation of a subset of the population (the founder effect). The result of genetic drift tends to be a reduction in the variation within the population, and an increase in the divergence between populations. In many cases, the effects of natural selection on a given allele are directional. The allele either confers a selective advantage, or spreads throughout the gene pool, or it confers a selective disadvantage, and disappears from it. In other cases, however, selection acts to preserve multiple alleles within the gene pool and a balanced equilibrium is observed. This situation, labelled balanced polymorphism, can arise because of a selective advantage for individuals heterozygous for a given allele. For example, the disease sickle cell anaemia is caused by a mutation in one of the genes responsible for the production of haemoglobin. Individuals with two copies of the mutant gene for sickle haemoglobin (HbS/HbS) develop the disease. Individuals that are heterozygous - one copy of the sickle gene and one copy of the normal gene (HbS/HbA) - are carriers of the condition. It is believed that these heterozygous individuals are more resistant to malaria than individuals homozygous for the normal gene (HbA/HbA), and that this selective advantage maintains the presence of the HbS gene in the population. As a result of balanced polymorphism, the gene pools of most populations contain a number of deleterious alleles that reduces the overall fitness of the population (known as the genetic load).

Genetic variation within populations and species can now be analysed at the level of nucleotide sequences in DNA (genome analysis) and the amino acid sequences of

proteins (proteome analysis). The genetic differences between species can be used to infer evolutionary history, on the basis that the closest relatives will have gene pools that are most similar. Recent advances in the sequencing of genomes, allied to computer-based techniques for storing and comparing this information, have led to the construction of detailed evolutionary trees. The use of molecular clocks - nucleotide sequences (or amino acid sequences) in which evolutionary change accumulates at a constant rate - allows dates to be attached to the points at which populations start to diverge to form new species. These approaches are also proving useful in other useful areas (for example, in tracing the transmission routes of infectious diseases).

If two populations of a given species become genetically distinct enough that they can no longer interbreed, they are regarded as new species (a process called speciation).

Darwinian Population Genetics

Darwinian population genetics, or simply population genetics, is a central feature, if not the central feature, of the modern evolutionary synthesis, or neo-Darwinism. The modern evolutionary synthesis is a combination of Darwin's theory of the evolution of species by natural selection, Mendel's theory of genetics as the basis for biological inheritance, and mathematical population genetics. Put together by dozens of scientists throughout the 1930s and 1940s, Darwinian population genetics is our best model of the process that incrementally created all life on Earth, evolution and natural selection. Population genetics is the study of the genetic distribution and change of allele frequency within a given species -- basically, which genes are more or less prominent within that species. This distribution and the way it changes can be understood comprehensively through five forces: natural selection, genetic drift, mutation, migration and nonrandom mating. Mathematical population genetics is a formalism that allows us to precisely predict future allele frequencies if complete information about these five forces is available. Of course, it never is, although approximations can be exceedingly useful. Population genetics is among the best-verified models in science.

Population genetics can explain the process by which new adaptations emerge and become fixed in a species, how speciation occurs, why some adaptive traits evolve more readily than others, the evolution of evolvability and many other matters of scientific interest. Because evolution mostly takes place over millions of years (i.e., evolutionary timescales), direct experimentation can be difficult. However, the principles of population genetics have been tested and verified with species whose replicative times are very short, such as bacteria, which can reproduce in half an hour given suitable space and nutrients.

The five evolutionary forces quantified and studied by population genetics are intuitively easy to understand. Natural selection occurs when an organism is killed off by its environment, rival species, or members of the same species, for whatever reason. If death occurs before reproduction, the organism is said to be evolutionarily

unsuccessful. Whatever traits are associated with the premature death will eventually be selected out of the gene pool. Natural selection is perhaps the most powerful of the evolutionary forces, and the most widely understood.

Genetic drift occurs randomly in offspring. When a certain trait doesn't effect an organism's ability to survive or reproduce one way or the other, it may evolve and become fixed purely as a consequence of the genetic lottery. Mutations are also a side effect of the copying errors of DNA, and rarely become visible or significant, although occasionally they may result in an advantageous survival trait.

Migration occurs when members of a species move from one place to another, cutting off reproductive ties with other members of the species. When a part of a species becomes reproductively isolated from another part by a geographic barrier, the two groups eventually speciate into different variants. This was famously observed on the Galapagos Islands by Charles Darwin. Nonrandom mating is another very powerful force in population genetics. More attractive members of a species generally acquire more mates for longer periods, and experience greater reproductive success as a result.

Population genetics is a huge field which has been studied by many millions of scientists, not only biologists, and will continue to be studied by millions to come. Unfortunately, it is only taught at the most basic level in the public school systems.

Natural Selection

Natural selection can be viewed either narrowly or broadly. Narrowly conceived, it is simply one class of violations of the assumptions of the Hardy-Weinberg Laws, namely the cases in which viability or fertility depends on genotype. Broadly conceived, it is the primary force which causes evolution to be adaptive, the creative and progressive element in the evolutionary process. A comprehensive theory of evolution, one which does not yet exist, would integrate ecological processes (which determine the range of environments and the finesses of phenotypes), developmental processes (which determine the effect of genotype on phenotype), and population genetics (which tells us the changes in genetic composition of a population when the finesses of the genotypes are known). Lacking the other elements of this future theory, we concentrate here on the population genetics. We first examine the mathematics of gene frequency change, to see how much change of gene frequency is caused by a given pattern of differential viability and fertility, to see what the pattern of gene frequency change through time will be, and to see how selection acting on diploid genotypes affects gene frequencies. The selection with constant finesses causes an equilibrium genetic composition to be maintained. The third part concerns the effect of natural selection on the mean fitness of individuals in the population. It was a central tenet of Darwin's thinking that natural selection had an average tendency to increase adaptedness, but it is a nontrivial matter to investigate whether or not it actually does so in simple model situations. All of this development takes place under the assumption that finesses are constants which do not change either as a function of time or of the composition of the population.

Simply put, selection occurs when some genotypes in the population are on average more successful in reproduction. These genotypes may survive better, produce more offspring, or be more successful in attracting mates; the alleles responsible for these traits are then passed on to offspring. There is broad theoretical consensus and abundant empirical data to suggest that selection can change populations radically and quickly. If one genetic variant, or allele, increases survivorship or fertility, selection will increase the frequency of the favored allele, and concurrently eliminate other alleles. This type of selection, called directional selection, decreases the amount of genetic variation in populations.

Alternatively, an individual carrying two different alleles for the same gene (a heterozygote) may have advantages, as exemplified by the well-known example of the sickle-cell allele in Africa, in which **heterozygotes** are more resistant to malaria. In this case, called over dominant selection, genetic variation is preserved in the population. Although a number of similar examples are known, directional selection is much more common than over dominant selection; this implies that the common action of selection is to decrease genetic variation within populations. It is equally clear that if different (initially similar) populations occupy different habitats, selection can create differences among populations by favoring different alleles in different areas.

Heterozygote is an individual having two different alleles of a particular gene or genes, and so giving rise to varying offspring.

Gene Doping

Gene doping, use of substances or techniques to manipulate cells or genes in order to improve athletic performance. Since the latter half of the 20th century, the manipulation of human genes has formed an important area of biomedical research, with much effort focused in particular on refining gene therapy for the treatment of diseases such as cystic fibrosis and anemia. In the early 21st century, however, members of the international sports community became concerned that athletes seeking to gain physical advantage in competition would abuse gene therapy and similar technologies. In 2003, although no athletes were known to have experimented with gene doping, the World Anti-Doping Agency, which regulates the use of substances in sports, added the transfer of cells, DNA, or RNA and the use of all other gene-altering agents, biological or pharmacological, to its list of prohibited substances and methods. One example of a gene therapy with potential for doping in sports is Repoxygen,

which was originally developed for the treatment of anemia but was never tested on humans. Repoxygen consists of a segment of DNA designed to stimulate the synthesis of erythropoietin, a hormone normally produced and released by the kidneys that acts on bone marrow to augment the production of red blood cells (erythrocytes). The DNA is packaged in a viral delivery system, which enables the foreign DNA to be transferred to host cells, where it becomes integrated into the cells' own DNA. The gene insert then instructs cells to synthesize erythropoietin. Increased numbers of red cells increase the oxygen-carrying capacity of blood, which in turn improves aerobic capacity.

In addition to the erythropoietin gene, other genes have also been identified as possible targets for manipulation, including those that encode insulin-like growth factor 1, human growth hormone, phosphoenolpyruvate carboxykinase, and peroxisome proliferator-activated receptor-delta. Whereas gene therapies such as Repoxygen presumably can be detected through DNA sequencing of suspected abusers, other substances and techniques used to manipulate genes may not be easily distinguished from substances that occur naturally in the body.

Genetic Drift

Often overlooked by the layperson, genetic drift is given a place of importance in population genetics. While some analyses of genetic drift quickly become complicated, the basic process of drift is simple and involves random changes in allele frequency. In sexual species, the frequency of alleles contained in the progeny may not perfectly match the frequency of the alleles contained in the parents. As an analogy, consider flipping a coin twenty times. Although one might expect ten heads and ten tails, the actual outcome may be slightly different; in this example, the outcome (progeny) does not perfectly represent the relative frequency of heads and tails (the parents).

What does this mean for populations? Start by considering neutral alleles, which have no impact on survival or reproduction. (An example is the presence or absence of a widow's peak hairline.) The frequency of a neutral allele may shift slightly between generations, sometimes increasing and sometimes decreasing. What outcomes are expected from this process? Suppose that a particular allele shifts frequency at random for a number of generations, eventually becoming very rare, with perhaps only one copy in the population. If the individual carrying this allele does not pass it on to any offspring or fails to have any offspring, the allele will be lost to the population. Once lost, the allele is gone from the population forever. In this light, drift causes the loss of genetic variation over time. All populations are subject to this process, with smaller populations more strongly affected than larger ones.

Perhaps better known than the pervasive, general effects of genetic drift are special examples of drift associated with unusually small populations. Genetic bottlenecks occur when a small number of individuals from a much larger population are the sole contributors to future generations; this occurs when a catastrophe kills most of the population, or when a few individuals start a new population in different area. Genetic bottlenecks reduce the genetic variation in the new or subsequent population relative

to the old. Cheetahs, which have very little genetic variation, are presumed to have gone through several genetic bottlenecks. Occasionally, these new populations may have particular alleles that are much more common than in the original population, by chance alone. This is usually called the founder effect.

Genetic drift is a mechanism of evolution that occurs by random chance rather than natural selection. In genetic drift, a population experiences a change in the frequency of a given allele, prompted by random luck rather than a need for adaptation. This differs from natural selection, in which allelic frequency is altered based on the fittest genes surviving to reproduce and the weaker genes dying off. Genetic drift tends to be a phenomenon amongst smaller populations, while natural selection holds sway in larger populations. An allele, or genetic variant, is a component of a gene that produces a certain trait. Imagine there are both red worms and white worms in the same population. If a red worm mates with a white worm, each will pass one allele to its offspring, red or white, to form a gene. The dominant, or strong, allele will determine which trait the baby worm presents. If white is dominant, the baby worm will be white, if red is dominant, the baby worm will be red, and if the baby worm receives two of the same recessive alleles, it will exhibit that recessive feature. Genetics is far more complex than this example allows, but this is the general concept.

Now imagine these worms live in a swamp filled with red mud and are surrounded by birds that want to eat them. The red worms are more likely to survive because they are camouflaged by the mud and won't be as easily seen by predators. Therefore, more red worms will live to reproduce and more red alleles will be passed on to offspring, increasing the red allelic frequency. More white worms, which are easily seen by birds, will be eaten before they are able to pass on their genes, thus decreasing their allelic frequency. This is natural selection.

Now, imagine there are ten red worms and ten white worms with equal chances of surviving to reproduce. A tree falls on the swamp, killing eight worms; six white and two red. Then suppose two white worms and one red worm get sick and die. By chance, there are now seven red worms and only two white worms left. This is an example of genetic drift.

Genetic drift can also occur through a random sampling error. A sampling error occurs when a sample exhibits different results than the entire population would. For example, say there are fifty red worms and fifty white worms in a population, and scientists randomly select ten worms to observe. Because the sample is smaller, the alleles passed on in the group of ten may not even out as they would in a group of one hundred. Also, if the group contains more red worms than white, the presentation of alleles in the offspring will be skewed.

Genetic drift becomes fixed when one allele replaces another entirely or one allele dies off. Imagine the seven red worms and two white worms left in the swamp after the tree catastrophe and sickness killed the other eleven worms. As the worms reproduce, less white worms will appear until finally there are no white worms left. Genetic drift will then be fixed, because all future generations will be red. Because genetic drift works much more rapidly in small populations, a population bottleneck

or founder effect can increase the process of genetic drift. A population bottleneck occurs when a population suddenly undergoes a dip in size. The tree falling on the swamp and killing nearly half the worm population is an example of the bottleneck effect. A founder effect occurs when a small portion of a population becomes isolated from the rest of the group and evolves separately.

Effect of Genetic Diversity

Genetic diversity is the variation of heritable characteristics present in a population of the same species. It serves an important role in evolution by allowing a species to adapt to a new environment and to fight off parasites. It is applicable to domesticated species, which typically have low levels of diversity. Studying genetic diversity in humans can help researchers form theories on human origins.

Living things contain in their cells the basic instructions, or blueprints, for their own development. Many of these instructions, which are called genes, result in physical characteristics that affect the way organisms interact with their environment. Variations in such characteristics within the same species give rise to genetic diversity. For a species to adapt to an ever-changing **ecosystem**, a significant level of variation must be present. Those individuals that possess favorable characteristics will go on to reproduce, while those that don't do not tend to pass on their characteristics to many offspring.

Ecosystem includes all of the living things (plants, animals and organisms) in a given area, interacting with each other, and also with their non-living environments (weather, earth, sun, soil, climate, and atmosphere).

Domesticated species often have low levels of genetic diversity. This is caused by the artificial selection, or preferential breeding, of crops and animals for traits that humans find preferable. While this can have positive short-term results, such as a richer harvest, low diversity among domesticated species poses risks. A newly-evolved virus or bacteria strand can invade a population of nearly identical organisms very rapidly. The protection that diversity generally offers in wild populations is lost in this scenario.

The Irish Potato Famine between 1845 and 1852 was caused by a parasite invading a large population of nearly identical potatoes. The parasite was a water mold named Phytophthora infestans. This famine caused Ireland's population, which was widely dependent on potatoes for food, to decrease by 20 to 25 percent. Human genetic diversity typically varies based on the geographic location of a population. This has led biologists and anthropologists to study these diversity levels in the quest to

understand human origins. Genetic diversity levels in Africa, for example, have been found to be higher than in many of areas of the world. Researchers have developed models of human origins from this evidence. The recent Out-of-Africa model, which suggests modern humans have a common origin in Africa, is one such example.

Biodiversity refers to the level of variation of all living things within an ecosystem. The importance of biodiversity to an ecosystem is analogous to the importance of genetic diversity to a population. Both forms of diversity contribute to the health and robustness of the larger system. When these levels of diversity fall, both systems are less able to adapt to a changing environment.

Gene Pool

The gene pool is a reflection of the total number of alleles shared among the breeding population of a species. An allele is one form of a gene: for example, the gene which codes for eye color in humans include alleles for blue, green, brown, and hazel eyes, among others. People can refer to the gene pool for a specific gene, such as hair color, or for a species as a whole.

As a general rule, the bigger the gene pool is, the better things are for a species, because it reflects diversity. Diversity tends to encourage hardiness. When a population is not very diverse, negative alleles can become amplified, which can be dangerous for the population in the long term. For example, in a species of animal in which many animals have a allele which weakens the bones in the legs, future generations will be more likely to have weak legs, while in a diverse population where only a few individuals have this allele, fewer numbers of animals will have this problem in the future because there are fewer individuals to pass it on.

A number of things can influence the size of a gene pool. Overall number of organisms is important, as are the relationships between organisms. The more total representatives of a species, the more possibility for diversity, and the fewer direct genetic relationships between animals, the less of a potential for negative alleles to become amplified as a result of inbreeding. Spontaneous mutations in a population can also increase the size of a gene pool, with some species being famous for their high mutation rate.

Gene pools are of interest for a number of reasons. Among livestock farmers and animal breeders, a diverse set of genes is desired to create a healthy population of breeding animals. This is an especially large concern for breeders who work with purebred animals, as purebreds tend to develop issues associated with inbreeding due to restrictions on the animals allowed to breed. Conservationists are also concerned with gene selection, as endangered species can reach a point of critical mass beyond which they will not survive as their populations dwindle and the gene pool becomes less diverse.

Researchers interested in the growth and shrinkage of gene pools often work with fruit flies, because they breed quickly and they have long been used in the study of genetics. Working with carefully controlled populations of fruit flies, geneticists and

other researchers can examine the impact of a variety of influences on the genetic health of a species.

Method of Population Genetics

Because the processes of variation and selection take place at the population level, the basic theory of the genetics of evolutionary change is contained in the general area known as population genetics.

A simple way of viewing evolutionary change at the genotypic level would be to invent some hypothetical ancestral genotype, such as AAbbccDDEE, and an "evolved" derivative, such as aaBBccddee. (For illustrative purposes, only five genes are used, and these are assumed to be all homozygous.) Also, for simplicity it can be assumed that in both the ancestral and the evolved populations all individuals are identical. Clearly for all the genes except cc, a new allele completely replaces the original allele, and the new alleles can be either dominant or recessive. For example, in the case of the first gene, in the ancestral population all alleles are A, and in the evolved population all are a.

For a to replace A, the population must go through stages in which there are mixtures of A and a alleles present in the population at the same time. In population genetics, allele frequency is the measurement of the commonness of an allele. The convention is to let the frequency of a dominant allele be p and that of a recessive allele q. Both are generally expressed as decimal fractions. In the above example, p changes from 1 to 0, and q changes from 0 to 1. Since there are only two alleles in this example, p + q must always equal 1. In the intermediate stages, there must be times when there are intermediate allele frequencies, for example when p = 0.4 and q = 0.6.

What can be said about genotype frequencies in the intermediate populations? In the ancestral and derived populations there must have been the following genotypic frequencies:

Ancestral AA = 1, Aa = 0, aa = 0

Evolved AA = 0, Aa = 0, aa = 1

Intuitively it seems that, in the intermediate stages, there must be more-complex proportions, including some heterozygotes. One possible intermediate stage is as follows:

AA = 0.30, Aa = 0.20, aa = 0.50

The allele frequencies at such an intermediate stage can be calculated by "adding up" the alleles. Hence, the frequency of A will be 0.30 plus $^{1}/_{2}$ of 0.20 because the heterozygotes only carry one A allele. This is written

$$p = 0.30 + {}^{0.20}/_{2} = 0.40$$

Similarly, $q = 0.50 + {}^{0.20}/_{2} = 0.60$

In general, if D = frequency of homozygous dominants, R = frequency of homozygous recessives, and H = frequency of heterozygotes, then

$$p = D + H/_2$$

and

$$q = R + H/_2$$

This section has shown the importance of the concepts of allele frequency and genotype frequency in describing the genetic structure of populations. Of these, allele frequency is the simpler descriptor, and it forms the central tool of population genetics. Hence, the genetic basis of evolutionary change at the population level is described in terms of changes of allele frequencies.

Hardy-Weinberg equilibrium

It is a curious fact that populations show no inherent tendency to change allele or genotype frequencies. In the absence of selection or any of the other forces that can drive evolution, a population with given values of p and q will settle into a special stable set of genotypic proportions called a Hardy-Weinberg equilibrium. This principle was first realized by Godfrey Harold Hardy and Wilhelm Weinberg in 1908. The Hardy-Weinberg equilibrium of a population with allele frequencies p and q is defined by the set of genotypic frequencies p^2 of AA, 2pq of Aa, and q^2 of aa.

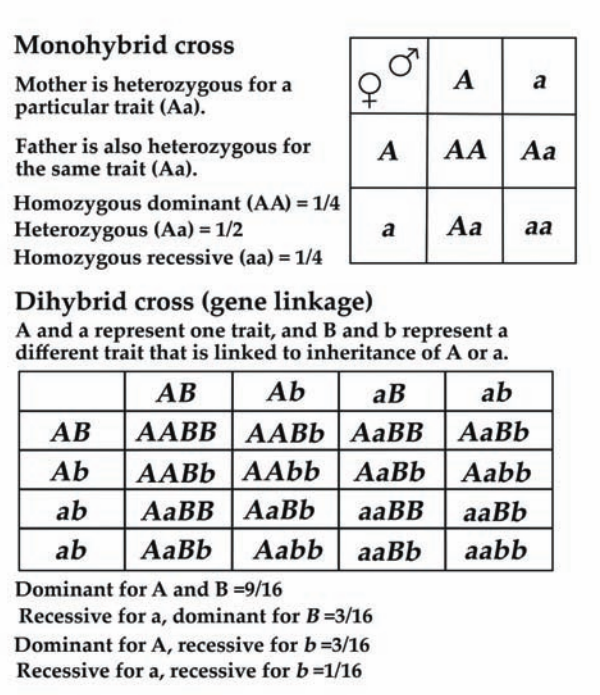


Figure 1. Punnett square diagrams are used to predict all the possible gene combinations that could result from the mating of parents with known genotypes for a particular trait. A monohybrid cross represents the inheritance pattern of a single trait, whereas a dihybrid cross represents the inheritance patterns of two traits that are linked.

When such a population reproduces itself to make a new generation, the lack of change is made apparent. It is intuitive that the allele frequencies p and q in the population are also measures of the frequencies of eggs and sperm used in creating a new generation. The new generation produced from the zygotes has exactly the same genotypic proportions as the first generation (the parents of the zygote).

Some specific allele frequencies, 0.7 for p and 0.3 for q, can be used to illustrate the calculation of the genotypic frequencies that constitute the Hardy-Weinberg equilibrium:

$p \times p = 0.7 \times 0.7 = 0.49$ of AA

$2 \times p \times q = 2 \times 0.7 \times 0.3 = 0.42$ of Aa

$q \times q = 0.3 \times 0.3 = 0.09$ of aa

When this population reproduces, there will be 0.49 + 0.21 = 0.7 of A gametes and 0.09 + 0.21 = 0.3 of a gametes, and, when these gametes combine, the population in the next generation will clearly have the same genotypic proportions as the previous one.

These simple calculations rely on several underlying assumptions. Perhaps the most crucial one is that there is random mating, or mating regardless of the genotype of the partner. In addition, the population must be large, and there can be no other pressures, such as selection, that can change allele frequencies. Despite these stringent requirements, many natural populations that have been studied are in **Hardy-Weinberg equilibrium** for the genes under investigation. The Hardy-Weinberg equilibrium constitutes an important benchmark for population genetic analysis.

Hardy–Weinberg equilibrium states that allele and genotype frequencies in a population will remain constant from generation to generation in the absence of other evolutionary influences.

If the Hardy-Weinberg principle of population genetics shows that there is no inherent tendency for evolutionary change, then how does change occur? This is considered in the following sections.

Changes in gene frequencies

Selection

One assumption behind the calculation of unchanging genotypic frequencies in Hardy-Weinberg equilibrium is that all genotypes have the same fitness. In genetics, fitness does not necessarily have to do with muscles; fitness is a measure of the ability to produce fertile offspring. In reality, the fitnesses of different genotypes are highly variable. The genotype with the greatest fitness is given the fitness value (w) of 1, and the lesser

fitnesses are fractions of 1. For example, if snails of genotypes *AA* and *Aa* were to have an average of 100 offspring but those of genotype *aa* only 70, then the fitnesses of these three genotypes would be 1, 1, and 0.7, respectively. The proportional difference from the most fit is called the selection coefficient, *s*. Hence, $s = 1 - w$.

Alleles carried by less-fit individuals will be gradually lost from the population, and the relevant allele frequency will decline. This is the fundamental way in which natural selection operates in a population. Selection against dominant alleles is relatively efficient, because these are by definition expressed in the phenotype. Selection against recessive alleles is less efficient, because these alleles are sheltered in heterozygotes. Even though populations under selection technically are not in Hardy-Weinberg equilibrium, the proportions of the formula can be used as an approximation to show the relative proportions of homozygous recessives and heterozygotes. If a rare deleterious recessive allele is of frequency $^1/_{50}$ in the population, then $(^1/_{50})^2$, or 1 out of 2,500, individuals will express the recessive phenotype and be a candidate for negative selection. Heterozygotes will be at a frequency of $2pq = 2 \times {}^{49}/_{50} \times {}^1/_{50}$, or about 1 in 25. In other words, the heterozygotes are 100 times more common than recessive homozygotes; hence, most of the recessive alleles in a population will escape selection.

Because of the sheltering effect of heterozygotes, selection against recessive phenotypes changes the frequency of the recessive allele slowly. Even if the most severe level of selection is imposed, giving the recessive phenotype a fitness of zero (no fertile offspring), the recessive allele frequency (expressed as a fraction of the form $^1/_x$) will increase in denominator by 1 in every generation. Therefore, to halve an allele frequency from $^1/_{50}$ to $^1/_{100}$ would proceed slowly from $^1/_{50}$ to $^1/_{51}$, $^1/_{52}$, $^1/_{53}$, and so on and would take 50 generations to get to $^1/_{100}$. For lower intensities of selection, the progress would be even slower.

A different type of natural selection occurs when the fitness of a heterozygote exceeds the fitness of both homozygotes. The maintenance in human populations of the severe hereditary disease sickle cell anemia is owing to this form of selection. The disease allele (HbS) produces a specific type of hemoglobin that causes distortion (sickling) of the red blood cells in which the hemoglobin is carried. (Normal hemoglobin is coded by another allele, HbA). Accordingly, the possible genotypes are HbAHbA, HbAHbS, and HbSHbS. The latter individuals are homozygous for the sickle cell allele and will develop severe anemia because the oxygen transporting property of their blood is compromised. While the condition is not lethal before birth, such individuals rarely survive long enough to reproduce. On these grounds it might be expected that the disease allele would be selected against, driving the allele frequency to very low levels. However, in tropical areas of the world, the allele and the disease are common. The explanation is that the HbAHbSheterozygote is fitter and capable of leaving more offspring than is the homozygous normal HbAHbA in an environment containing the falciparum form of malaria. This extra measure of protection is evidently provided by the sickle cell hemoglobin, which is detrimental to the malaria parasite. In malarial environments, therefore, populations that contain

the sickle cell gene have advantages over populations free of this gene. The former populations are in less danger from malaria, although they "pay" for this advantage by sacrificing in every generation some individuals who die of anemia.

Mutation

Genetics has shown that mutation is the ultimate source of all hereditary variation. At the level of a single gene whose normal functional allele is A, it is known that mutation can change it to a nonfunctional recessive form, a. Such "forward mutation" is more frequent than "back mutation" (reversion), which converts a into A. Molecular analysis of specific examples of mutant recessive alleles has shown that they are generally a heterogeneousset of small structural changes in the DNA, located throughout the segment of DNA that constitutes that gene. Hence, in an example from medical genetics, the disease phenylketonuria is inherited as a recessive phenotype and is ascribed to a causative allele that generally can be called k. However, sequencing alleles of many independent cases of phenylketonuria has shown that this k allele is in fact a set of many different kinds of mutational changes, which can be in any of the protein-coding regions of that gene.

Recessive deleterious mutations are relatively rare, generally in the order of 1 per 105 or 106 mutant gametes per generation. Their constant occurrence over the generations, combined with the even greater rarity of back mutations, leads to a gradual accumulation in the population. This accumulation process is called mutational pressure.

Since mutational pressure to a deleterious recessive allele and selection pressure against the homozygous recessives are forces that act in opposite directions, another type of equilibrium is attained that effectively sets the value of q. Mathematically, q is determined by the following expression in which u is the net mutation rate of A to a, and s is the selection coefficient presented above:

$$q^2 = (u/s), \text{ or } q = (u/s)^{1/2}$$

Nonrandom mating

Many species engage in alternatives to random mating as normal parts of their cycle of sexual reproduction. An important exception is sexual selection, in which an individual chooses a mate on the basis of some aspect of the mate's phenotype. The selection can be based on some display feature such as bright feathers, or it may be a simple preference for a phenotype identical to the individual's own (positive assortative mating).

Two other important exceptions are inbreeding (mating with relatives) and enforced outbreeding. Both can shift the equilibrium proportions expected under Hardy-Weinberg calculations. For example, inbreeding increases the proportions of homozygotes, and the most extreme form of inbreeding, self-fertilization, eventually eliminates all heterozygotes. Inbreeding and outbreeding are evolutionary strategies adopted by plants and animals living under certain conditions. Outbreeding brings

gametes of different genotypes together, and the resulting individual differs from the parents. Increased levels of variation provide more evolutionary flexibility. All the showy colors and shapes of flowers are to promote this kind of exchange. In contrast, inbreeding maintains uniform **genotypes**, a strategy successful in stable ecological habitats.

Genotype is the part of the genetic makeup of a cell, and therefore of an organism or individual, which determines one of its characteristics (phenotype).

In humans, various degrees of inbreeding have been practiced in different cultures. In most cultures today, matings of first cousins are the maximal form of inbreeding condoned by society. Apart from ethical considerations, a negative outcome of inbreeding is that it increases the likelihood of homozygosity of deleterious recessive alleles originating from common ancestors, called homozygosity by descent. The inbreeding coefficient F is a measure of the likelihood of homozygosity by descent; for example, in first-cousin marriages, $F = {}^{1}/_{16}$. A large proportion of recessive hereditary diseases can be traced to first-cousin marriages and other types of inbreeding.

Random genetic drift

In populations of finite size, the genetic structure of a new generation is not necessarily that of the previous one. The explanation lies in a sampling effect, based on the fact that a subsample from any large set is not always representative of the larger set. The gametes that form any generation can be thought of as a sample of the alleles from the parental one. By chance the sample might not be random; it could be skewed in either direction. For example, if p = 0.600 and q = 0.400, sampling "error" might result in the gametes having a p value of 0.601 and a q of 0.399. If by chance this skewed sampling occurs in the same direction from generation to generation, the allele frequency can change radically. This process is known as random genetic drift. As might be expected, the smaller the population, the greater chance of sampling error and hence significant levels of drift in any one generation. In extreme cases, drift over the generations can result in the complete loss of one allele; in these occurrences the other is said to be fixed.

Other cases of sampling error occur when new colonies of plants or animals are founded by small numbers of migrants (founder effect) and when there is radical reduction in population size because of a natural catastrophe (population bottleneck). One inevitable effect of these processes is a reduction in the amount of variation in the population after the size reduction. Two species that have gone through drastic

bottlenecks with the associated reduction of genetic variation are cheetahs (Africa) and northern elephant seals (North America).



CORE CONCEPT: PROBLEMS OF POPULATION GENETICS

Population geneticists pursue their goals by developing abstract mathematical models of gene frequency dynamics, trying to extract conclusions from those models about the likely patterns of genetic variation in actual populations, and testing the conclusions against empirical data.

The science of population genetics is the auto mechanics of evolutionary biology. Organic evolution, in the Darwinian scheme, is a consequence of the conversion of variation among members of an ensemble into differences between ensembles in time and space. Classically those ensembles are collections of individual organisms, populations in the usual sense, but the Darwinian scheme can be applied as well to ensembles of organelles within cells or to collections of populations that make up a species. The essential features of the Darwinian scheme that determine both the subject and form of population genetics are that

- There are processes that produce variation among individuals within a population,
- There are processes that result in changes in the relative frequencies of the variants within a generation,
- There is a hereditary process across generations that may result in further change in the relative frequencies of the different variants, but population frequencies are correlated across generational lines so that the frequency distribution of variants in any time interval is some nontrivial function of their distribution in the interval of time.

It is the relation between the processes of the generation and the modulation of variation within generations in features 1 and 2 and the processes occurring between generations in feature 3 that both create the science of population genetics and pose its methodological dilemmas and its shape as an inquiry. These relations are most easily understood as the transformations in the state of the population, shown in Figure 2. The population change can be represented both in a genotypic and a phenotypic space, and the complete laws of transformation make use of both these spaces.

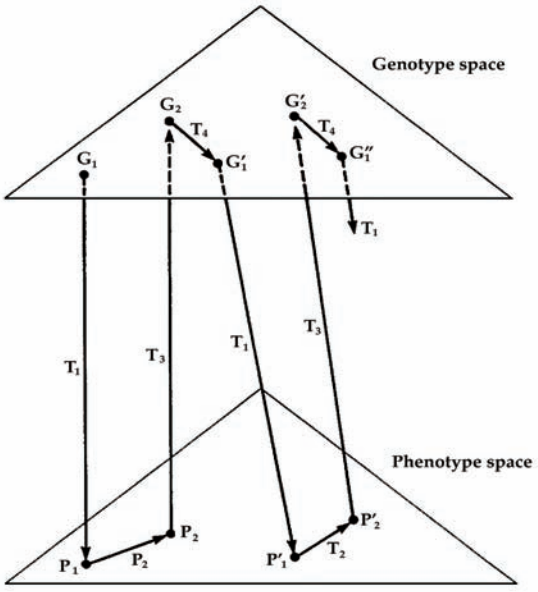


Figure 2 Schematic representation of the paths of transformation of population genotype from one generation to the next. G and P are the spaces of genotypic and phenotypic description. G_1, G'_1, G_2, and G'_2 are genotypic descriptions at various points in time within successive generations. P_1, P'_1, P_2, and P'_2 are phenotypic descriptions. T_1, T_2, T_3, and T_4 are laws of transformation. Details are given in the text.

G_1, G_2, P_1, and P_2 are genotypic and phenotypic descriptions at different points within a generation, while the same quantities with primes are the descriptions of the equivalent states in the next generation. The laws of transformation are

T_1: a set of epigenetic laws that gives the distribution of phenotypes, including fitness, that results from the development of various genotypes in various environments.

T_2: the laws of mating, migration, and natural selection including stochastic elements that transform the phenotypic array of potential reproducing units in a population within a generation.

T_3: the set of reverse epigenetic transformations that allows inferences about the distribution of genotypes G_2 corresponding to the distribution of phenotypes P_2.

T_4: the phenomenology of genetics such as segregation, recombination, mutation, horizontal transfer, etc., including stochastic elements, that allows us to predict the probability distribution of genotypes in the next generation produced from gametogenesis and fertilization, given an array of parental genotypes.

This representation of the structure of population genetic inference points immediately to two major sources of difficulty that have shaped the problems of population from its beginnings and still do so. First there is the problem of development. T_4, the complete apparatus of genetics that carries the population transformation across generations, is operating in the genotypic space to transform genotypic frequencies and cannot be framed in phenotypic terms. But T_2 that specifies the transformation of

frequencies within generations is a set of physiological, ecological, and behavioral relations that operates in the phenotypic space. In the absence of a knowledge of the epigenetic transformations, the circuit is cut and nothing can be done. Nor can one depend on developmental genetics to provide the missing laws, because the main direction of developmental biology has never been, and shows no promise of ever being, to understand the environmental contingency of variations in development. Much of the past and the present problems of population genetics can be understood only as an attempt to finesse the unsolved problem of an adequate description of development. The second problem that appears in Figure 2 is the question of rates. Generally the movement in the phenotypic and the genotypic spaces is extremely small, the number of interacting genetic and phenotypic variables is large, and the number of generations and individuals over which observations can be taken is extremely limited. Thus the possibility of inferring the forces that are operating from the differences between G_1 and G_2 or G'_1 or between P_1 and P_2 is very remote. This has led to the other major preoccupation of population genetics, the possibility of estimating dynamic forces without ever actually measuring them. To the extent that population genetics has succeeded in this effort, it has provided a uniquely powerful methodology to evolutionary reconstruction that is not available from studies of physiology, ecology, and behavior, whose resolving power is necessarily low compared with the size of the actual forces operating in nature.

Finally, there is a problem that is not so much a difficulty within population genetics, but rather lies in its relationship with other branches of biology and with the rest of evolutionary theory. In these latter fields there is a strong value on finding universals, or at least in claiming universals even when there is no compelling evidence for them. So, even in a historical science like evolutionary biology, an overwhelming emphasis has been placed on selective, and even adaptive, explanations for all phenomena, to the exclusion of other possibilities, partly as a consequence of the belief that a science is validated to the degree that it can make universal claims. But population genetics has developed in a different direction. The schematic given in Figure 2 is one of the historical trajectories of a character change as a result of the stochastic realizations of the interaction of a large number of forces that alter gene frequencies. Moreover, the strength of these various forces and the signal-to-noise ratio in their dynamics depend on circumstances that are different for different species and different genes. There are no universal claims, or even great generalizations, that can be made about selection intensities, population sizes, migration rates, or about the temporal history of these forces, which are certainly varying in time, or about the amount of genetic variance likely to be present at any moment, on which these forces act. Thus the population genetic research problem is short on general hypotheses to be tested. There are some very large statistical characterizations that can be achieved by repeated observations on a variety of species and genes, and, at the other extreme, there is the

> The concentration on a molecular genetic description of population variation has de-pauperized the problematic of population genetics by marginalizing studies of the phenotype.

possibility of finding examples of the operation of particular forces like balancing selection or neutral evolution, but neither statistical generalities nor individual cases satisfy the model of what a powerful science is supposed to be.

Coping with Development

Population genetics has taken two opposite paths to solving the problem of reconciling the phenotypic and the genotypic spaces, paths that essentially finesse the problem by operating in only one of them. The first of these, characteristic of investigation before the emergence of molecular biology but now almost entirely abandoned, is to operate entirely in the phenotypic space. Phenotypic observations were used to make rough inferences about the genotypic space, but without providing a detailed description of the population in terms of genotypic frequencies. The tools and concepts were those of biometrical genetics: selection experiments on phenotypic traits in a variety of environments and a variety of selection schemes, estimates of components of genetic variance and covariance for phenotypic characters, studies of norms of reaction of fitness components, observations on fitness components or morphological traits from segregations, and recombination's involving large parts of the genome. While lacking in any genetic details, these observations and experiments produced an immense richness of information about heritable phenotypic variation in populations and left unresolved a large number of important problems that are untouched by the more recent observations of molecular population genetics. A few among the many discoveries of phenotypic population genetics that have yet to be dealt with at the level of the genotype are

1. Populations sampled from nature can usually be artificially selected for almost any morphological character, with some notable exceptions such as directional asymmetry, and some physiological characters such as alcohol tolerance, while many physiological characters such as temperature and salt tolerance or speed of development are much less selectable. The relevant evidence on genic polymorphism is either in the opposite direction or almost entirely lacking. Enzyme loci are more genetically variable than genes coding for structural proteins, and artificial selection for adult alcohol tolerance in Drosophila melanogaster produces no change in allele frequencies at the Adh locus. While homeobox genes and other similar loci active in early development have been extensively studied, nothing is known about the loci responsible for continuous selectable variation in either morphological or physiological phenotype.
2. Little is known about the number of loci segregating and the distribution of gene effects for phenotypic traits with heritable variation. Older gene mapping studies in which sparsely distributed markers were used could localize some gene effects to chromosome arms and occasionally localize a large effect. Modern **quantitative trait locus (QTL)** mapping studies in which much more densely distributed molecular markers are used

Quantitative trait locus is a section of DNA which correlates with variation in a phenotype. Usually the QTL is linked to, or contains, the genes which control that phenotype.

have a much greater discriminatory power, but it is a long way from localizing the loci responsible for phenotypic differences between two highly divergent lines to giving a description in allelic frequency terms of the standing heritable variation in a population.

3. Characters with no phenotypic variation may nevertheless have considerable underlying genetic variation that is revealed under conditions of extreme developmental stress. Despite the clear evidence that such canalized characters are common and that the degree of developmental plasticity of a character can also be selected for, including, for example, the amount of fluctuating asymmetry in characters that show no directional asymmetry, nothing is known at the genotypic level. Some glimpse of the possibilities is offered by the recent finding of Gibson and Hogness that selection for genetic assimilation of the bithorax ether phenocopy is accompanied by a change in the frequency of sequence variants at the Ubx locus.

4. Many correlations between morphological characters or between morphology and fitness are a consequence of linkage disequilibrium, but some are not or at least have not been broken by recombination and selection. The entire phenomenon of the resistance to artificial selection by countervailing fitness (usually fertility) effects needs to be investigated at the genotypic level. What is the explanation, for example, of the appearance in the classic selection experiments of Mather and Harrison of balanced sterile lines at intermediate levels of the selected phenotype so that no selection progress could be made in either direction, despite the presence of genetic variance for the character?

5. There are constraints on the pathway that phenotypic evolution can take under selection. As shown, for example, by Hall for the ebg locus in Escherichia coli, the same selected phenotype may have different genotypic bases because of different phenotypically similar mutations, some

of which allow for further selectable mutations while others are dead ends for selection because the subsequent one-step mutations do not give appropriate phenotypes

The mapping of genotypic mutational space onto phenotypic space is a major task for population genetics if an adequate genetic explanation of phenotypic selection and evolution is to be given. The alternative solution to the problem of development has been to study characters for which there is no morphogenesis, that is, to study the genome directly. This approach is an attempt to collapse the two spaces in Figure 2. It is this ploy that is the contribution of molecular population genetics, beginning, imperfectly, with electrophoretic and immunological studies of proteins and ending with the complete genotypic information contained in DNA sequences. Electrophoretic studies eliminate much of the problem of morphogenesis but there remains a residual ambiguity about genotype because even the most discriminating sequential electrophoresis will not distinguish some amino acid substitutions and there is no information on which or how many amino acid differences separate two electro morphs. There is, of course, no information about silent substitutions. Restriction enzyme studies are completely ambiguous about amino acid substitutions but at least provide information on silent sites. It is only with complete DNA sequences that a complete collapse of the phenotypic into genotypic description is possible.

The euphoria that has accompanied our ability, finally, to give a complete and unambiguous description of the allelic composition of a natural population for some arbitrarily chosen piece of the genome has hindered us from seeing that the problem of epigenesis has not been eliminated. While it is possible now to give a genotypic description of the variation in a population, the scheme of evolutionary transformation in Figure 2 still requires phenotypic information on whole organisms. Formally, what DNA sequencing does is to allow a complete description of G_1 and so we have the illusion that we do not need the morphogenetic rules T_1. But the rules of transformation T_2, the rules of mating, migration, and natural selection, are rules about phenotypes. Gene sequences do not mate, they do not migrate, they do not live or die differentially, except as a consequence of the physiology and the metabolism of their organismic carriers in interaction with the physical and biotic environment. To eliminate totally the problem of morphogenesis from our observations, we must avoid all phenotypic descriptions above the level of genotype, including fitness. A direct attempt to measure the demography and the fitness of genotypes once again introduces the contingencies of development. The epigenetic transformation T_1 is still needed to predict the distribution of phenotypes P_1 that are transformed into the selected phenotypes P_2, and we still require T_3 for carrying the phenotypes back into the genotypic space.

The problem of directly measuring the fitness of genotypes has been considerable, and the results not generally encouraging. The direct estimation of net finesses of genotypes in nature and the discrimination of fitness differences from nonrandom mating, segregation biases, migration, and stochastic elements demand the possibility of observing mating pairs and associating their offspring with them, so that fertility as

well as viability estimates can be made. The most favorable material would be annual seed plants or an animal with single clutches that can be associated with parental pairs like birds or ovoviviparous fish. The most complete and sophisticated attempt to measure all the components of the life cycle in nature for genotypes has been, in fact, the work of Christiansen on the esterase enzyme polymorphism in an ovoviviparous fish, Zoarces, but no strong fitness differences were found. Nor should this be very surprising. Given the immense amount of genotypic polymorphism present in natural populations, it is unlikely that large fitness variance can be associated with many individual polymorphic loci, and even when there are cases of strongly selected polymorphisms, it is unlikely that an arbitrary polymorphism, chosen for its ease of observation, will be one of these cases.

As substitutes for net fitness measurements, there have been attempts to measure the physiological properties of defined genotypes and then to argue that such properties are manifest as fitness differences in nature. An example of this approach is the work on the phosphoglucose isomerase enzyme polymorphisms in Colias butterflies, in which it has been demonstrated that activity differences associated with different alleles appear to cause differences in male flight ability, which in turn must influence male mating success. The question that remains open, however, and a deep difficulty of all such work on fitness in nature, is that this is a ceteris paribus argument: "All other things being equal," male mating ability is an important component of fitness, but we do not know whether male mating ability is, in fact, an important component of the standing variance in fitness in nature. The problem is similar to the situation of melanism in Biston betularia, in which melanic forms on dark backgrounds are certainly less likely to be taken by predators, but it is by no means clear that predation pressure is sufficiently high to represent an important cause of differential mortality. Moreover, the laboratory demonstration of fitness differences associated with allelic variation makes use of deliberately constructed stress conditions that may be irrelevant in nature. The polymorphic alleles at the Adh locus in D. melanogaster have very different activities on an ethanol substrate and are selected in the presence of toxic levels of ethanol in the medium. These ethanol levels, however, are above any known in nature, including those in fermentation sheds in wineries. The classical work of the Dobzhansky school on fitness in D. pseudoobscura was almost entirely in terms of larval competitive viability, yet in nature larvae of this species are never found in crowded conditions, and it seems more likely that the differential ability of females to find an egg-laying site and the ability of larvae to pupate before the small food sources dry up are much more important determinants of fitness variance. The problem of directly measuring fitness differences in nature is one that has plagued population genetics for nearly a century, and our ability to identify genotypes at the DNA level has not made it disappear.

The Problem of Rates

The difficulty of measuring fitness differences in nature is one facet of the more general issue that the individual forces operating on genotypic composition of a population

at a locus are small so that the change from one generation to another, even the real stochastic change, is usually well below the observable limit in samples of feasible size. The question then is how we are to reconstruct the dynamic process in Figure 2 when G and G' are not observably different. The answer has been to invert the process of deduction and to appeal to the cumulative effects of weak forces over long periods. From theoretical considerations we can predict the probability distribution of changes in genotypic composition under various models from one generation to the next. If the forces have remained constant for a very long period, then the population genotypic composition will come to a steady state under these forces, and, if we are lucky, that steady state will bear the unique signature of the forces, making an inference about them possible. We can use, for this purpose, not only the genotypic composition of a given population at a given locus, but the variation among loci, among populations, and among closely related species. Such synchronic or static data replace the direct measurement of the forces, and the direct diachronic observation of temporal change as the preferred method of population genetics.

Inference from static data has a long history in population genetics. In its simplest form, an observed intermediate allelic frequency of a polymorphism could be taken as evidence of balancing selection or of an equilibrium between mutation and selection. In the case of sickle-cell hemoglobin, in which the homozygote S/S is lethal, the inferred mutation rate would have to be of the order of 10^{-1} in West Africa but virtually absent elsewhere to explain the facts of the polymorphism. Clearly the postulate of balancing selection would make more sense even in the absence of the geographical correlation with malaria. The use of B/A and D/L ratios in the 1950s and 1960s were attempts to make inferences about the dominance of deleterious fitness effects when these effects could not be directly measured. But the wholesale use of static data as the bread and butter of population genetics has depended, first, on the provision of a rich and precise genotypic data set by molecular survey methods and second, on the development of an articulated stochastic theory of gene frequency evolution that can, in principle, predict different static signatures for different mixtures of forces. The first attempts to match stochastic theory to molecular data involved the use of the Ewens and Watterson distributions to test whether the frequency distributions of electrophoretic alleles at multiallelic loci like Est5-B and

Drosophila is a genus of flies, belonging to the family Drosophilidae, whose members are often called "small fruit flies" or (less frequently) pomace flies, vinegar flies, or wine flies, a reference to the characteristic of many species to linger around overripe or rotting fruit.

Adh in **Drosophila** were compatible with nonselective processes, with balanced polymorphism, or with purifying selection.

This method assumed, as do all subsequent forms of the same approach, that

- The experimental method distinguishes all genotypes,
- The forces of selection and effective population size have been constant for a long time so that
- The current allelic frequency distribution is at steady state, and
- The current allelic frequency distribution is independent of the initial conditions

These earlier attempts were not continued partly because the method of electrophoresis does not fulfill requirement 1 and partly because the tests of Watterson and Ewens turned out not to have sufficient power for such data sets to distinguish many selective hypotheses from nonselective hypotheses. In the case of Est-5b the data were not compatible with any of the alternatives, suggesting that the steady-state assumption was wrong or that a more complex selective hypothesis needed to be invoked. It soon became apparent that electrophoretic data simply did not have the necessary structure to distinguish hypotheses among forces. In the first place, most polymorphic loci are, at the electrophoretic level, biallelic or triallelic, and there is usually insufficient evidence in the frequencies of two or three alleles to test most hypotheses. There are exceptions. For example, the repeatable altitudinal clines in the two common alleles of the Adh and similar geographical clines in different continents of Gpdh and Tpi electromorphs make a selective rather than historical explanation extremely likely. Second, even for the occasional locus like Xdh or Esterase in Drosophila, which has 20 or more alleles present in a single sample in which the pattern is of one or two common alleles and a large menu of rare alleles, extreme similarity of distribution between different populations can be explained either as repeatable selection or as a consequence of a small amount of migration between the populations. Third, the bulk of loci that are electrophoretically homozygous provide no information at all about the causes of their homozygosity. It is this problem that lies at the crux of the difficulty, a problem that has been solved by a fortunate quirk of nature.

The problem of genetic identity is to distinguish two possible cases for apparent allelic identity. It may be that two gene copies are identical because they stem from a recent common ancestor, so recent that no mutations have appeared since their genealogical splitting. That is, they are identical by descent. The other possibility is that their most recent common ancestor was indeed very ancient and that many mutations have occurred since that ancestor, but these mutations have been selected against, leaving only the allowable gene form. That is, the present genes are identical by state. Nor does this distinction apply at only a completely monomorphic locus. If there are two alleles in intermediate frequency at a polymorphic locus, one might be the ancient form kept uniform by a constant purifying selection, while the other might be a recent mutation that is selectively favored and has spread rapidly since its origin,

having reached a polymorphic equilibrium with the older allele or being in the process of replacing it. On the other hand, both alleles may have been around for a very long time by chance because both are acceptable by natural selection while all other mutations have been rejected. The difficulty is that two alleles do not carry with them any information about how long they have been separated, and this is true whether they are identified by electrophoresis or by total amino acid sequencing. If only alleles also had temporal information that allowed us to distinguish identity by recent descent from enforced identity by state. But, of course, they do as a consequence of a lucky quirk of nature: the redundancy of the DNA code.

Alleles that are identical in the amino acid state may nevertheless differ in their nucleotides and if, as a first order of approximation, natural selection does not distinguish among alternative codons for the same amino acid, then gene copies can accumulate silent nucleotide substitutions unimpeded by natural selection. The number of such silent differences is then an index to the time that has passed since their common ancestor and can be used to distinguish among competing theories. It is vital to understand that DNA sequencing as the method of choice for population genetics does not stem from the fact that somehow DNA is a more basic or fundamental state description than amino acid sequences, but from the many–one correspondence between DNA sequences and amino acid sequences. Had the relation turned out to be one–one, there would be no advantage to DNA sequencing except as a cheap way to determine protein sequence.

More broadly, it is the presence, side by side in the same short gene region, of coding sites together with silent sites, with introns, and with untranslated and non-transcribed flanking DNA, all of which are subject to the same temporal processes of mutation and reproduction, but that differ from one another in the physiological consequences of their variation, that allows us to distinguish selective from purely historical similarities and differences. While the rationale for DNA sequencing is often stated to be that silent sites in codons are selectively neutral and so can serve as a selection-free molecular clock against which amino acid substitutions can be calibrated, that neutrality is not necessary (fortunately, since it is not true). All that we require is that different DNA positions have physiological constraints operating at different levels of intensity and in different contextual situations, so that the static variation observed within and between populations bears the distinguishable signatures of different temporal processes.

The most striking use of the temporal information in DNA sequences and the one with the greatest implications for our understanding of dynamical processes was, in fact, the first application of the method by Kreitman in his study of the standing variation at the Adh locus in D. melanogaster. Except for the single Fast/Slow (F/S) electrophoretic polymorphism that was deliberately introduced into the sample by the choice of lines, there was not a single amino acid polymorphism in a geographically very diverse sample, yet silent sites and introns were approximately 6% polymorphic (14% in exon 3). Thus the identity of the amino acid sequences could not be the result of recent common ancestry, but must be the outcome of a selection process in which all

amino acid replacements, except the single widespread F/S polymorphism, have been removed by selection. This result was made even more striking by that of Schaeffer and Miller, which showed that in a sample of 99 genomes of D. pseudoobscura, there is no amino acid polymorphism at all at Adh, with the exception of one line with an isoleucine/valine change, despite the fact that there is ~6% silent nucleotide polymorphism. The implications of these results for our understanding of natural selection are powerful.

Three-quarters of all random single nucleotide changes in coding regions will cause an amino acid substitution, yet every one of these has been screened out by natural selection in a protein that is, for example, 28% leucine/isoleucine/valine. While selection differences need not be large for any substitution, since population sizes for these species are estimated to be, conservatively, in excess of 10^5, selection has been discriminating enough to weed out every substitution. The problem is how every amino acid substitution can make a physiological difference that is ultimately translated into an average difference in viability and fertility. Nor is this discriminatory power of natural selection confined to a small enzyme like alcohol dehydrogenase.

When similar studies were done on the very large and electrophoretically very polymorphic enzymes xanthine dehydrogenase and esterase-5b a similar but less extreme result was found. Approximately 90% of all amino acid substitutions in xanthine dehydrogenase and 85% in esterase-5b have been rejected by natural selection. If these results are indeed general, then a major problem has been posed for cellular physiology and metabolism and for evolutionary biology. Given environmental contingency and the accidents of development, how can the amino acid composition of proteins be translated so exquisitely into differential fertility and viability of individuals? If that translation does occur, how can the amino acid compositions of enzyme proteins change by multiple amino acids during speciation, as they indeed have? Of course, speciation may be accompanied by drastic reduction in population size so that Ns is temporarily small and random fixations of amino acids differences may occur. But that cannot be the whole story.

As pointed out by McDonald and Kreitman, neutral fixations during speciation should reproduce between species the same ratio of silent to replacement substitutions as are polymorphic within species. Yet in some cases, for example glucose-6- phosphate dehydrogenase in D. melanogaster and D. simulans there is a great excess of amino acid differences compared with the intraspecific polymorphism. The McDonald–Kreitman approach to detecting selectively driven divergence between species is another opening into an old problem made possible by the redundancy of the DNA code. It is widely recognized by systematists, even those trained to avoid typologies, that many characters that distinguish taxa have no detectable variation within species. The genus Drosophila is distinguished from other members of the Drosophilidae by the possession, by every individual in every species in the genus, of one proclinate and two reclinate orbital bristles. Given the lack of intrataxon variation, where did the variation for taxonomic differentiation come from and where did it go after the event? Many more or less plausible stories can be told, but there is no available

method for demonstrating that natural selection was involved. For nucleotide sequences the possibility of implicating natural selection exists because we can make internal comparisons of polymorphism and species divergence for **nucleotide** positions of different function.

Nucleotides are organic molecules that serve as the monomer units for forming the nucleic acid polymers deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), both of which are essential biomolecules in all life-forms on Earth.

Regional heterogeneity in the density of polymorphism along a DNA sequence can arise for reasons other than the direct constraints on the sequence itself because of the extremely tight linkage among the nucleotides in short genomic sections. Thus a region of unusually high nucleotide variation can be taken as evidence that random fixation of mutations has been resisted by some form of balancing selection. Kreitman and Hudson used this approach to infer heterosis for the F/S electrophoretic alleles of Adh in D. melanogaster, although their data show an excess nucleotide polymorphism even within the Slow allelic class, so that something else besides heterosis of the two electrophoretic forms must also be happening.

In contrast, extremely low variation in a genomic region might be the equilibrium signature of a high functional constraint on the region, but it might also be the historical leftover of a recent rapid replacement of one or a few sites that carried adjacent nucleotides to fixation by hitchhiking. This explanation by Berry of nucleotide homogeneity of the fourth chromosome must surely be the correct one. Such a rapid fixation, however, might be the result of a selectively advantageous mutation's having been fixed, a so-called selective sweep, but it might also be the consequence of a segregation distortion. Because the segregation distorters (SDs) that have been studied are necessarily those that are in some sort of balanced equilibrium in populations, like the SD in Drosophila and the t locus in mice, we do not know how often distorters, unopposed by selection, may sweep through populations.

Whether or not one can distinguish selective sweeps from distorter sweeps depends on the rapidity with which the two processes occur and on the time since the event. In the absence of any ancillary evidence about replacement speed and relaxation time, it is hard to see how these processes can be separated. The problem of distinguishing selection from segregation distortion applies as well to the inference of adaptive evolution between species. Both are illustrations of a deep structural problem in making inferences from static variation. Such data are, at the first level, information on the

times of divergence from a common ancestral sequence, in mutational units, between genomes in populations. But the distribution of divergence (coalescence) times is a dual representation of the distribution of rates, in mutational units, of fixation of particular sequences through the population. An unusually low fixation rate of some stretch of genome compared with the average for the genome may mean that there are forces holding sequences in equilibrium against random fixation or that the time scale for the production of variation in the region is more compressed. Such a compression can come from a local excess in mutability or from a migration into the population of genomes that were divergent in this genic region. In a like manner, a deficiency of variation means either that some genome has spread unusually rapidly through the population, for whatever reason, or that the time scale for the production of variation has been extended because of selective constraints. The decision from among these alternatives and the biological interpretation of the causes of the changes in time rates must be made on other grounds not contained in the static data themselves.

The discovery by molecular biology of different functional classes of nucleotides has not only led to the possibility of obtaining strong evidence about old problems, but has revealed a new set of problems for population genetics, problems that arise from considering the standing variation within those nucleotide classes. These are the problems of explaining regional heterogeneity in DNA sequences even within functional types. Thus, at the dpp locus in Drosophila, a gene that sits at the center of many of the processes of early embryonic development, there is a major heterogeneity in the pattern of nucleotide polymorphism in the large intron. There is a concentration of polymorphism at the two ends of the intron and little variation in the middle of the intron both within and between species.

Some of this conservation of the middle of the intron coincides with multiple short repressor motifs in the region, but much of it remains to be explained on functional grounds. There is also a region of perfect conservation within and between species in the 30 untranslated region of Intron 3, in the middle of a region of considerable nucleotide divergence and polymorphism. What are the functional sources of such conservations and how frequent are they? One of the most interesting and general unresolved problems arising from observations of genomic heterogeneity is the question of codon bias. A reasonable explanation of unequal codon usage is that, especially for highly transcribed genes, unequal availability of tRNA species constrains the codon usage to match it for the most efficient translation. But this explanation raises many problems when different species and different genes are compared. For example, the AUA codon is totally avoided for isoleucine in Adh in all species of Drosophila, yet for other genes in D. melanogaster, it is used on average ~5% of the time. In D. virilis, however, it comprises 38% of the isoleucine codons for the dpp gene. Moreover, although a clear bias toward G/C ending alternative codons is general for genes in D. melanogaster, in D. virilis the dpp gene is biased toward A/T in eight amino acid groups. Have there been major changes in tRNA availability from one species to another, and, if so, why do different genes in the same species not agree in their codon bias differences from other species? The existence of specific usages of gene and species make simple explanations of the evolution of codon usage doubtful. The issue

of codon usage raises a problem of constraint on DNA sequences that is not ordinarily considered. The secondary structure of message, its stability, and the time course of its translation will be influenced by the nucleotide sequence. We should not suppose that a maximum rate of translation is necessarily optimal for the translation process. The final folding of a protein occurs as a consequence of the formation of folding intermediates during the process of translation, and, presumably, some slowing down and speeding up of chain elongation at appropriate places is necessary for a realization of local free-energy minimizations during folding. These variations in speed will be affected by the particular DNA sequence, partly in its relation to the availability of tRNA, so that the problem is more complicated than that of simply maximizing the use of common tRNA species. Nor should we even suppose that fidelity of translation is an unmitigated good.

The translational process, like any molecular mechanism, makes mistakes. There must be a population of variant protein molecules in the cell. Are these variants of some physiological use? After all, if one argues that heterosis comes from the advantage of having two slightly different copies of a molecule in each cell, then why not have more than two different copies and in unequal numbers? But all of these considerations lead us to suppose that the DNA sequence itself, even in coding regions, may be constrained over stretches of various length. It is the common assumption that constraints on amino acid variation are the consequence of requirements on the physiological function of the protein in question. But if we suppose that DNA sequence itself is constrained over even short stretches, then the amino acid sequence will be held constant as a consequence of the constancy of the underlying nucleotide sequence.

At the moment we do not know how much of the variation and the conservation of protein sequences is a secondary consequence of requirements on the nucleotide sequence itself. Nor do we know how to find out. There is yet another possibility opened by data from DNA sequence that returns us to a stage in the history of inference. The attempt to make judgements about forces from static data on electrophoretic variation foundered for three reasons: ambiguity about the actual genetic identity of the observed classes, insufficiently rich data sets even for the few multiallelic loci, and a stochastic theory that was derived from the consideration of the moments of stationary distributions rather than the richer structure of gene phylogenies. The first two problems are dealt with directly by the nature of DNA sequence data and by an immense amount of haplotypic diversity that appears even in modest data sets. The third problem has been attacked by developments in stochastic theory that take into account the details of haplotypic phylogenies, although still suffering from the need to make many simplifying assumptions. Essentially these developments attempt to compare the observed distribution of DNA haplotypic differences within a sample with those expected from purely neutral evolution. The null hypothesis that produces the expected distribution of differences among haplotypes consists, necessarily, of a long list of assumptions, so what is being tested when the data are fit to the null hypothesis is the conjunction of a large number of neutral assertions. Some of these are background assumptions about population history and linkage that are not of

primary interest in the test. Others are relevant to whether natural selection needs to be invoked to explain the observed pattern of differences. For example, the method given by Hudson involves the differences between two species for two or more genomic regions and the polymorphism within these regions.

The background assumptions are that generations are discrete, that mutations at different sites occur independently and with a common mean for sites within a locus, that the species are at a stationary state, that they diverged from an ancestral population whose population size was the average of the stationary sizes of the current species, and that there is complete linkage within gene regions but free recombination between them. The selectively relevant assumptions are that the segregating mutations that are actually segregating within the species are selectively equivalent, that the degree of purifying selection may differ between regions but is the same between species, and that no change in selective regime occurred in the process of speciation. We do not know, in general, how deviations from the background assumptions affect the power of the test compared with the selective assumptions, nor do we know, in general, what the power of the test is to detect one sort of selective deviation as opposed to another. Thus we do not know at present how to interpret observations of significant or non-significant test results. In practice, such tests can be subject to an extensive investigation of their operating characteristics. Simonsen have shown for the simpler test of Tajima that the power to detect selection is low, a result that is reminiscent of the earlier situation of similar tests on electrophoretic data. The important question is less whether the tests are powerful in any absolute sense, but what their relative power is to detect deviations from the background assumptions as opposed to deviations from the selective assumptions. Tests of low power are only useless. Tests that powerfully detect the wrong deviations are destructive.

Problems of Generality

Twenty years of electrophoretic surveys and 15 years of DNA sequencing and restriction analysis make it clear that there is a lot of standing genetic variation in populations both at the amino acid level and the DNA level, but there is a great deal of difference among species and gene regions in how much genetic variation is to be found. It is also possible from theoretical considerations to make some predictions about correlations, as for example that regions of the genome with low recombination should turn out to be regions of lower genetic variation, all other things being equal, as indeed has been observed for a couple of chromosomal regions in D. melanogaster. How well this correlation holds in other organisms with different demographic histories remains to be seen. The alternative to statistical generalization is the exemplification of particular phenomena without any strong implication about their frequency in nature. So, for example, a convincing case has been made that the Adh electrophoretic variation in D. melanogaster is a balanced polymorphism, but, unlike the universalist claims about heterosis that were common in population genetics in the 1950s and 1960s, it has never been suggested that Adh in D. melanogaster provides any general insight into genetic variation in general. In like manner, the strong case made by

Eanes that G6pdh diverged selectively in D. melanogaster and D. simulans is not taken as a general refutation of claims of the neutral evolution of **amino acid** sequences. But producing an example, or even several, of selective divergence, neutral evolution, and balanced polymorphism, a selective sweep-out of variation, or linkage equilibrium or disequilibrium fails to match the Universalist model on which so much of science is built. It is, moreover, a self-terminating research program since its objective, to demonstrate that certain evolutionary cases do indeed occur in nature under appropriate circumstances, is soon realized.

Amino acids are organic compounds containing amine ($-NH_2$) and carboxyl (-COOH) functional groups, along with a side chain (R group) specific to each amino acid.

There is a third pathway of research in population genetics that is exceedingly difficult of accomplishment partly for reasons of technique and partly for reasons of the intellectual division of labor in biology. That pathway is to bring population genetics into the main methodological tradition of functional biology by attempting to measure the forces operating in populations directly, rather than by statistical inference, and by carrying out functional experiments that lead to a knowledge of the physiological, ecological, and behavioral mediations for evolutionary change. To do so means solving the problem of rates.

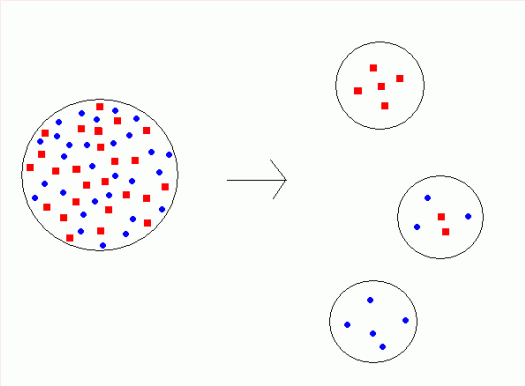
First, it is necessary to make cellular and organismic physiological measurements of the effects of nucleotide substitutions, by use of the available apparatus of in vitro mutagenesis and recombination followed by a molecular and physiological analysis of the relevant metabolic functions including transcription and translation, as has been attempted for the Adh locus in Drosophilae. But the demonstration of some difference in some parameter of cellular activity is not sufficient. It must be possible to show that these physiological differences mediate fitness differences whose values are at least 2 orders of magnitude smaller than what has been measured in the laboratory. The problem is not one of sufficient experimental population size because even in Drosophilae populations of effective size, 105 can be managed. The major problem is the detection, in sufficiently large samples taken over sufficiently long time periods, of the small changes brought about by selection intensities of the order of 10^{-3} or smaller.

The time problem can be solved by the choice of an organism with generation times of the order of a day. The sample size problem can only be solved by labor or by mechanization. And to what end? If population genetics can only illustrate a

few exemplars of processes or find very general correlations, what purpose is served by being able to provide the complete path of mediation from DNA to fitness? The answer lies in the essentially methodological nature of the science. While progress in molecular biology has consisted both of learning how things actually are in nature (the double-helical nature of DNA) and in finding methods for asking such questions (how to sequence DNA), the discoveries, both theoretical and experimental, that have had the most impact on population genetics are not those that revealed a particular fact about the world, but those that have provided methodologies for answering particular questions when desired. That is why theoretical investigations have played such a central role in population genetics as opposed to any other branch of biology. Population genetics is a science of methodologies that provides the concepts and techniques that can be used to resolve, in favorable circumstances, problems in evolution, human genetics, plant and animal breeding, and ecology. The greatest methodological challenge that it now faces is to connect the observations about the outcome of evolutionary processes to the tradition of experimental functional biology.

CONCEPT OF GENETIC DRIFT

Genetic drift is a fundamental concept in population genetics that describes the random changes in the frequency of genetic variants (alleles) within a population over time. Unlike natural selection, which acts on specific traits that confer advantages or disadvantages in survival and reproduction, genetic drift is a stochastic process driven by chance events. It becomes particularly influential in small populations, where random fluctuations can have a significant impact on the genetic makeup. As individuals reproduce and pass on their genes to the next generation, some alleles may become more prevalent while others decrease or may even be lost entirely. Over successive generations, genetic drift can lead to the fixation of certain alleles (where they become the only variant present in the population) or the loss of others, reducing genetic diversity. Genetic drift is essential in understanding the evolutionary history of populations, the effects of population bottlenecks, and the genetic health of endangered species. Its influence on the genetic makeup of populations makes it a crucial component of population genetics research, complementing other evolutionary forces such as mutation, natural selection, and gene flow.



SUMMARY

- Population genetics is a field of biology that studies the genetic composition of biological populations, and the changes in genetic composition that result from the operation of various factors, including natural selection.
- Genetic drift is a random process that can significantly impact small populations. It occurs when genetic variation is influenced by chance events rather than natural selection.
- The collection of all the alleles of all of the genes found within a freely interbreeding population is known as the gene pool of the population. Each member of the population receives its alleles from other members of the gene pool (its parents) and passes them on to other members of the gene pool (its offspring).
- Genetic drift is a mechanism of evolution that occurs by random chance rather than natural selection. In genetic drift, a population experiences a change in the frequency of a given allele, prompted by random luck rather than a need for adaptation.
- Recessive deleterious mutations are relatively rare, generally in the order of 1 per 105 or 106 mutant gametes per generation. Their constant occurrence over the generations, combined with the even greater rarity of back mutations, leads to a gradual accumulation in the population.
- Alleles that are identical in the amino acid state may nevertheless differ in their nucleotides and if, as a first order of approximation, natural selection does not distinguish among alternative codons for the same amino acid, then gene copies can accumulate silent nucleotide substitutions unimpeded by natural selection.
- Regional heterogeneity in the density of polymorphism along a DNA sequence can arise for reasons other than the direct constraints on the sequence itself because of the extremely tight linkage among the nucleotides in short genomic sections.

MULTIPLE CHOICE QUESTIONS

1. Approximately 4% of the turtle population in the pond shows the recessive phenotype long nose (nn). What is the frequency of the dominant allele (N) in the population?

a. 0.16
b. 0.20
c. 0.40
d. 0.32
e. 0.80

2. Within the squirrel population at park, 16% show the recessive phenotype of a curled tail (tt). What is the frequency of the dominant allele (T) in the population?

a. 0.40
b. 0.16
c. 0.26
d. 0.60
e. 0.32

3. The effects of natural selection may be countered by

a. gene flow
b. genetic drift
c. mutation
d. inbreeding

4. Mating with relatives is called

a. inbreeding
b. outcrossing
c. random mating
d. clines
e. polymorphic

5. The random loss of alleles in a population is called

a. mutation
b. selection
c. genetic drift
d. electrophoresis
e. gene flow

Answer

1. (e) 2. (d) 3. (a) 4. (a) 5. (c)

REVIEW QUESTIONS

1. Discuss about Darwinian population genetics.
2. Give an overview on gene doping.
3. Explain the effect of genetic diversity.
4. What do you understand by the 'Hardy-Weinberg equilibrium'?
5. How to find problems of generality in population genetics?
6. Define a population gene pool and explain how the size of the gene pool can affect the evolutionary success of a population.
7. What is random genetic drift?

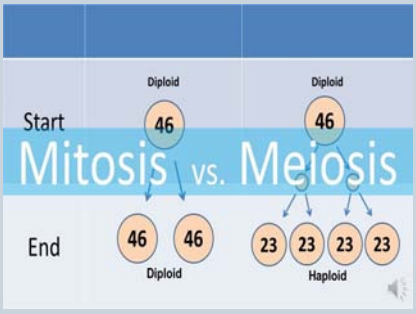
REFERENCES

1. A. V. Cano, H. Rozhonova, A. Stoltzfus, D. M. McCandlish and J. L. Payne (2022-02-10). “Mutation bias shapes the spectrum of adaptive substitutions”. Proc Natl Acad Sci U S A. 119 (7).
2. Crow, James F. (2010). “Wright and Fisher on Inbreeding and Random Drift”. Genetics. 184 (3): 609–611.
3. Gillespie, John (2004). Population Genetics: A Concise Guide (2nd ed.). Johns Hopkins University Press.
4. H. Maughan, J. Masel, C. W. Birky, Jr. and W. L. Nicholson (2007). “The roles of mutation accumulation and selection in loss of sporulation in experimental populations of Bacillus subtilis”. Genetics. 177 (2): 937–48.
5. Hastings, P. J.; Lupski, J. R.; Rosenberg, S. M.; Ira, G. (2009). “Mechanisms of change in gene copy number”. Nature Reviews Genetics. 10 (8): 551–564.
6. Masel, J.; King, O. D.; Maughan, H. (2007). “The loss of adaptive plasticity during long periods of environmental stasis”. American Naturalist. 169 (1): 38–46.
7. Miko, I. (2008). “Epistasis: Gene interaction and phenotype effects”. Nature Education. 1 (1): 197.



CHAPTER 4

MITOSIS AND MEIOSIS



It is also called indirect cell division or somatic cell division or equational division. In this, mature somatic cell divides in such a way that chromosomes number is kept constant in daughter cells equal to those in parent cell, so the daughter cells are quantitatively as well as qualitatively similar to the parental cell.

It is a special type of division in which the chromosomes duplicate only once, but cell divides twice. So one parental cell produces 4 daughter cells; each having half the chromosome number and DNA amount than normal parental cell. So meiosis is also called reductional division.

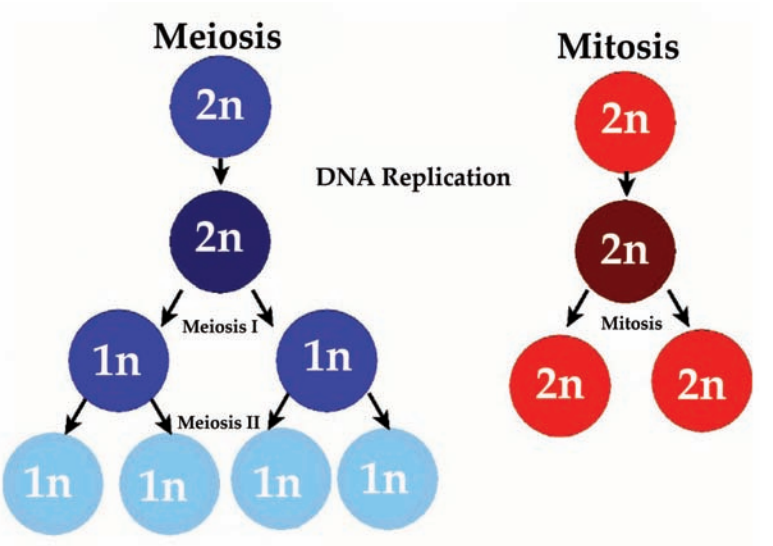
After studying this chapter, you will understand the core concepts of:

- Mitosis
- Meiosis

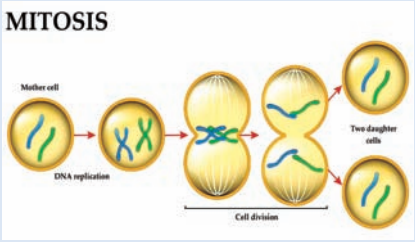
INTRODUCTION

Mitosis and meiosis are two fundamental processes involved in cell division, playing essential roles in growth, development, and reproduction. Both processes are critical for the maintenance of genetic stability and diversity in living organisms. Mitosis is the type of cell division that occurs in somatic cells, resulting in the production of two identical daughter cells with the same genetic information as the parent cell. The process of mitosis consists of several stages, including prophase, metaphase, anaphase, and telophase, during which the duplicated chromosomes are distributed equally between the daughter cells. Mitosis is crucial for the growth, repair, and maintenance of tissues and is responsible for the development of a multicellular organism from a single fertilized egg.

On the other hand, meiosis is a specialized type of cell division that occurs in reproductive cells (gametes), leading to the production of haploid cells with half the number of chromosomes as the parent cell. Meiosis consists of two consecutive divisions: meiosis I and meiosis II, each consisting of prophase, metaphase, anaphase, and telophase stages. Meiosis I results in the separation of homologous chromosomes, while meiosis II separates the sister chromatids of each chromosome. The end result is the production of four genetically diverse haploid gametes, which are essential for sexual reproduction. During fertilization, two haploid gametes from different parents combine to form a diploid zygote with a complete set of chromosomes, ensuring genetic diversity in the offspring.



The differences between mitosis and meiosis lie in their outcomes and the way they distribute genetic information. Mitosis produces genetically identical daughter cells, while meiosis generates genetically diverse gametes. Additionally, mitosis preserves the diploid chromosome number of the parent cell, while meiosis reduces it by half to ensure the proper chromosome number in the offspring. These distinct processes are crucial for the development and survival of all sexually reproducing organisms.



CORE CONCEPT: MITOSIS

Mitosis is the process in which a eukaryotic cell nucleus splits in two, followed by division of the parent cell into two daughter cells. The word "mitosis" means "threads," and it refers to the threadlike appearance of chromosomes as the cell prepares to divide. Early microscopists were the first to observe these structures, and they also noted the appearance of a specialized network of microtubules during mitosis.

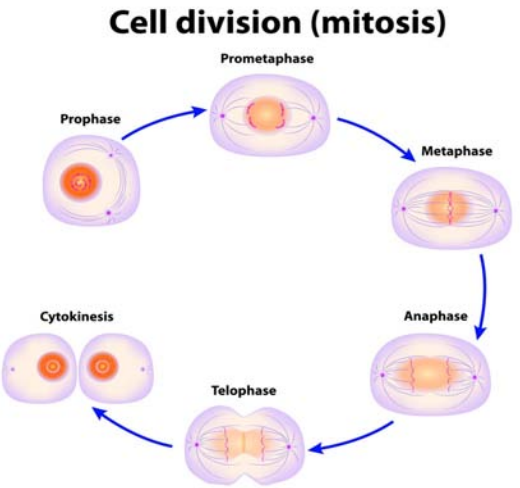
Mitosis, a process of cell duplication, or reproduction, during which one cell gives rise to two genetically identical daughter cells. Strictly applied, the term mitosis is used to describe the duplication and distribution of chromosomes, the structures that carry the genetic information.

Prior to the onset of mitosis, the chromosomes have replicated and the proteins that will form the mitotic spindle have been synthesized. Mitosis begins at prophase with the thickening and coiling of the chromosomes. The nucleolus, a rounded structure, shrinks and disappears. The end of prophase is marked by the beginning of the organization of a group of fibres to form a spindle and the disintegration of the nuclear membrane.

The chromosomes, each of which is a double structure consisting of duplicate chromatids, line up along the midline of the cell at metaphase. In **anaphase** each chromatid pair separates into two identical chromosomes that are pulled to opposite ends of the cell by the spindle fibres. During telophase, the chromosomes begin to decondense, the spindle breaks down, and the nuclear membranes and nucleoli re-form. The cytoplasm of the mother cell divides to form two daughter cells, each containing the same number and kind of chromosomes as the mother cell. The stage, or phase, after the completion of mitosis is called interphase.

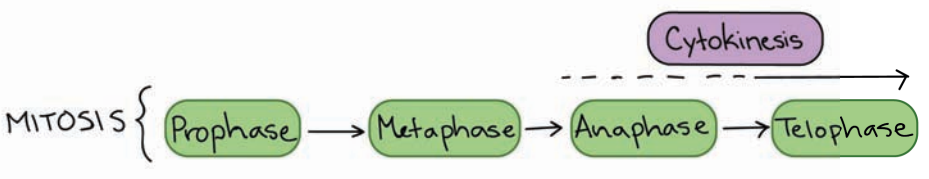
Anaphase is the stage of mitosis after the metaphase when replicated chromosomes are split and the daughter chromatids are moved to opposite poles of the cell.

Mitosis is absolutely essential to life because it provides new cells for growth and for replacement of worn-out cells. Mitosis may take minutes or hours, depending upon the kind of cells and species of organisms. It is influenced by time of day, temperature, and chemicals.



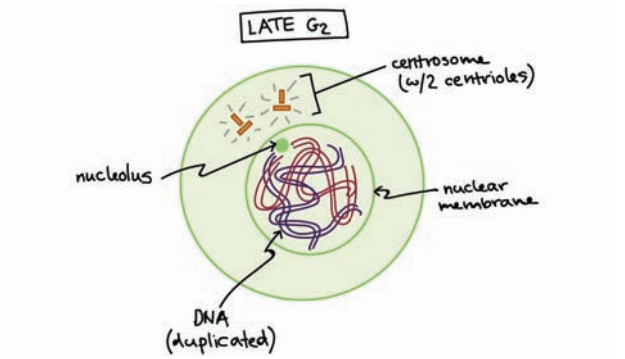
Phases of mitosis

Mitosis consists of four basic phases: prophase, metaphase, anaphase, and telophase. Some list five, breaking prophase into an early phase (called prophase) and a late phase (called prometaphase). These phases occur in strict sequential order, and cytokinesis - the process of dividing the cell contents to make two new cells - starts in anaphase or telophase.



Stages of mitosis: prophase, metaphase, anaphase, telophase. Cytokinesis typically overlaps with anaphase and/or telophase.

You can remember the order of the phases with the famous mnemonic: [Please] Pee on the MAT. But don't get too hung up on names – what's most important to understand is what's happening at each stage, and why it's important for the division of the chromosomes.

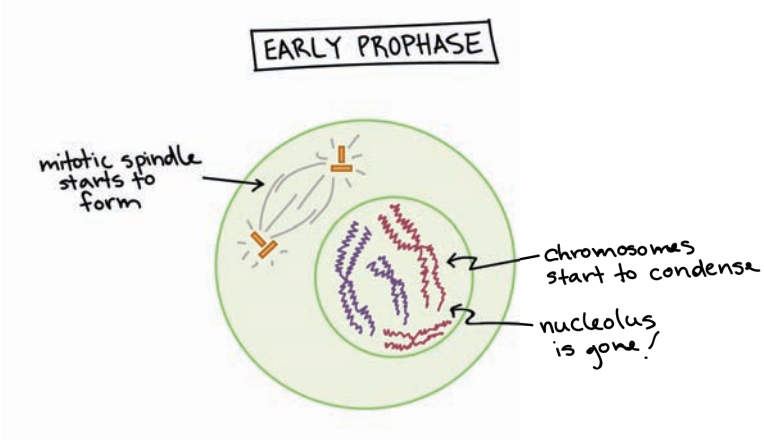


Late G2 phase. The cell has two centrosomes, each with two centrioles, and the DNA has been copied. At this stage, the DNA is surrounded by an intact nuclear membrane, and the nucleolus is present in the nucleus.

Let's start by looking at a cell right before it begins mitosis. This cell is in interphase and has already copied its DNA, so the chromosomes in the nucleus each consist of two connected copies, called **sister chromatids**. You can't see the chromosomes very clearly at this point, because they are still in their long, stringy, decondensed form.

Sister Chromatid refers to the identical copies (chromatids) formed by the replication of a chromosome, with both copies joined together by a common centromere.

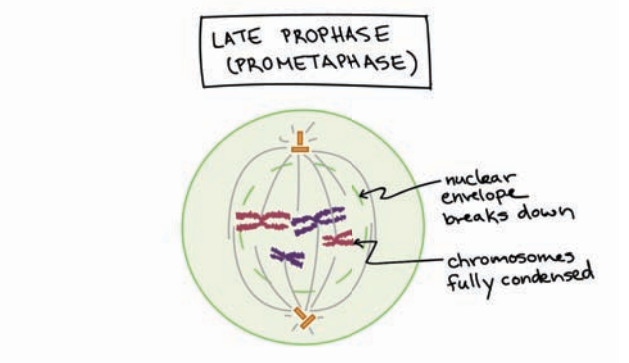
This animal cell has also made a copy of its centrosome, an organelle that will play a key role in orchestrating mitosis, so there are two centrosomes. (Plant cells generally don't have centrosomes with centrioles, but have a different type of microtubule organizing center that plays a similar role.)



In early prophase, the cell starts to break down some structures and build others up, setting the stage for division of the chromosomes.

- The chromosomes start to condense
- The **mitotic spindle** begins to form. The spindle is a structure made of microtubules, strong fibers that are part of the cell's "skeleton." Its job is to organize the chromosomes and move them around during mitosis. The spindle grows between the centrosomes as they move apart.
- The nucleolus (or nucleoli, plural), a part of the nucleus where ribosomes are made, disappears. This is a sign that the nucleus is getting ready to break down.

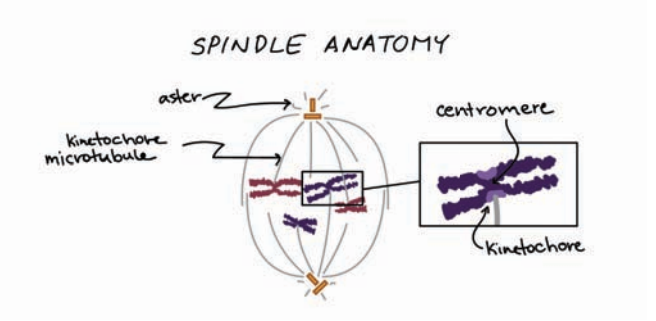
Mitotic Spindle refers to the cytoskeletal structure of eukaryotic cells that forms during cell division to separate sister chromatids between daughter cells.



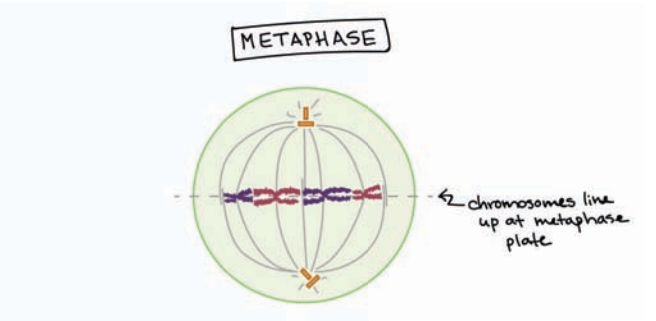
Late prophase (prometaphase). The nuclear envelope breaks down and the chromosomes are fully condensed.

In late prophase (sometimes also called prometaphase), the mitotic spindle begins to capture and organize the chromosomes.

- The chromosomes finish condensing, so they are very compact.
- The nuclear envelope breaks down, releasing the chromosomes.
- The mitotic spindle grows more, and some of the microtubules start to "capture" chromosomes.



Microtubules can bind to chromosomes at the kinetochore, a patch of protein found on the centromere of each sister chromatid. (Centromeres are the regions of DNA where the sister chromatids are most tightly connected.)

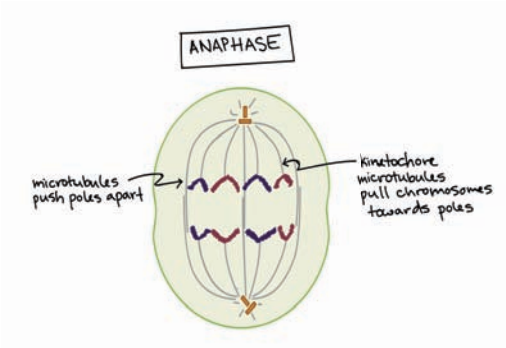


Microtubules that bind a chromosome are called kinetochore microtubules. Microtubules that don't bind to kinetochores can grab on to microtubules from the opposite pole, stabilizing the spindle. More microtubules extend from each

centrosome towards the edge of the cell, forming a structure called the aster. In metaphase, the spindle has captured all the chromosomes and lined them up at the middle of the cell, ready to divide.

- All the chromosomes align at the metaphase plate (not a physical structure, just a term for the plane where the chromosomes line up).
- At this stage, the two kinetochores of each chromosome should be attached to microtubules from opposite spindle poles.

Before proceeding to anaphase, the cell will check to make sure that all the chromosomes are at the metaphase plate with their kinetochores correctly attached to microtubules. This is called the spindle checkpoint and helps ensure that the sister chromatids will split evenly between the two daughter cells when they separate in the next step. If a chromosome is not properly aligned or attached, the cell will halt division until the problem is fixed.

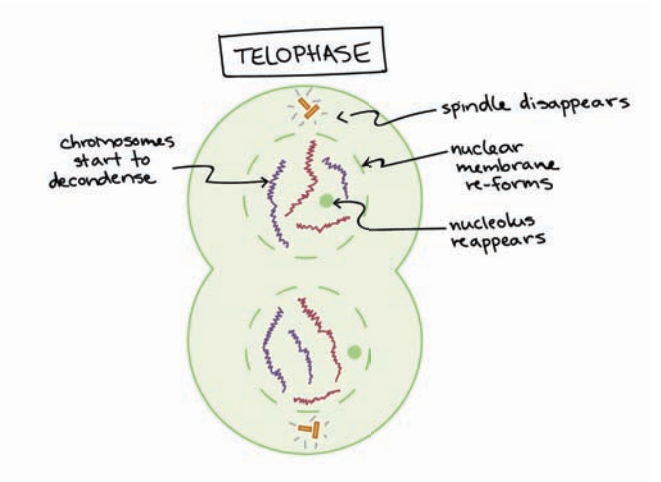


Anaphase. The sister chromatids separate from one another and are pulled towards opposite poles of the cell. The microtubules that are not attached to chromosomes push the two poles of the spindle apart, while the kinetochore microtubules pull the chromosomes towards the poles. In anaphase, the sister chromatids separate from each other and are pulled towards opposite ends of the cell.

- The protein "glue" that holds the sister chromatids together is broken down, allowing them to separate. Each is now its own chromosome. The chromosomes of each pair are pulled towards opposite ends of the cell.
- Microtubules not attached to chromosomes elongate and push apart, separating the poles and making the cell longer.

Motor proteins are a class of molecular motors that can move along the cytoplasm of animal cells.

All of these processes are driven by **motor proteins**, molecular machines that can "walk" along microtubule tracks and carry a cargo. In mitosis, motor proteins carry chromosomes or other microtubules as they walk.



In telophase, the cell is nearly done dividing, and it starts to re-establish its normal structures as cytokinesis (division of the cell contents) takes place.

- The mitotic spindle is broken down into its building blocks.
- Two new nuclei form, one for each set of chromosomes. Nuclear membranes and nucleoli reappear.
- The chromosomes begin to decondense and return to their "stringy" form.

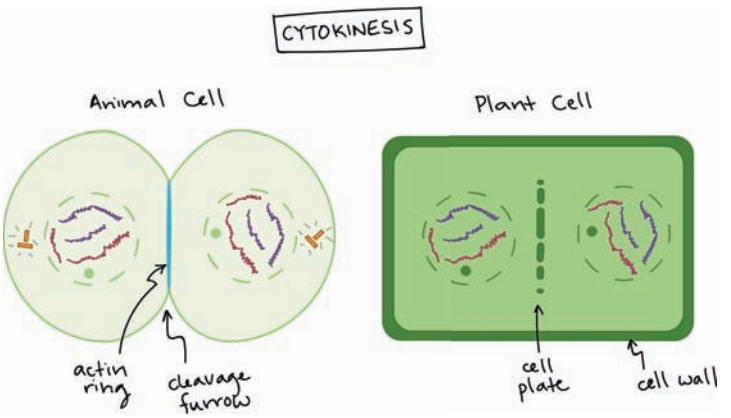
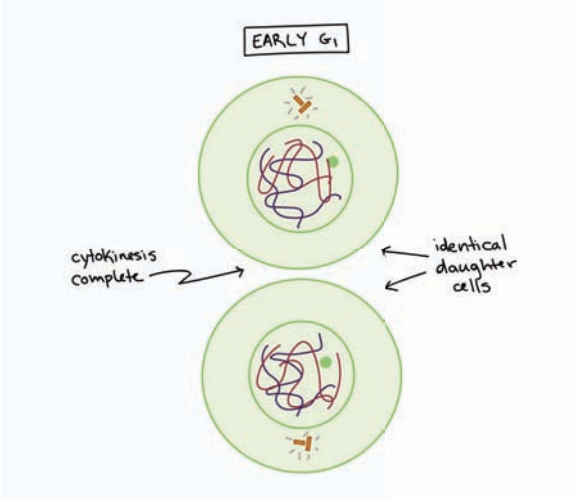


Figure 1: Cytokinesis, the division of the cytoplasm to form two new cells, overlaps with the final stages of mitosis. It may start in either anaphase or telophase, depending on the cell, and finishes shortly after telophase.

In animal cells, cytokinesis is contractile, pinching the cell in two like a coin purse with a drawstring. The "drawstring" is a band of filaments made of a protein called actin, and the pinch crease is known as the cleavage furrow. Plant cells can't be divided like this because they have a cell wall and are too stiff. Instead, a structure called the cell plate forms down the middle of the cell, splitting it into two daughter cells separated by a new wall.



When division is complete, it produces two daughter cells. Each daughter cell has a complete set of chromosomes, identical to that of its sister (and that of the mother cell). The daughter cells enter the cell cycle in G1.

When cytokinesis finishes, we end up with two new cells, each with a complete set of chromosomes identical to those of the mother cell. The daughter cells can now begin their own cellular "lives," and – depending on what they decide to be when they grow up – may undergo mitosis themselves, repeating the cycle.

Stages

Prophase

The DNA in the nucleus has already been duplicated in the previous stage of cell division, so by the time prophase starts, the nucleus contains two complete identical sets of DNA. As prophase begins, the chromatin, which are normally spread throughout the nucleus, begin to condense into an X shape, held together in the middle with a specific sequence of DNA called a **centromere**. Each half of the X is one replicated half of DNA. Once they coil together into the X, they're called mitotic chromosomes. Towards the end of prophase, the material enclosing the nucleus and the cytoskeleton disappear, except in the case of some fungi, algae, and similar organisms, in which the process happens entirely inside the nuclear membrane. This is called closed mitosis.

Centromere is the specialized DNA sequence of a chromosome that links a pair of sister chromatids (a dyad).

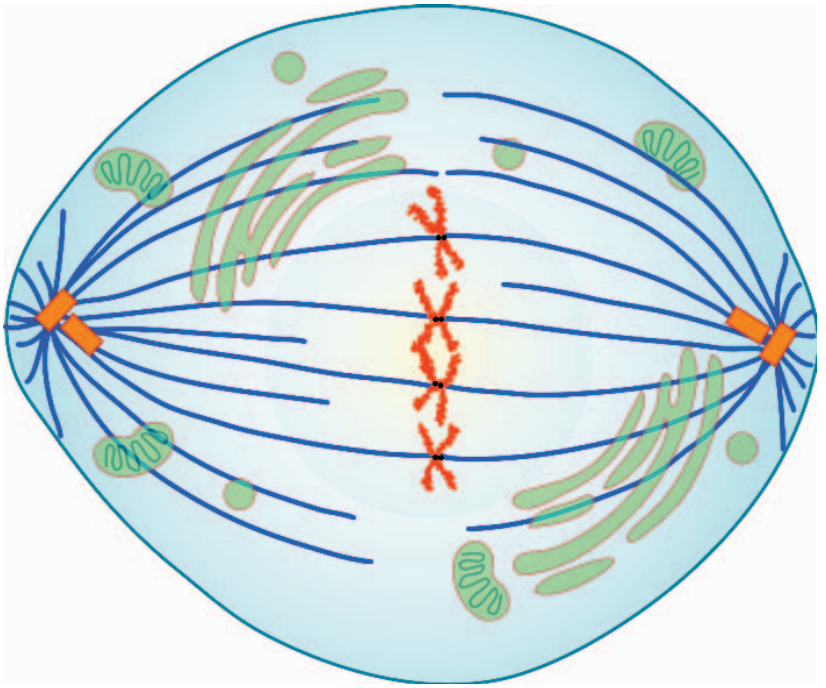
Once the material enclosing the nucleus dissolves, or, in the case of closed mitosis, after the DNA forms into Xs, structures

called centrioles move to opposite ends of the cell and help make a spindle apparatus of microtubules, which is essentially like ropes running across the cell. The chromosomes also develop structures in the middle called kinectochores.

The process of mitosis is divided into stages corresponding to the completion of one set of activities and the start of the next. These stages are prophase, prometaphase, metaphase, anaphase, and telophase. During mitosis, the chromosomes, which have already duplicated, condense and attach to spindle fibers that pull one copy of each chromosome to opposite sides of the cell.

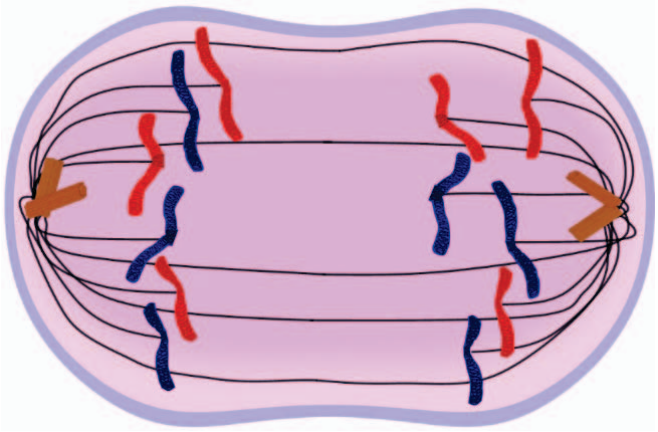
Metaphase

As prophase finishes and metaphase begins, the rope-like microtubules connect to the kinetochores on each side of the chromosome, so that they can later pull them apart. The chromosomes align themselves with the spindle apparatus, which is spread around the cell like the vertical lines on a globe. The soon to be divided chromosomes are symmetrically positioned on the metaphase plate, which is essentially the equator of the parent cell. At the end of metaphase, each chromosome has microtubules connected to both of its halves, and they are lined up in a straight line along the equator of the cell.



Anaphase

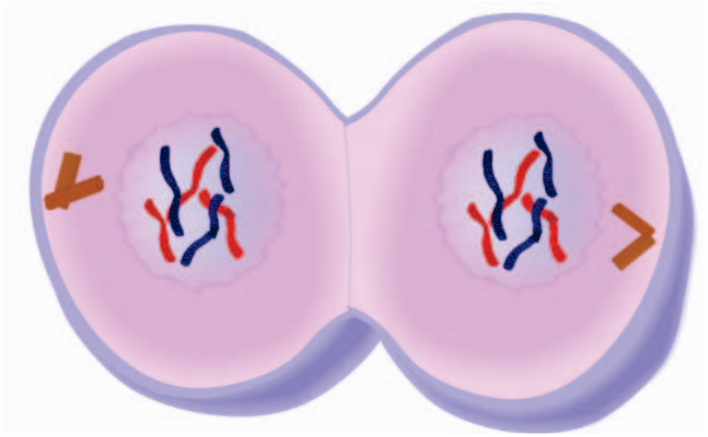
Once the chromosomes get lined up properly, the spindle apparatus immediately pulls the two identical DNA halves apart from one another and moves them to opposite sides of the cell. These two sets of chromosomes will develop into the nuclei of two daughter cells which are perfectly identical to each other and the parent cell.



Telophase

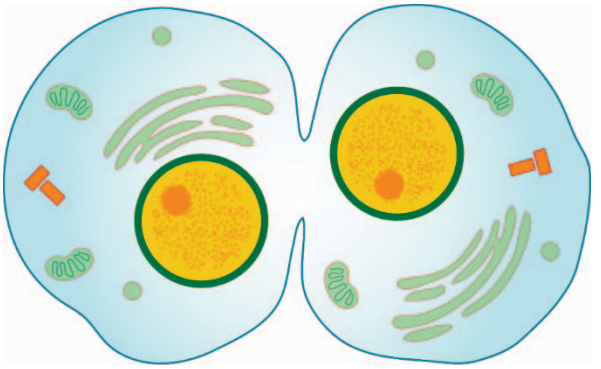
After the chromosomes arrive at the ends of the cell, they start to uncoil and spread out again, as they were before they formed into Xs. This is basically the opposite of the beginning of prophase. While this happens, the spindle apparatus is broken down. After that, the **nuclear membrane**, which encases the nucleus, forms again around the chromosomes, unless it never dissolved in the first place, as in closed mitosis. Although this is the last phase, cell division is not complete until cytokinesis happens.

Nuclear Membrane is made up of two lipid bilayer membranes which surrounds the nucleus, and in eukaryotic cells it encases the genetic material.



Cytokinesis

Cytokinesis is the next stage of cellular development, and is similar to mitosis, except it involves the other parts of the cell instead of the nucleus. During this phase, the equator-like metaphase plate of the cell pinches together, separating the cell into two new cells. Once this is completed, there are two functioning, identical cells.



Purpose

One of the main purposes of this process is the natural growth of the parent organism. It's also done to replace cells that are worn out, damaged, or just at the end of their natural lifespan. For instance, a person continually sloughs of dead skin cells, so the body has to divide cells to make new ones. Some animals also use this process to regenerate parts of themselves, like lizards who can regrow their tails after losing them. Additionally, some animals undergo this process as part of asexual reproduction.

Problems

Problems with mitosis are devastating for cells, and can result in their death. Even if the cell doesn't die, the chromosomes can be damaged or altered, which can lead to genetic disorders; Down syndrome, for instance, is caused by a chromosomal problem that's connected to mitosis. Additionally, damage to the chromosomes or problems with the timing of how the cell divides can lead to growths and sometimes cancer. This can also happen if the chromosome isn't pulled apart properly.

Meiosis

Mitosis and meiosis are both means of cellular division, but they differ in a few key ways. First, meiosis only happens in specific types of reproductive cells called gametes — in humans, eggs and sperm — and spores. Also, in meiosis, the DNA from each contributing cell gets mixed up, with little pieces of DNA from different cells forming parts of the X. This is different from mitosis where the two halves of the Xs are identical. Also, meiosis ends up with four cells that are completely genetically unique, whereas in mitosis, the end result is two completely identical cells.

Types of Mitosis

(i) Anastral mitosis: It is found in plants in which spindle has no aster.

(ii) Amphiastral mitosis: It is found in animals in which spindle has two asters, one at each pole of the spindle. Spindle is barrel-like.

(iii) Intranuclear or Promitosis: In this nuclear membrane is not lost and

spindle is formed inside the nuclear membrane e.g. Protozoans (Amoeba) and yeast. It is so as centriole is present within the nucleus.

(iv) Extranuclear or Eumitosis: In this nuclear membrane is lost and spindle is formed outside nuclear membrane e.g. in plants and animals.

(v) Endomitosis: Chromosomes and their DNA duplicate but fail to separate which lead to polyploidy e.g. in liver of man, both diploid (2N) and polyploid cells (4N) have been reported. It is also called endoduplication and endopolyploidy.

(vi) Dinomitosis: In which nuclear envelope persists and microtubular spindle is not formed. During movement the chromosomes are attached with nuclear membrane.

Mitotic Abnormalities

Various mitotic abnormalities are formed as a result of placing the cells in unfavourable conditions during different phases.

Physical agents, such as temperature, radiations and chemicals like narcotics and enzyme inhibitors, can easily produce mitotic deformities.

Cytasteral mitosis

Wilson noticed formation of numerous asters in cytoplasm of unfertilized eggs called cytasters by placing them in hypertonic sea water.

In the eggs of marine invertebrates, production of many small cytoplasmic asters is of frequent occurrence. Intermixing of nucleoplasm and hyaloplasm seems to be necessary condition for aster formation. Costello showed that asters can be produced only if germinal vesicle has broken down.

Multipolar and caternar mitosis:

Mitotic divisions showing several spindles and Centrosome are common in many Protozoa and fish eggs. Multipolarity is usually caused by uneven division of Centrosome as well as irregular distribution of chromatids on different spindles. It results in the formation of cells as aneuploidy (uneven chromosome number).

Caternar mitoses have been described by Dalcq and Simon in amphibian eggs. During this, dividing cell forms a large number of asters without spindles or chromosomes to link them. It suggests that asters are capable of dividing in the absence of spindle or nucleus in an autonomous way.

Achrosomal mitosis

Sometimes, mitosis is induced by various agents without use of nucleus or chromosomes. Briggs in fertilized eggs with heavily x-rayed sperms, removed the egg's maturation spindle by pricking and sucking. In this way, he obtained very nice

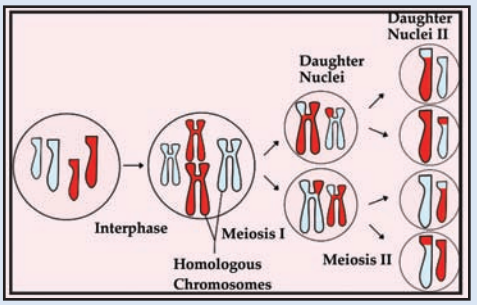
non-nucleate blastulae. Similarly, Stauffer obtained normal non-nucleated Axolotl blastula.

Anastral mitosis

In plant cells and many oocytes, generally asters around the Centrosome are absent, thus producing Anastral mitosis. Bataillon and Tchou Su described in details the Anastral mitosis in amphibian inter—specific hybrids.

C-Mitosis (abnormal spindle formation)

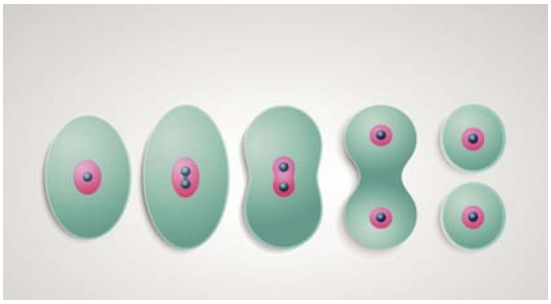
Brachet has described that colchicine inhibits mitosis by disorganizing spindle formation. The results are the formation of polyploid cells, and reduplication of chromosomes. Besides, during mitosis, sometimes cell nucleus is either deformed into a single spheroid compact mass (pycnosis) or becomes broken down (karyorrhexis).



CORE CONCEPT: MEIOSIS

Meiosis is a specialized type of cell division that reduces the chromosome number by half, creating four haploid cells, each genetically distinct from the parent cell that gave rise to them. This process occurs in all sexually reproducing single-celled and multicellular eukaryotes, including animals, plants, and fungi.

As an important component in the process of sexual reproduction, meiosis is the means by which one cell divides into four different cells. Here are some details on the process of meiosis and why meiosis is necessary in order to ensure reproduction. Meiosis takes its name from the Greek word meioun, which is translated to mean to make something smaller. In effect, that is what happens when meiosis occurs. A diploid eukaryotic cell will begin a process of dividing into four more or less equal portions, referred to as haploid cells. Generally referred to as gametes, the number of chromosomes contained in each of the gametes created from diploids is reduced from the parent diploid cell. Chromosomes contain the basic DNA chain that will determine the physical characteristics of the child in the event that a pregnancy takes place.

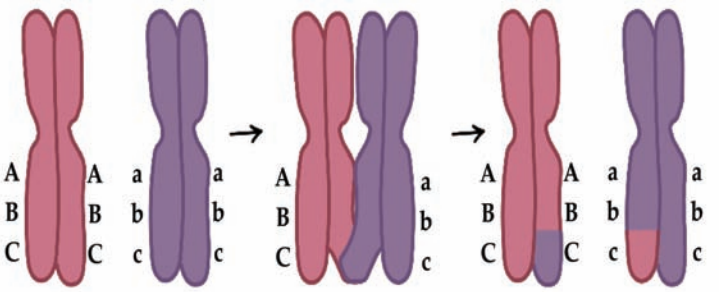


During the process of meiosis, the genome of the diploids begins to undergo two distinct rounds of division that ultimately result in the four haploid cells. Each of the haploid cells or gametes will contain one fully complete strand of chromosomes, which equates to half of the chromosomes supplied by the previous division. The gametes are capable of meshing or fusing with other opposite gender haploid cells during the process of fertilization. If the fertilization is successful, this leads to the creation of a **zygote**. This is important to note, as the union of these two different gender diploids will contained the DNA strands of both of the haploid cells, resulting in a unique physiological makeup for the new life.

Zygote is a eukaryotic cell formed by a fertilization event between two gametes.

The need for a combination of two haploid cells to form the zygote is one of the characteristics that set the process of meiosis apart from other means of reproduction, such as mitosis. While both meiosis and mitosis will rely on some of the same mechanisms to aid in chromatin distribution, it is important to remember that mitosis occurs as a form of asexual reproduction, meaning there is no need for a combination of opposite gender gametes. Meiosis sets the stage for sexual reproduction that results in genetic variations that are based on the unique composition of the two opposite gametes, rather than the creation of new life from a single source.

While there is ongoing research to determine if there is some way to manipulate the process of meiosis, modern science still has a long way to go. At present there are no effective methods to potentially prevent the formation of birth defects, or even create the possibility of altering the basic DNA. This means that modern science does not currently offer a means to produce a desired physical aspect. Until much more is understood about the basic cell structures that divide and change during meiosis, how chromatin is produced, what can create an unhealthy gamete and other related matters, science will have to be content with observing the process as it now stands.



Meiosis and Sexual Reproduction

The ability to reproduce *in kind* is a basic characteristic of all living things. In kind means that the offspring of any organism closely resemble their parent or parents. Sexual reproduction requires fertilization: the union of two cells from two individual organisms. Haploid cells contain one set of chromosomes. Cells containing two sets of chromosomes are called diploid. The number of sets of chromosomes in a cell is called its ploidy level. If the reproductive cycle is to continue, then the diploid cell must somehow reduce its number of chromosome sets before fertilization can occur again or there will be a continual doubling in the number of chromosome sets in every generation. Therefore, sexual reproduction includes a nuclear division that reduces the number of chromosome sets.

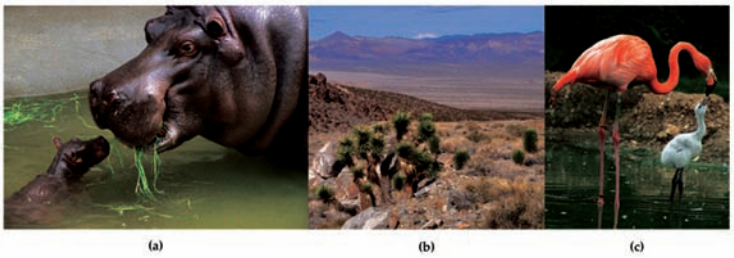


Figure 2: Offspring Closely Resemble Their Parents: *In kind* means that the offspring of any organism closely resemble their parent or parents. The hippopotamus gives birth to hippopotamus calves (a). Joshua trees produce seeds from which Joshua tree seedlings emerge (b). Adult flamingos lay eggs that hatch into flamingo chicks (c).

Sexual reproduction is the production of haploid cells (gametes) and the fusion (fertilization) of two gametes to form a single, unique diploid cell called a zygote. All animals and most plants produce these gametes, or eggs and sperm. In most plants and animals, through tens of rounds of mitotic cell division, this diploid cell will develop into an adult organism.

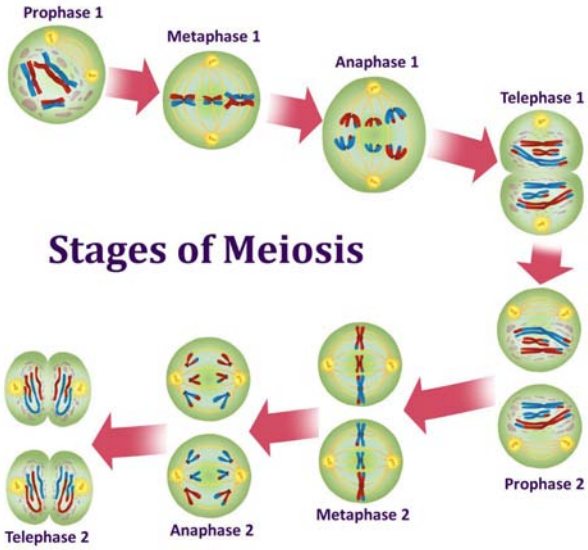
Haploid cells that are part of the sexual reproductive cycle are produced by a type of cell division called meiosis. Meiosis employs many of the same mechanisms as mitosis. However, the starting nucleus is always diploid and the nuclei that result at the end of a meiotic cell division are haploid, so the resulting cells have half the chromosomes as the original. To achieve this reduction in chromosomes, meiosis consists of one round of chromosome duplication and two rounds of nuclear division.

Because the events that occur during each of the division stages are analogous to the events of mitosis, the same stage names are assigned. However, because there are two rounds of division, the major process and the stages are designated with a "I" or a "II." Thus, meiosis I is the first round of meiotic division and consists of prophase I, prometaphase I, and so on. Meiosis II, the second round of meiotic division, includes prophase II, prometaphase II, and so on.

Meiosis I

Meiosis is preceded by an interphase consisting of three stages. The G_1 phase (also called the first gap phase) initiates this stage and is focused on cell growth. The S phase is next, during which the DNA of the chromosomes is replicated. This replication produces two identical copies, called sister chromatids, that are held together at the centromere by **cohesin** proteins. The centrosomes, which are the structures that organize the microtubules of the meiotic spindle, also replicate. Finally, during the G_2 phase (also called the second gap phase), the cell undergoes the final preparations for meiosis.

Cohesin is a protein complex that regulates the separation of sister chromatids during cell division, either mitosis or meiosis.



Prophase I

During prophase I, chromosomes condense and become visible inside the nucleus. As the nuclear envelope begins to break down, homologous chromosomes move closer together. The synaptonemal complex, a lattice of proteins between the homologous chromosomes, forms at specific locations, spreading to cover the entire length of the chromosomes. The

tight pairing of the homologous chromosomes is called synapsis. In synapsis, the genes on the chromatids of the homologous chromosomes are aligned with each other. The synaptonemal complex also supports the exchange of chromosomal segments between non-sister homologous chromatids in a process called crossing over. The crossover events are the first source of genetic variation produced by meiosis. A single crossover event between homologous non-sister chromatids leads to an exchange of DNA between chromosomes. Following crossover, the synaptonemal complex breaks down and the cohesin connection between homologous pairs is also removed. At the end of prophase, I, the pairs are held together only at the chiasmata; they are called tetrads because the four sister chromatids of each pair of homologous chromosomes are now visible.

Chiasmata are structures that physically link the homologous chromosomes after crossover and play an essential role in meiotic chromosome segregation.

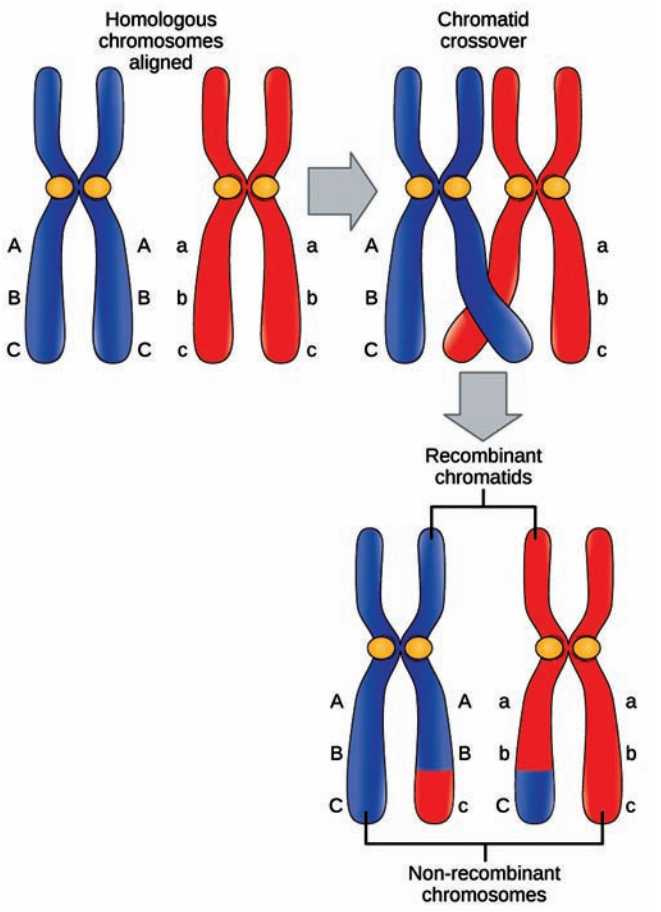


Figure 3: Crossover between homologous chromosomes: Crossover occurs between non-sister chromatids of homologous chromosomes. The result is an exchange of genetic material between homologous chromosomes.

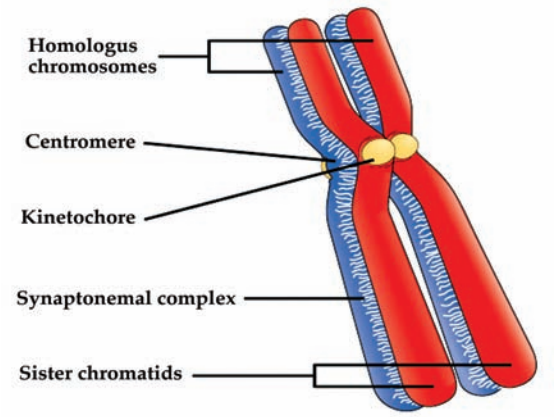


Figure 4: Synapsis holds pairs of homologous chromosomes together: Early in prophase I, homologous chromosomes come together to form a synapse. The chromosomes are bound tightly together and in perfect alignment by a protein lattice called a synaptonemal complex and by cohesin proteins at the centromere.

Prometaphase I

The key event in prometaphase I is the formation of the spindle fibre apparatus where spindle fiber microtubules attach to the kinetochore proteins at the centromeres. Microtubules grow from centrosomes placed at opposite poles of the cell. The **microtubules** move toward the middle of the cell and attach to one of the two fused homologous chromosomes at the kinetochores. At the end of prometaphase I, each tetrad is attached to microtubules from both poles, with one homologous chromosome facing each pole. In addition, the nuclear membrane has broken down entirely.

Microtubules are a component of the cytoskeleton, found throughout the cytoplasm.

Metaphase I

During metaphase I, the tetrads move to the metaphase plate with kinetochores facing opposite poles. The homologous pairs orient themselves randomly at the equator. This event is the second mechanism that introduces variation into the gametes or spores. In each cell that undergoes meiosis, the arrangement of the tetrads is different. The number of variations is dependent on the number of chromosomes making up a set. There are two possibilities for orientation at the metaphase plate. The possible number of alignments, therefore, equals 2n, where n is the number of chromosomes per set. Given these two mechanisms, it is highly unlikely that any two haploid cells resulting from meiosis will have the same genetic composition.

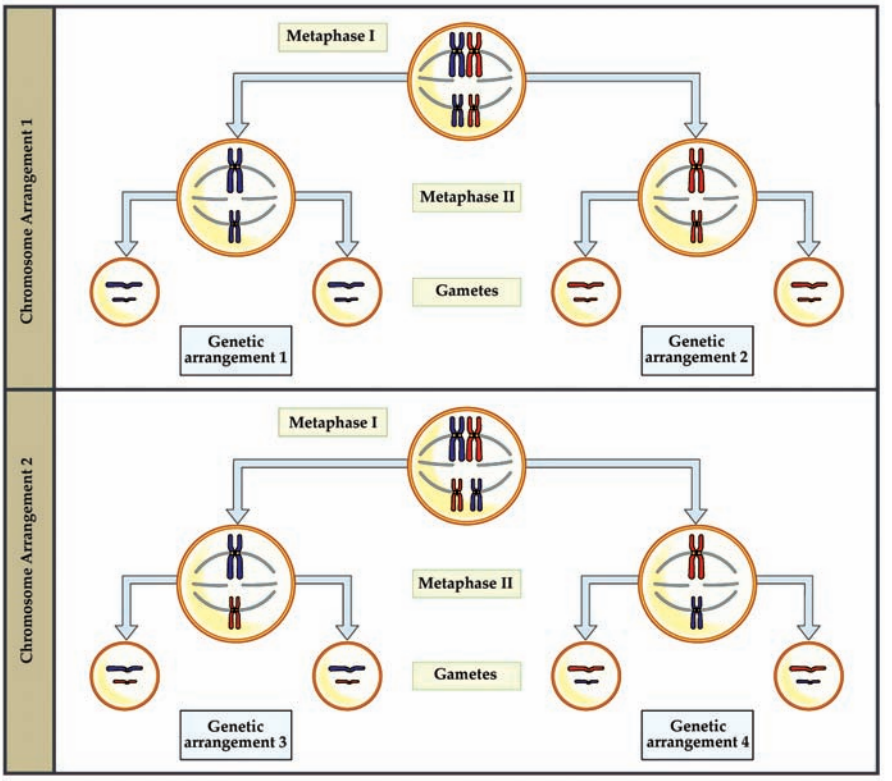
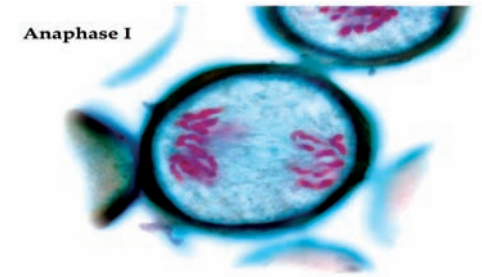


Figure 5: Meiosis I ensures unique gametes: Random, independent assortment during metaphase I can be demonstrated by considering a cell with a set of two chromosomes (n = 2). In this case, there are two possible arrangements at the equatorial plane in metaphase I. The total possible number of different gametes is 2n, where n equals the number of chromosomes in a set. In this example, there are four possible genetic combinations for the gametes. With n = 23 in human cells, there are over 8 million possible combinations of paternal and maternal chromosomes.

Anaphase I

Chromosome is a DNA molecule with part or all of the genetic material (genome) of an organism.

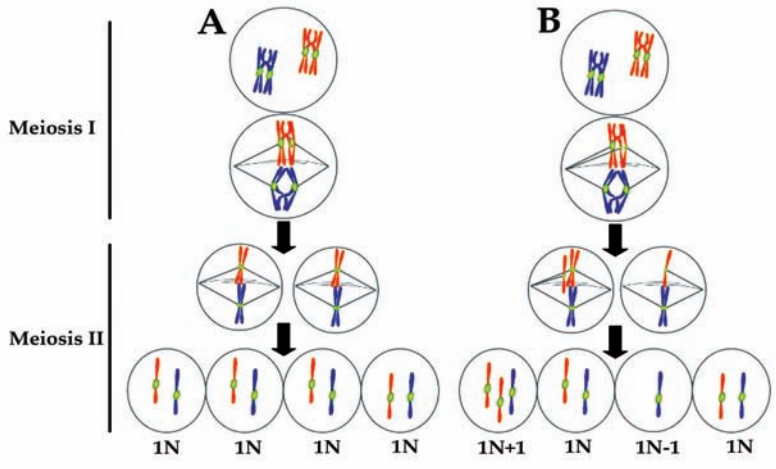
In anaphase I, the microtubules pull the attached **chromosomes** apart. The sister chromatids remain tightly bound together at the centromere. The chiasmata are broken in anaphase I as the microtubules attached to the fused kinetochores pull the homologous chromosomes apart.



Telophase I and Cytokinesis

In telophase I, the separated chromosomes arrive at opposite poles. In some organisms, the chromosomes decondense and nuclear envelopes form around the chromatids in telophase I. Then cytokinesis, the physical separation of the cytoplasmic components into two daughter cells, occurs without reformation of the nuclei. In nearly all species of animals and some fungi, cytokinesis separates the cell contents via a cleavage furrow (constriction of the actin ring that leads to cytoplasmic division). In plants, a cell plate is formed during cell **cytokinesis** by Golgi vesicles fusing at the metaphase plate. This cell plate will ultimately lead to the formation of cell walls that separate the two daughter cells.

Cytokinesis is the part of the cell division process during which the cytoplasm of a single eukaryotic cell divides into two daughter cells.

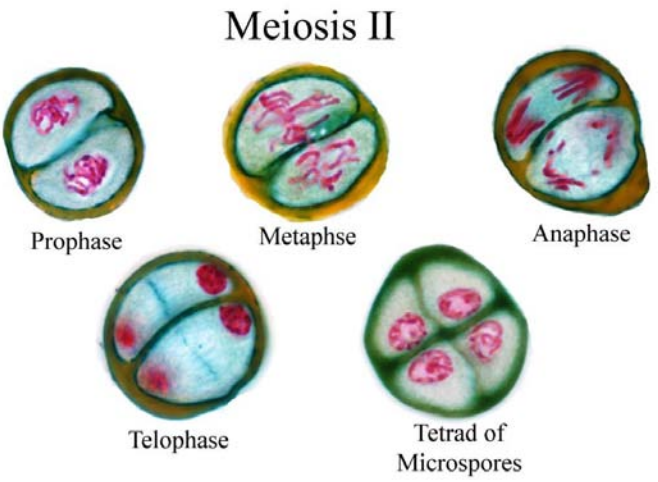


Two haploid cells are the end result of the first meiotic division. The cells are haploid because at each pole there is just one of each pair of the homologous chromosomes. Therefore, only one full set of the chromosomes is present. Although there is only one chromosome set, each homolog still consists of two sister chromatids.

Meiosis II

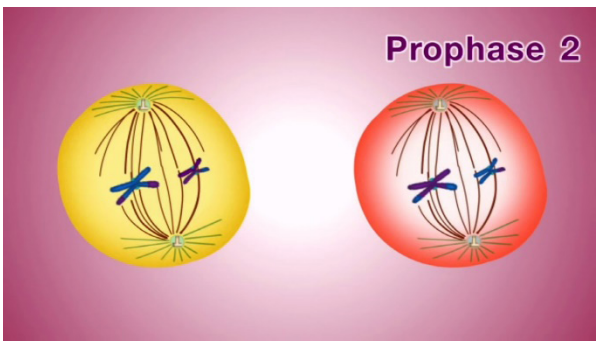
Meiosis II initiates immediately after cytokinesis, usually before the chromosomes have fully decondensed. In contrast to meiosis I, meiosis II resembles a normal mitosis. In some species, cells enter a brief interphase, or interkinesis, before entering meiosis II. Interkinesis lacks an S phase, so chromosomes are not duplicated. The two cells produced in meiosis I go through the events of meiosis II together. During meiosis II, the sister chromatids within the two daughter cells

separate, forming four new haploid gametes. The mechanics of meiosis II is similar to mitosis, except that each dividing cell has only one set of homologous chromosomes.



Prophase II

If the chromosomes decondensed in telophase I, they condense again. If nuclear envelopes were formed, they fragment into vesicles. The centrosomes that were duplicated during interphase I move away from each other toward opposite poles and new spindles are formed.

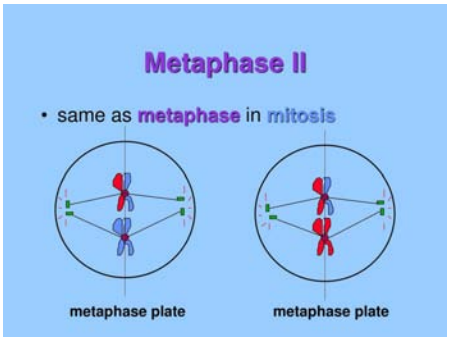


Prometaphase II

The nuclear envelopes are completely broken down and the spindle is fully formed. Each sister chromatid forms an individual kinetochore that attaches to microtubules from opposite poles.

Metaphase II

The sister chromatids are maximally condensed and aligned at the equator of the cell.



Anaphase II

The sister chromatids are pulled apart by the kinetochore microtubules and move toward opposite poles. Non-kinetochore microtubules elongate the cell.

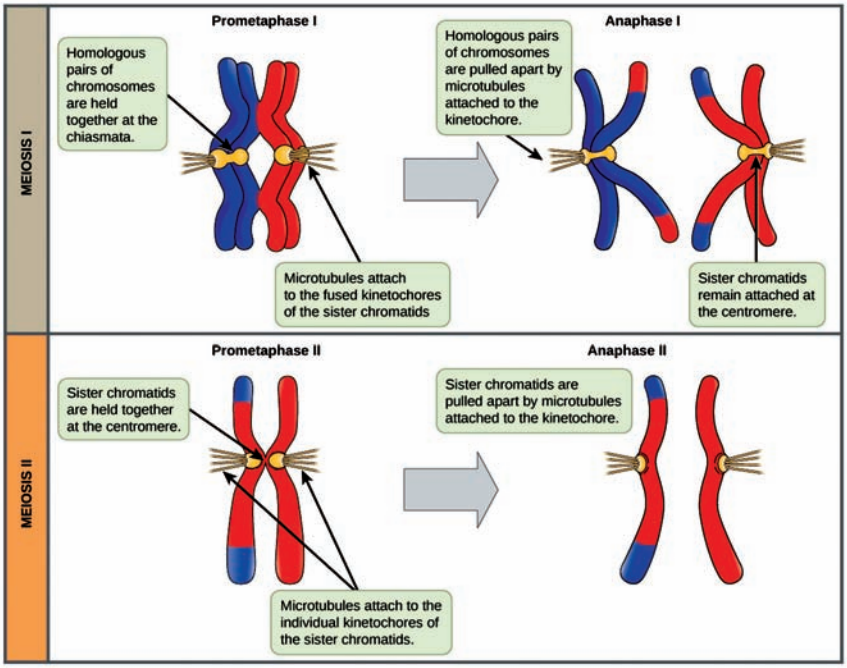


Figure 6: Meiosis I vs. Meiosis II: The process of chromosome alignment differs between meiosis I and meiosis II. In prometaphase I, microtubules attach to the fused kinetochores of homologous chromosomes, and the homologous chromosomes are arranged at the midpoint of the cell in metaphase I. In anaphase I, the homologous chromosomes are separated. In prometaphase II, microtubules attach to the kinetochores of sister chromatids, and the sister chromatids are arranged at the midpoint of the cells in metaphase II. In anaphase II, the sister chromatids are separated.

Telophase II and Cytokinesis

The chromosomes arrive at opposite poles and begin to decondense. Nuclear envelopes form around the chromosomes. Cytokinesis separates the two cells into four unique haploid cells. At this point, the newly-formed nuclei are both haploid. The

cells produced are genetically unique because of the random assortment of paternal and maternal homologs and because of the recombining of maternal and paternal segments of chromosomes (with their sets of genes) that occurs during crossover.

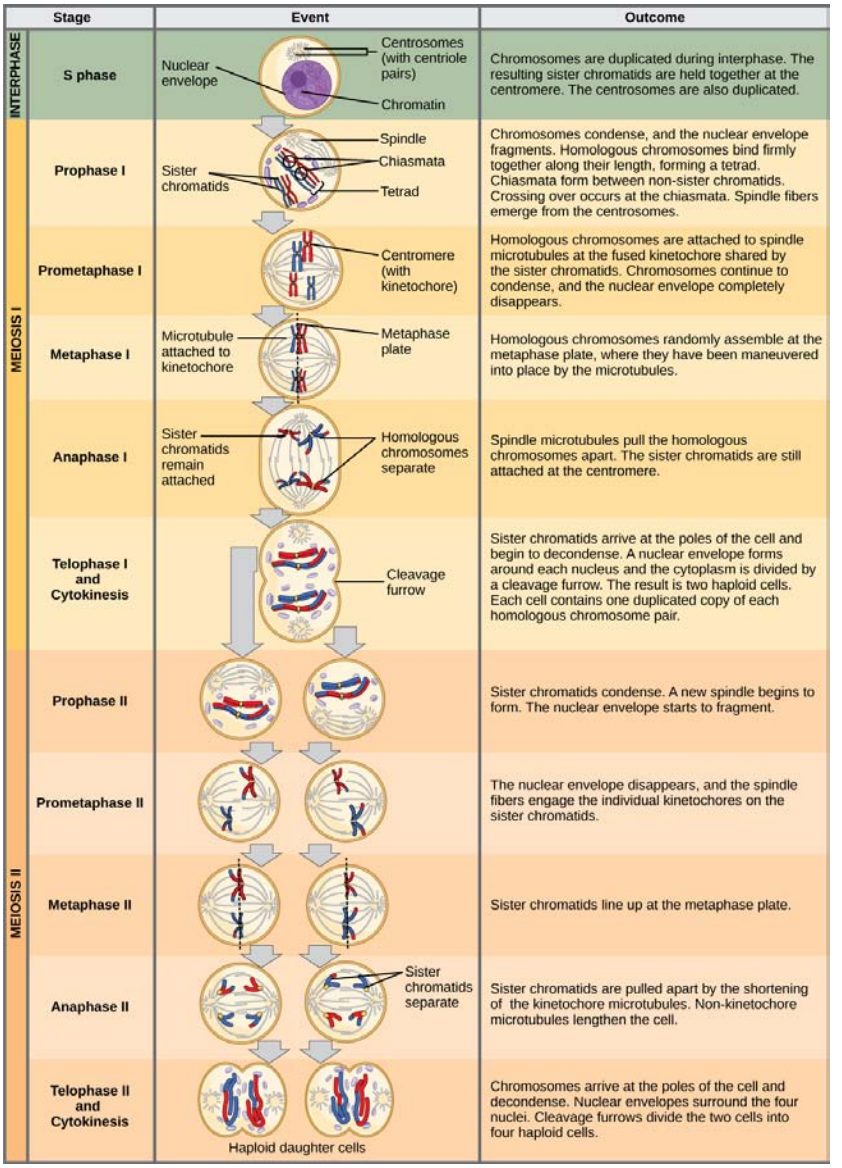


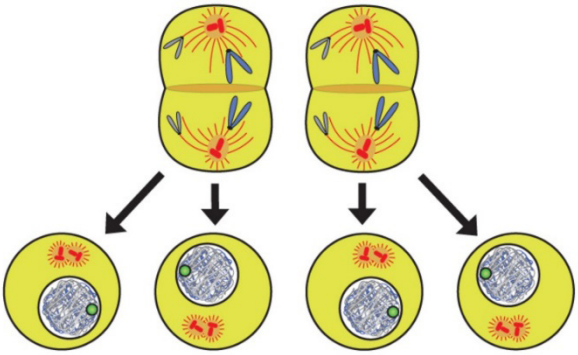
Figure 7: Complete Stages of Meiosis: An animal cell with a diploid number of four (2n = 4) proceeds through the stages of meiosis to form four haploid daughter cells.

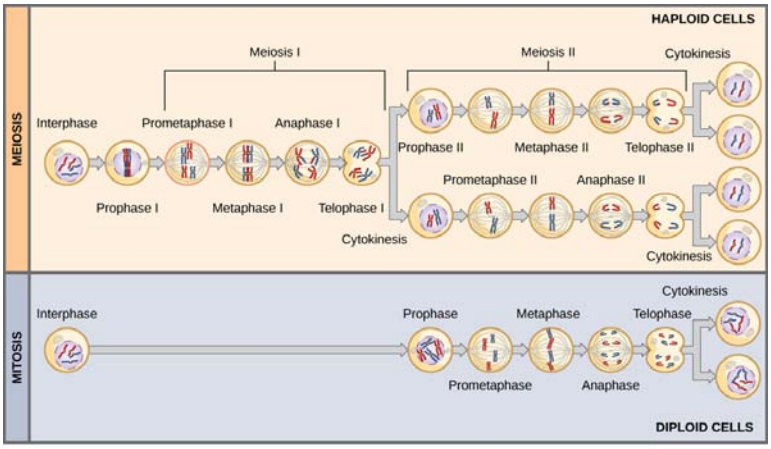
Comparing Meiosis and Mitosis

Mitosis and meiosis are both forms of division of the nucleus in eukaryotic cells. They share some similarities, but also exhibit distinct differences that lead to very different outcomes. The purpose of mitosis is cell regeneration, growth, and asexual reproduction, while the purpose of meiosis is the production of gametes for sexual reproduction. Mitosis is a single nuclear division that results in two nuclei that are usually

partitioned into two new daughter cells. The nuclei resulting from a mitotic division are genetically identical to the original nucleus. They have the same number of sets of chromosomes, one set in the case of haploid cells and two sets in the case of diploid cells. In most plants and all animal species, it is typically diploid cells that undergo mitosis to form new **diploid cells**. In contrast, meiosis consists of two nuclear divisions resulting in four nuclei that are usually partitioned into four new haploid daughter cells. The nuclei resulting from meiosis are not genetically identical and they contain one chromosome set only. This is half the number of chromosome sets in the original cell, which is diploid.

Diploid cells have two homologous copies of each chromosome, usually one from the mother and one from the father.





						OUTCOME
PROCESS	DNA synthesis	Synapsis of homologous chromosomes	Crossover	Homologous chromosomes line up at metaphase plate	Sister chromatids line up at metaphase plate	Number and genetic composition of daughter cells
MEIOSIS	Occurs in S phase of interphase	During prophase I	During prophase I	During metaphase I	During metaphase II	Four haploid cells at the end of meiosis II
MITOSIS	Occurs in S phase of interphase	Does not occur in mitosis	Does not occur in mitosis	Does not occur in mitosis	During metaphase	Two diploid cells at the end of mitosis

Figure 8: Comparing Meiosis and Mitosis: Meiosis and mitosis are both preceded by one round of DNA replication; however, meiosis

includes two nuclear divisions. The four daughter cells resulting from meiosis are haploid and genetically distinct. The daughter cells resulting from mitosis are diploid and identical to the parent cell.

The main differences between mitosis and meiosis occur in meiosis I. In meiosis I, the homologous chromosome pairs become associated with each other and are bound together with the synaptonemal complex. Chiasmata develop and crossover occurs between homologous chromosomes, which then line up along the metaphase plate in tetrads with kinetochore fibers from opposite spindle poles attached to each kinetochore of a homolog in a tetrad. All of these events occur only in meiosis I.

When the tetrad is broken up and the homologous chromosomes move to opposite poles, the ploidy level is reduced from two to one. For this reason, meiosis I is referred to as a reduction division. There is no such reduction in ploidy level during mitosis.

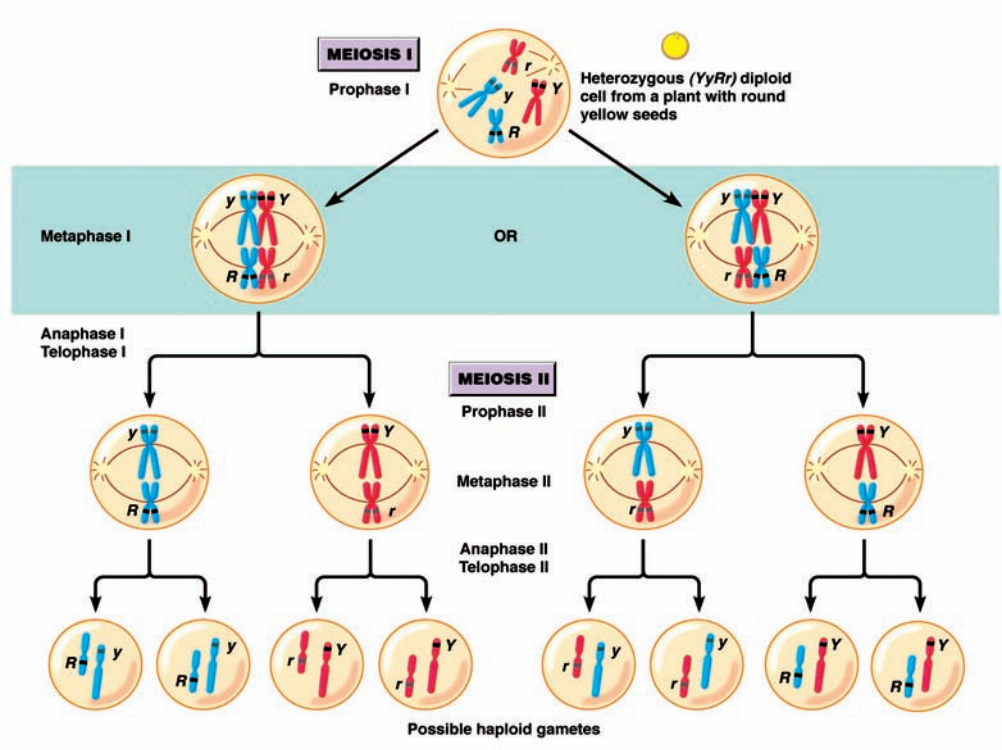
Meiosis II is much more similar to a mitotic division. In this case, the duplicated chromosomes (only one set, as the homologous pairs have now been separated into two different cells) line up on the metaphase plate with divided kinetochores attached to kinetochore fibres from opposite poles. During anaphase II and mitotic anaphase, the kinetochores divide and sister chromatids, now referred to as chromosomes, are pulled to opposite poles. The two daughter cells of mitosis, however, are identical, unlike the daughter cells produced by meiosis. They are different because there has been at least one crossover per chromosome. Meiosis II is not a reduction division because, although there are fewer copies of the genome in the resulting cells, there is still one set of chromosomes, as there was at the end of meiosis I. Meiosis II is, therefore, referred to as equatorial division.

Sexual Reproduction

Sexual reproduction was an early evolutionary innovation after the appearance of eukaryotic cells. During sexual reproduction, the genetic material of two individuals is combined to produce genetically-diverse offspring that differ from their parents. The fact that most eukaryotes reproduce sexually is evidence of its evolutionary success. In many animals, it is actually the only mode of reproduction. The genetic diversity of sexually-produced offspring is thought to give species a better chance of surviving in an unpredictable or changing environment.

Scientists recognize some real disadvantages to sexual reproduction. On the surface, creating offspring that are genetic clones of the parent appears to be a better system. If the parent organism is successfully occupying a habitat, offspring with the same traits would be similarly successful. Species that reproduce sexually must maintain two different types of individuals, males and females, which can limit the ability to colonize new habitats as both sexes must be present. Therefore, there is an obvious benefit to an organism that can produce offspring whenever circumstances are favorable by asexual budding, fragmentation, or asexual eggs. These methods of asexual reproduction do not require another organism of the opposite sex. Indeed, some organisms that lead a solitary lifestyle have retained the ability to reproduce asexually.

In addition, in asexual populations, every individual is capable of reproduction. In sexual populations, the males are not producing the offspring themselves. In theory, an asexual population could grow twice as fast.



Nevertheless, multicellular organisms that exclusively depend on asexual reproduction are exceedingly rare. Why is sexuality (and meiosis) so common? This is one of the important unanswered questions in biology and has been the focus of much research beginning in the latter half of the twentieth century. There are several possible explanations, one of which is that the variation that sexual reproduction creates among offspring is very important to the survival and reproduction of the population. Thus, on average, a sexually-reproducing population will leave more descendants than an otherwise similar asexually-reproducing population. The only source of variation in asexual organisms is mutation. This is the ultimate source of variation in sexual organisms, but, in addition, those different mutations are continually reshuffled from one generation to the next when different parents combine their unique genomes and the genes are mixed into different combinations by the process of meiosis. Meiosis is the division of the contents of the nucleus, dividing the chromosomes among gametes.

The process of meiosis produces unique reproductive cells called gametes, which have half the number of chromosomes as the parent cell. Fertilization, the fusion of haploid gametes from two individuals, restores the diploid condition. Thus, sexually-reproducing organisms alternate between haploid and diploid stages. However, the ways in which reproductive cells are produced and the timing between meiosis and fertilization vary greatly. There are three main categories of sexual life cycles: diploid-dominant, demonstrated by most animals; haploid-dominant, demonstrated by all

fungi and some algae; and the alternation of generations, demonstrated by plants and some algae.

The Red Queen Hypothesis

It is not in dispute that sexual reproduction provides evolutionary advantages to organisms that employ this mechanism to produce offspring. But why, even in the face of fairly stable conditions, does sexual reproduction persist when it is more difficult and costly for individual organisms? Variation is the outcome of sexual reproduction, but why are ongoing variations necessary? Possible answers to these questions are explained in the Red Queen hypothesis, first proposed by Leigh Van Valen in 1973.

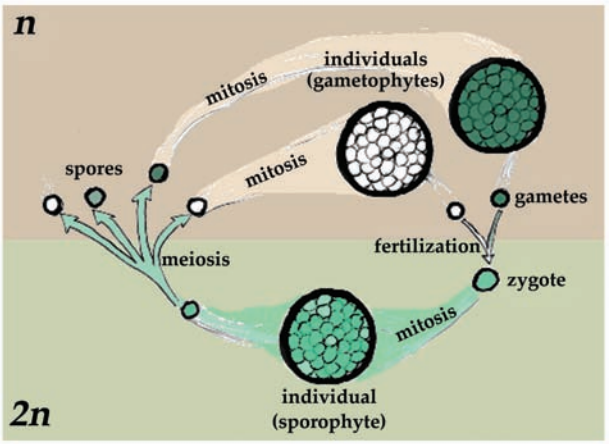


Figure 9: The Sexual Life Cycle: In animals, sexually-reproducing adults form haploid gametes from diploid germ cells. Fusion of the gametes gives rise to a fertilized egg cell, or zygote. The zygote will undergo multiple rounds of mitosis to produce a multicellular offspring.

All species co-evolve with other organisms; for example, predators evolve with their prey and parasites evolve with their hosts. Each tiny advantage gained by favorable variation gives a species an edge over close competitors, predators, parasites, or even prey. The only method that will allow a co-evolving species to maintain its own share of the resources is to also continually improve its fitness. As one species gains an advantage, this increases selection on the other species; they must also develop an advantage or they will be out-competed. No single species progresses too far ahead because genetic variation among the progeny of sexual reproduction provides all species with a mechanism to improve rapidly. Species that cannot keep up become extinct. The Red Queen's catchphrase was, "It takes all the running you can do to stay in the same place." This is an apt description of co-evolution between competing species.

Life Cycles of Sexually Reproducing Organisms

In sexual reproduction, the genetic material of two individuals is combined to produce genetically diverse offspring that differ from their parents. Fertilization and meiosis alternate in sexual life cycles. What happens between these two events depends upon the

organism? The process of meiosis, the division of the contents of the nucleus that divides the chromosomes among gametes, reduces the chromosome number by half, while fertilization, the joining of two haploid gametes, restores the diploid condition. There are three main categories of life cycles in eukaryotic organisms: diploid-dominant, haploid-dominant, and alternation of generations.

In meiosis, DNA replication is followed by two rounds of cell division to produce four daughter cells, each with half the number of chromosomes as the original parent cell. The two meiotic divisions are known as Meiosis I and Meiosis II. Before meiosis begins, during S phase of the cell cycle, the DNA of each chromosome is replicated so that it consists of two identical sister chromatids, which remain held together through sister chromatid cohesion.

Diploid-Dominant Life Cycle

In the diploid-dominant life cycle, the multicellular diploid stage is the most obvious life stage, as occurs with most animals, including humans. Nearly all animals employ a diploid-dominant life cycle strategy in which the only haploid cells produced by the organism are the gametes. Early in the development of the embryo, specialized diploid cells, called germ cells, are produced within the gonads (e.g. testes and ovaries). Germ cells are capable of mitosis to perpetuate the cell line and meiosis to produce gametes. Once the haploid gametes are formed, they lose the ability to divide again. There is no multicellular haploid life stage. Fertilization occurs with the fusion of two gametes, usually from different individuals, restoring the diploid state.

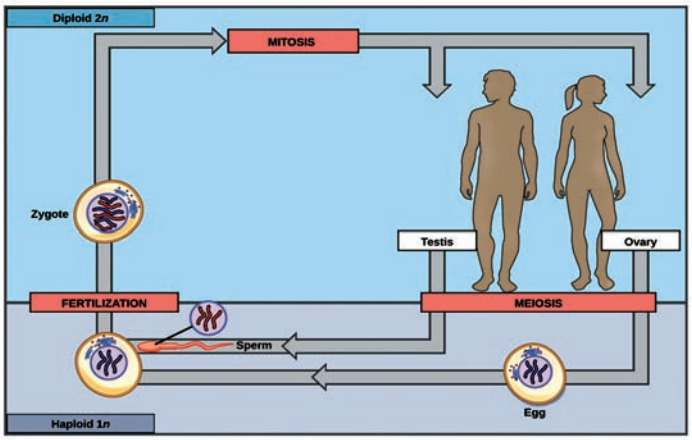


Figure 10: Diploid-Dominant Life Cycle: In animals, sexually-reproducing adults form haploid gametes from diploid germ cells. Fusion of the gametes gives rise to a fertilized egg cell, or zygote. The zygote will undergo multiple rounds of mitosis to produce a multicellular offspring. The germ cells are generated early in the development of the zygote.

Haploid-Dominant Life Cycle

Within haploid-dominant life cycles, the multicellular haploid stage is the most obvious life stage. Most fungi and algae employ a life cycle type in which the “body” of the

organism, the ecologically important part of the life cycle, is haploid. The haploid cells that make up the tissues of the dominant multicellular stage are formed by mitosis. During sexual reproduction, specialized haploid cells from two individuals, designated the (+) and (−) mating types, join to form a diploid zygote. The zygote immediately undergoes meiosis to form four haploid cells called spores. Although haploid like the "parents," these spores contain a new genetic combination from two parents. The spores can remain dormant for various time periods. Eventually, when conditions are conducive, the spores form multicellular haploid structures by many rounds of mitosis.

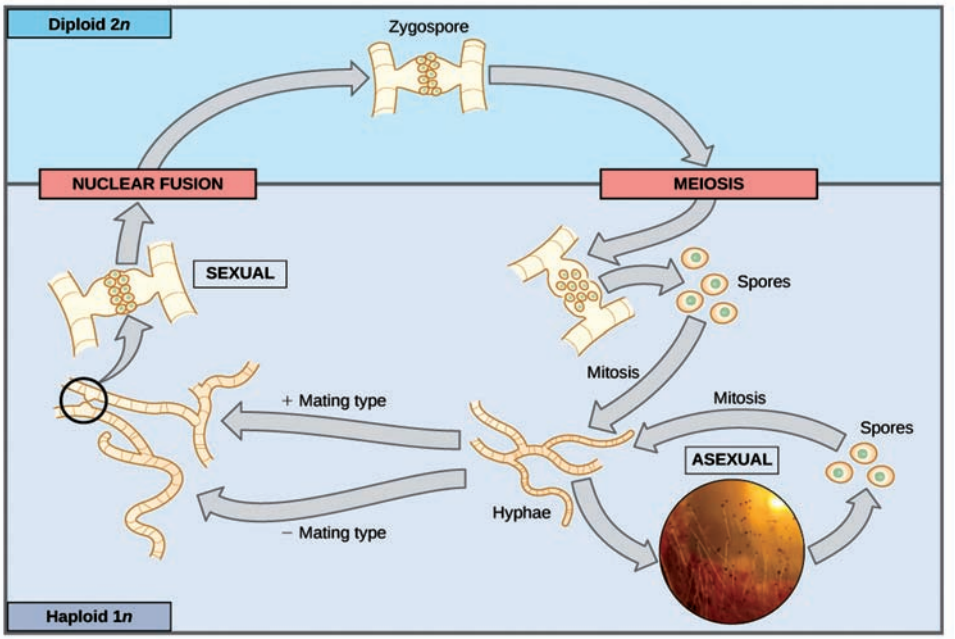


Figure 11: Haploid-Dominant Life Cycle: Fungi, such as black bread mold (Rhizopus nigricans), have haploid-dominant life cycles. The haploid multicellular stage produces specialized haploid cells by mitosis that fuse to form a diploid zygote. The zygote undergoes meiosis to produce haploid spores. Each spore gives rise to a multicellular haploid organism by mitosis.

Alternation of Generations

The third life-cycle type, employed by some algae and all plants, is a blend of the haploid-dominant and diploid-dominant extremes. Species with alternation of generations have both haploid and diploid multicellular organisms as part of their life cycle. The haploid multicellular plants are called gametophytes because they produce gametes from specialized cells. Meiosis is not directly involved in the production of gametes because the organism that produces the gametes is already a haploid. Fertilization between the gametes forms a diploid zygote. The zygote will undergo many rounds of mitosis and give rise to a diploid multicellular plant called a sporophyte. Specialized cells of the sporophyte will undergo meiosis and produce haploid spores. The spores will subsequently develop into the gametophytes.

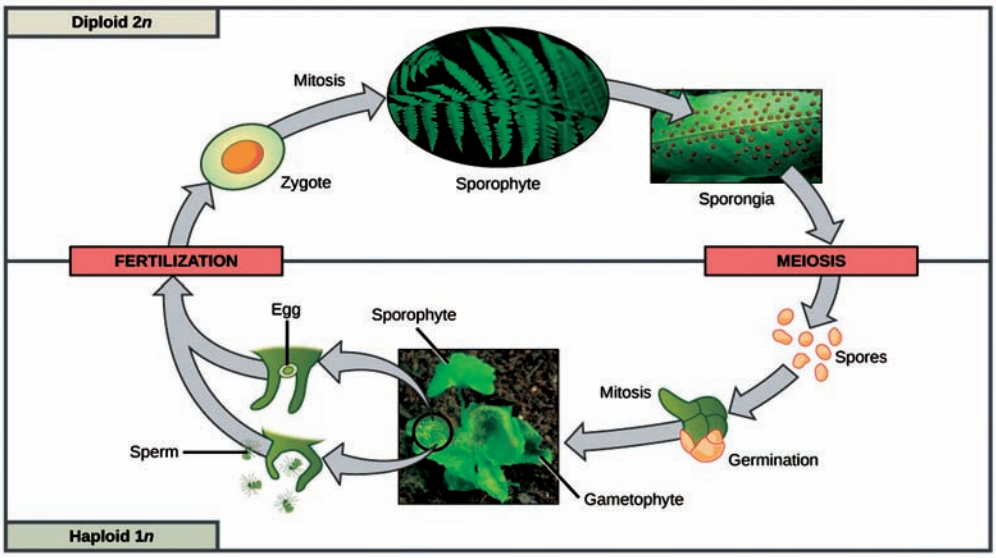


Figure 12: Alternation of Generations: Plants have a life cycle that alternates between a multicellular haploid organism and a multicellular diploid organism. In some plants, such as ferns, both the haploid and diploid plant stages are free-living. The diploid plant is called a sporophyte because it produces haploid spores by meiosis. The spores develop into multicellular, haploid plants called gametophytes because they produce gametes. The gametes of two individuals will fuse to form a diploid zygote that becomes the sporophyte.

Although all plants utilize some version of the alternation of generations, the relative size of the sporophyte and the gametophyte and the relationship between them vary greatly. In plants such as moss, the gametophyte organism is the free-living plant, while the sporophyte is physically dependent on the gametophyte. In other plants, such as ferns, both the gametophyte and sporophyte plants are free-living; however, the sporophyte is much larger. In seed plants, such as magnolia trees and daisies, the gametophyte is composed of only a few cells and, in the case of the female gametophyte, is completely retained within the sporophyte.

Sexual reproduction takes many forms in multicellular organisms. However, at some point in each type of life cycle, meiosis produces haploid cells that will fuse with the haploid cell of another organism. The mechanisms of variation (crossover, random assortment of homologous chromosomes, and random fertilization) are present in all versions of sexual reproduction. The fact that nearly every multicellular organism on earth employs sexual reproduction is strong evidence for the benefits of producing offspring with unique gene combinations, although there are other possible benefits as well.

Difference between Mitosis and Meiosis

S.No.	Characters	Mitosis	Meiosis
I. General			

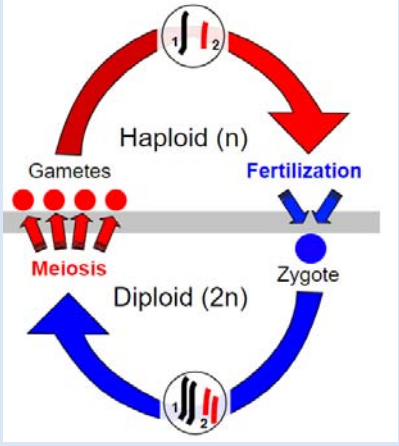
(1)	Site of occurrence	Somatic cells and during the multiplicative phase of gametogenesis in germ cells.	Reproductive germ cells of gonads.
(2)	Period of occurrence	Throughout life.	During sexual reproduction.
(3)	Nature of cells	Haploid or diploid.	Always diploid.
(4)	Number of divisions	Parental cell divides once.	Parent cell divides twice.
(5)	Number of daughter cells	Two.	Four.
(6)	Nature of daughter cells	Genetically similar to parental cell. Amount of DNA and chromosome number is same as in parental cell.	Genetically different from parental cell. Amount of DNA and chromosome number is half to that of parent cell.
II. Prophase			
(7)	Duration	Shorter (of a few hours) and simple.	Prophase-I is very long (may be in days or months or years) and complex.
(8)	Subphases	Formed of 3 subphases : early-prophase, mid-prophase and late-prophase.	Prophase-I is formed of 5 subphases: leptotene, zygotene, pachytene, diplotene and diakinesis.
(9)	Bouquet stage	Absent.	Present in leptotene stage.
(10)	Synapsis	Absent.	Pairing of homologous chromosomes in zygotene stage.
(11)	Chiasma formation and crossing over.	Absent.	Occurs during pachytene stage of prophase-I.

(12)	Disappearance of nucleolus and nuclear membrane	Comparatively in earlier part.	Comparatively in later part of prophase-I.
(13)	Nature of coiling	Plectonemic.	Paranemic.
III. Metaphase			
(14)	Metaphase plates	Only one equatorial plate	Two plates in metaphase-I but one plate in metaphase-II.
(15)	Position of centromeres	Lie at the equator. Arms are generally directed towards the poles.	Lie equidistant from equator and towards poles in metaphase-I while lie at the equator in metaphase-II.
(16)	Number of chromosomal fibres	Two chromosomal fibre join at centromere.	Single in metaphase-I while two in metaphase-II.
IV. Anaphase			
(17)	Nature of separating chromosomes	Daughter chromosomes (chromatids with independent centromeres) separate.	Homologous chromosomes separete in anaphase-I while chromatids separate in anaphase in anaphase-II.
(18)	Splitting of centromeres and development of inter-zonal fibres	Occurs in anaphase.	No splitting of centromeres. Inter-zonal fibres are developed in metaphase-I.
V. Telophase			
(19)	Occurrence	Always occurs	Telophase-I may be absent but telophase-II is always present.
VI. Cytokinesis			

(20)	Occurrence	Always occurs	Cytokinesis-I may be absent but cytokinesis-II is always present.
(21)	Nature of daughter cells	2N amount of DNA than 4N amount of DNA in parental cell.	1 N amount of DNA than 4 N amount of DNA in parental cell.
(22)	Fate of daughter cells	Divide again after interphase.	Do not divide and act as gametes.
VII. Significance			
(23)	Functions	Helps in growth, healing, repair and multiplication of somatic cells. Occurs in both asexually and sexually reproducing organisms.	Produces gametes which help in sexual reproduction.
(24)	Variations	Variations are not produced as it keeps quality and quantity of genes same.	Produces variations due to crossing over and chance arrangement of bivalents at metaphase-I.
(25)	In evolution	No role in evolution.	It plays an important role in speciation and evolution.

CONCEPT OF SEXUAL REPRODUCTION

Sexual reproduction is a fundamental biological process in which two parent organisms contribute genetic material to produce offspring with genetic diversity. Unlike asexual reproduction, which involves the production of genetically identical offspring from a single parent, sexual reproduction involves the fusion of specialized cells called gametes, typically sperm and egg. The gametes are produced through the process of meiosis, which halves the number of chromosomes in each cell to ensure the offspring have the correct number of chromosomes. During sexual reproduction, the male and female gametes combine to form a zygote, a fertilized cell with a complete set of chromosomes. This process introduces genetic variation, as each parent contributes a unique combination of genes to the offspring. Genetic diversity is advantageous as it enhances the adaptability and survival of the species in changing environments. Sexual reproduction is a prevalent and essential mode of reproduction in the animal and plant kingdoms, contributing to the richness of life's forms and evolutionary history.



SUMMARY

- Mitosis, a process of cell duplication, or reproduction, during which one cell gives rise to two genetically identical daughter cells. Strictly applied, the term mitosis is used to describe the duplication and distribution of chromosomes, the structures that carry the genetic information.
- As prophase begins, the chromatin, which are normally spread throughout the nucleus, begin to condense into an X shape, held together in the middle with a specific sequence of DNA called a centromere.
- The soon to be divided chromosomes are symmetrically positioned on the metaphase plate, which is essentially the equator of the parent cell.
- Sexual reproduction is the production of haploid cells (gametes) and the fusion (fertilization) of two gametes to form a single, unique diploid cell called a zygote. All animals and most plants produce these gametes, or eggs and sperm.
- In anaphase I, the microtubules pull the attached chromosomes apart. The sister chromatids remain tightly bound together at the centromere. The chiasmata are broken in anaphase I as the microtubules attached to the fused kinetochores pull the homologous chromosomes apart.
- Mitosis and meiosis are both forms of division of the nucleus in eukaryotic cells. They share some similarities, but also exhibit distinct differences that lead to very different outcomes. The purpose of mitosis is cell regeneration, growth, and asexual reproduction, while the purpose of meiosis is the production of gametes for sexual reproduction.
- Sexual reproduction was an early evolutionary innovation after the appearance of eukaryotic cells. During sexual reproduction, the genetic material of two individuals is combined to produce genetically-diverse offspring that differ from their parents.
- Within haploid-dominant life cycles, the multicellular haploid stage is the most obvious life stage. Most fungi and algae employ a life cycle type in which the “body” of the organism, the ecologically important part of the life cycle, is haploid.

MULTIPLE CHOICE QUESTIONS

1. The evolutionary advantage of meiosis can be best explained by which of these statements?

a. Meiosis alternates with mitosis from one to the next generation
b. Meiosis is essential for sexual reproduction
c. Passing of the same genetic system from one to next generation
d. Genetic recombination is possible from one to next generation

2. One of these events does not take place during meiosis

a. One successive division without any DNA replication
b. Chiasmata formation and crossing over
c. Segregation of homologous chromosomes
d. Separation of sister chromatids

3. The meiotic division takes place in

a. Meristematic cells
b. Conductive cells
c. Reproductive cells
d. Vegetative cells

4. Name the event wherein the paternal and maternal chromosomes change their material with each other in cell division

a. Crossing over
b. Synapsis
c. Dyad forming
d. Bivalent forming

5. Mendelian factor (Aa) is segregated during

a. Anaphase I
b. Anaphase II
c. Diplotene
d. Zygotene/Pachytene

Answer

1. (d) 2. (a) 3. (c) 4. (a) 5. (a)

REVIEW QUESTIONS

1. Explain different phases of mitosis.
2. Describe the importance of meiosis in sexual reproduction.
3. Describe the stages and results of meiosis II.
4. Describe the stages and results of meiosis I.
5. Explain the life cycles in sexual reproduction.
6. Describe the benefits of sexual reproduction.
7. Difference between mitosis and amitosis.

REFERENCES

1. Cell Division: Stages of Mitosis Learn Science at Scitable". www.nature.com. Archived from the original on 2015-11-14. Retrieved 2015-11-16.
2. Battaglia E (2009). "Caryoneme alternative to chromosome and a new caryological nomenclature" (PDF). Caryologia. 62 (4): 1–80. Archived from the original (PDF) on 2016-03-04.
3. Kops GJ, Weaver BA, Cleveland DW (October 2005). "On the road to cancer: aneuploidy and the mitotic checkpoint". Nature Reviews. Cancer. 5 (10): 773–85.
4. Chan GK, Liu ST, Yen TJ (November 2005). "Kinetochore structure and function". Trends in Cell Biology. 15 (11): 589–98.
5. De Souza CP, Osmani SA (September 2007). "Mitosis, not just open or closed". Eukaryotic Cell. 6 (9): 1521–7.
6. Kalatova B, Jesenska R, Hlinka D, Dudas M (January 2015). "Tripolar mitosis in human cells and embryos: occurrence, pathophysiology and medical implications". Acta Histochemica. 117 (1): 111–25.
7. Sandoz PA (December 2019). "Image-based analysis of living mammalian cells using label-free 3D refractive index maps reveals new organelle dynamics and dry mass flux". PLOS Biology. 17 (12): e3000553.



CHAPTER 5

DNA STRUCTURE AND ANALYSIS



DNA, abbreviation of deoxyribonucleic acid, organic chemical of complex molecular structure that is found in all prokaryotic and eukaryotic cells and in many viruses. DNA codes genetic information for the transmission of inherited traits.

After studying this chapter, you will understand the core concepts of:

- Properties of DNA
- DNA Structure
- DNA Functions

INTRODUCTION

The discovery of the structure of DNA by James Watson and Francis Crick in 1953 marked a groundbreaking moment in biology, revolutionizing our understanding of genetics and laying the foundation for modern molecular biology. DNA, short for deoxyribonucleic acid, is a complex molecule that carries the genetic information essential for the development, functioning, and reproduction of all living organisms. Its remarkable structure, a double helix, consists of two intertwined strands made up of nucleotides.

Each nucleotide in the DNA molecule is composed of three components: a sugar molecule called deoxyribose, a phosphate group, and one of four nitrogenous bases: adenine (A), cytosine (C), guanine (G), and thymine (T). The nitrogenous bases form the "rungs" of the DNA ladder, with adenine pairing with thymine and cytosine pairing with guanine through hydrogen bonds, creating the base pairs that stabilize the double helix structure.

The sequence of base pairs along the DNA molecule forms the genetic code, encoding the instructions for building proteins and carrying out various cellular processes. Understanding the structure of DNA has been crucial in deciphering the genetic information and unraveling the molecular mechanisms underlying life's diversity and complexity.

The analysis of DNA has become a cornerstone of modern biology and numerous applications across various fields. DNA analysis techniques have evolved significantly, providing powerful tools for studying genetic variation, identifying genetic disorders, solving crimes, and understanding evolutionary relationships among species.

One of the most important techniques in DNA analysis is polymerase chain reaction (PCR), which allows scientists to amplify specific regions of DNA, making it possible to study small samples of DNA and identify genetic markers associated with various traits and diseases. PCR has become an indispensable tool in medical diagnostics, forensic investigations, and genetic research.

DNA sequencing is another critical method that allows scientists to determine the exact order of nucleotides in a DNA molecule, revealing the genetic code and providing insights into an organism's genetic makeup. The development of high-throughput DNA sequencing technologies has revolutionized genomics, enabling the sequencing of entire genomes, including those of humans and various other organisms.

Furthermore, DNA fingerprinting, or DNA profiling, has become a widely used method for identifying individuals based on their unique DNA patterns. It has applications in forensic science, paternity testing, and the identification of human remains.

In addition to its applications in the medical and forensic fields, DNA analysis has been instrumental in studying evolutionary relationships among species. Comparing DNA sequences from different organisms helps scientists understand their genetic relatedness and trace their evolutionary history.

The study of DNA structure and analysis continues to advance with ongoing advancements in technology and bioinformatics. The integration of DNA analysis with other -omics approaches, such as transcriptomics, proteomics, and metabolomics, promises to provide a more comprehensive understanding of biological systems and their intricate networks.

The discovery of the structure of DNA and the development of DNA analysis techniques have been monumental achievements in biology. DNA serves as the blueprint of life, encoding the genetic information that shapes the characteristics and functions of all living organisms. The analysis of DNA has revolutionized numerous fields, from medicine and forensics to evolutionary biology and genetics. With continuous innovations and discoveries, DNA analysis remains at the forefront of scientific research, offering insights into the complexity and diversity of life and providing a powerful tool for solving biological mysteries and improving human health.



CORE CONCEPT: PROPERTIES OF DNA

DNA or deoxyribonucleic acid is the basic hereditary material present in all the cells of an organism and basically provides a blue print for the cell's functions, growth, reproduction and death.

DNA has a number of special physical and chemical properties that are important to its structure and functioning. While the terms may sound quite foreign and perhaps even complicated, they are significant to our understanding of DNA and its many applications today.

Physical Properties

In living organisms such as humans, DNA exists as a pair of molecules rather than a single molecule. These strands are entwined in the shape of a double helix and the helix is kept stable by hydrogen bonds, which can be found between the bases attached to the two strands. A long polymer, DNA is made up of smaller units called nucleotides. In turn, each nucleotide consists of a phosphate group, a sugar and a nitrogenous base. To get a sense of just how long a DNA polymer is, consider that just one human chromosome is several hundred million base pairs long.

Base Pairing

Cytosine is a pyrimidine, and one of the nitrogenous bases found in ribonucleic acid (RNA) and deoxyribonucleic acid (DNA).

Base pairing is a defining property of DNA and was particularly exciting when it was first discovered because it suggested the 'copying' mechanism for DNA. In DNA, bases are specific in that an adenine base, for example, only pairs with a thymine base. Following on that premise, a cytosine base will only bond to a guanine base. This base pairing is also known as complementary base pairing. The concept is quite simplistic but it is significant for DNA.

DNA Grooves

DNA has two kinds of grooves that play important roles in its functioning. Major and minor grooves are structures to allow for necessary proteins in your body to make contact with bases. You might wonder what makes it so important for proteins to attach to the bases. Some of these proteins are called transcription factors. While the word may well be a foreign one for those unfamiliar with DNA - which is most of the public - it is a key one. They are also significant for cell development and communication. As such, these DNA grooves seen in the structure of DNA facilitate the binding of proteins like transcription factors, which then serves to keep the cellular processes occurring effectively within your body.

DNA Supercoiling

If you try to picture a rope, you can get a visual image of DNA supercoiling. This coiling is a central property of DNA. DNA can be in a relaxed or coiled state and it is this coiling that allows our extremely long strands of DNA to fit or 'pack' into the comparatively much smaller cells in our bodies. This supercoiling property makes DNA more efficient by allowing it to make better use of space and fit neatly into a significantly smaller space relative to its relaxed, longer size.

DNA Conformations

DNA can exist in different conformations and these are important for a range of DNA mechanisms. These conformations interact with enzymes in your body and are also involved in aspects such as DNA repair.

DNA Sense and Antisense

Transcription factors have numerous roles such as telling your cells how large they should become or suppressing tumour growth.

To understand this aspect of DNA, you first need to be aware of RNA - ribonucleic acid. DNA holds the genetic 'codes' to make RNA. In turn, RNA is needed in your body because it holds the codes for the amino acids that make proteins. The antisense strand is the DNA strand that carries important information to make proteins by binding to the RNA. This antisense strand is the key for making proteins. In comparison, the sense strand is the one that does not code for RNA.

Despite the complex detail and challenges that come with understanding the properties of DNA, they are clearly important to its design and functioning in the human body. Since the discovery of DNA and our knowledge of its basic physical properties such as the double helix, we have progressed a great deal to learning about the chemical properties of DNA and interactions between DNA and its environment. Hopefully, we will learn even more detailed information about DNA structure and functioning in the future, which will have important applications in disease treatment and many other fields.



CORE CONCEPT: DNA STRUCTURE

DNA structure itself is actually quite fascinating, even perhaps to non-scientists! DNA can basically be broken down into smaller parts, which makes its important structure a bit easier to understand. If you can try to picture a helix, you have already got a sense of the structure of DNA. At a very simple level, DNA is a helical structure, with bases that are 'stacked' upon one another. These bases are part of the DNA subunits that are called nucleotides.

DNA or deoxyribonucleic acid is a helically twisted double chain polydeoxyribonucleotide macromolecule which constitutes the genetic material of all organisms with the exception of rhinoviruses. In prokaryotes it occurs in nucleoid and plasmids. This DNA is usually circular. In eukaryotes, most of the DNA is found in chromatin of nucleus.

DNA Is Composed of Polynucleotide Chains

The most important feature of DNA is that it is usually composed of two polynucleotide chains twisted around each other in the form of a double helix (Figure 1). The upper part

Double helix is a description of the molecular shape of a double-stranded DNA molecule.

of the figure (a) presents the structure of the double helix shown in a schematic form. Note that if inverted 180°, the **double helix** looks superficially the same, due to the complementary nature of the two DNA strands. The space-filling model of the double helix, in the lower part of the figure (b), shows the components of the DNA molecule and their relative positions in the helical structure. The backbone of each strand of the helix is composed of alternating sugar and phosphate residues; the bases project inward but are accessible through the major and minor grooves.

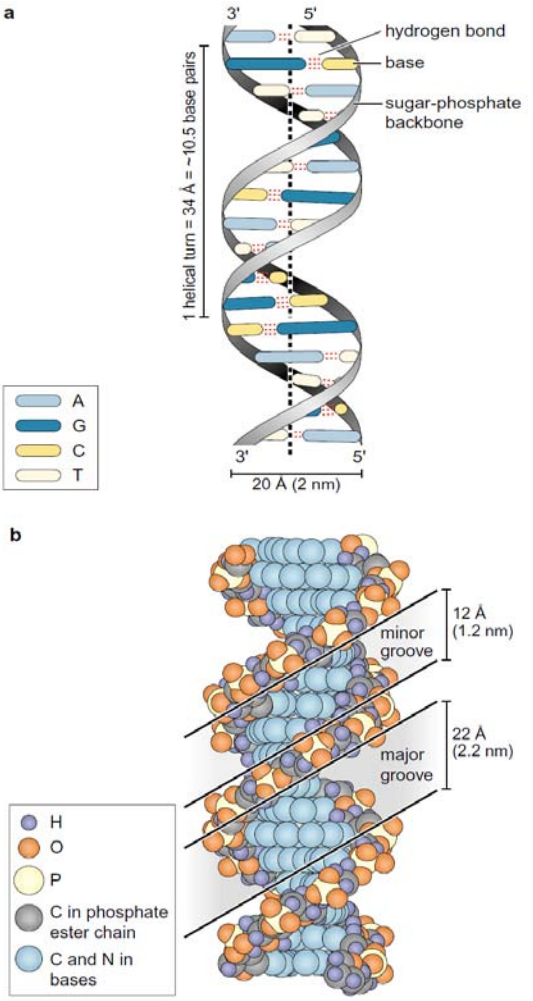


Figure 1: The Helical Structure of DNA. (a) Schematic model of the double helix. One turn of the helix (34 Å or 3.4 nm) spans approx. 10.5 base pairs. (b) Space-filling model of the double helix. The sugar and phosphate residues in each strand form the backbone, which are traced by the yellow, gray, and red circles, show the helical twist of the overall molecule. The bases project inward but are accessible through major and minor grooves.

Let us begin by considering the nature of the nucleotide, the fundamental building block of DNA. The nucleotide consists of a phosphate joined to a sugar, known as 2-deoxyribose, to which a base is attached. The phosphate and the sugar have the structures shown in Figure 2. The sugar is called 2-deoxyribose because there is no hydroxyl at position 2 (just two hydrogens). Note that the positions on the ribose are designated with primes to distinguish them from positions on the bases.

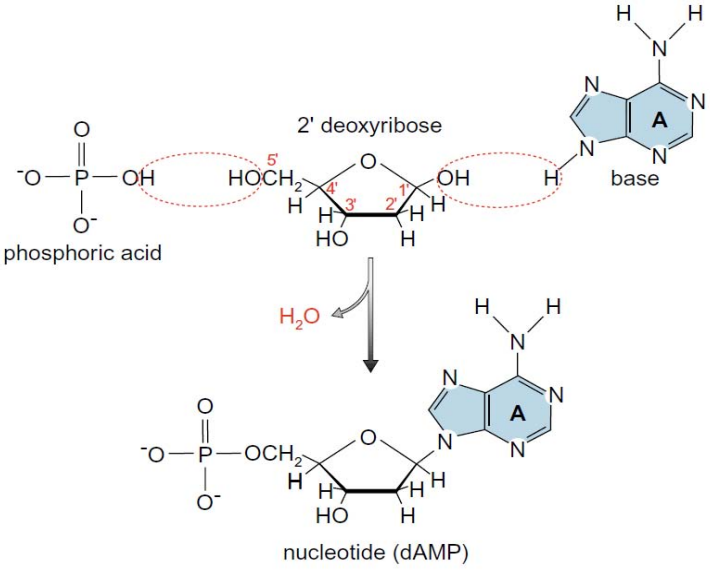


Figure 2: Formation of Nucleotide by Removal of Water. The numbers of the carbon atoms in 2' deoxyribose are labeled in red.

We can think of how the base is joined to 2-deoxyribose by imagining the removal of a molecule of water between the hydroxyl on the 1 carbon of the sugar and the base to form a glycosidic bond (Figure 2). The sugar and base alone are called a nucleoside. Likewise, we can imagine linking the phosphate to 2-deoxyribose by removing a water molecule from between the phosphate and the hydroxyl on the 5 carbon to make a 5 phosphomonoester. Adding a phosphate (or more than one phosphate) to a nucleoside creates a nucleotide. Thus, by making a glycosidic bond between the base and the sugar, and by making a phosphoester bond between the sugar and the phosphoric acid, we have created a nucleotide.

Nucleotides are, in turn, joined to each other in polynucleotide chains through the 3 hydroxyl of 2-deoxyribose of one nucleotide and the phosphate attached to the 5 hydroxyl of another nucleotide (Figure 3). This is a phosphodiester linkage in which the phosphoryl group between the two nucleotides

Phosphodiester bond occurs when exactly two of the hydroxyl groups in phosphoric acid react with hydroxyl groups on other molecules to form two ester bonds

has one sugar esterified to it through a 3 hydroxyl and a second sugar esterified to it through a 5 hydroxyl. **Phosphodiester** linkages create the repeating, sugar-phosphate backbone of the polynucleotide chain, which is a regular feature of DNA. In contrast, the order of the bases along the polynucleotide chain is irregular. This irregularity as well as the long length is, as we shall see, the basis for the enormous information content of DNA.

The phosphodiester linkages impart an inherent polarity to the DNA chain. This polarity is defined by the asymmetry of the nucleotides and the way they are joined. DNA chains have a free 5 phosphate or 5 hydroxyl at one end and a free 3 phosphate or 3 hydroxyl at the other end. The convention is to write DNA sequences from the 5 end (on the left) to the 3 end, generally with a 5 phosphate and a 3 hydroxyl.

Each Base Has Its Preferred Tautomeric Form

The bases in DNA fall into two classes, purines and pyrimidines. The purines are adenine and guanine, and the pyrimidines are cytosine and thymine. The purines are derived from the double-ringed structure shown in Figure 4. Adenine and guanine share this essential structure but with different groups attached. Likewise, cytosine and thymine are variations on the single-ringed structure shown in Figure 4. The figure also shows the numbering of the positions in the purine and pyrimidine rings. The bases are attached to the deoxyribose by glycosidic linkages at N1 of the pyrimidines or at N9 of the purines.

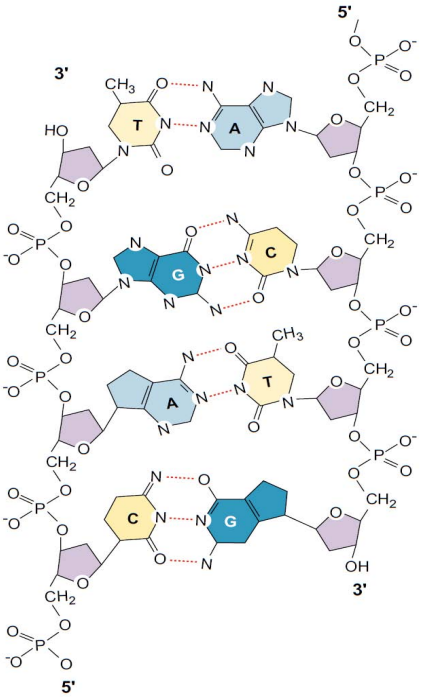


Figure 3: Detailed Structure of Polynucleotide Polymer. The structure shows base pairing between purines (in blue) and pyrimidines (in yellow), and the phosphodiester linkages of the backbone.

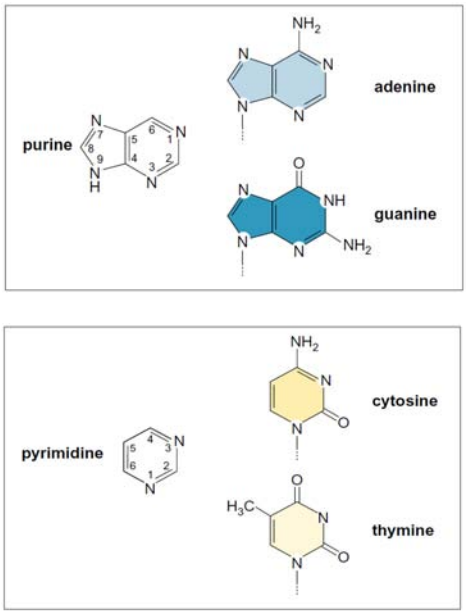


Figure 4: Purines and Pyrimidines. The dotted lines indicate the sites of attachment of the bases to the sugars.

Each of the bases exists in two alternative tautomeric states, which are in equilibrium with each other. The equilibrium lies far to the side of the conventional structures shown in Figure 4, which are the predominant states and the ones important for base pairing. The nitrogen atoms attached to the purine and pyrimidine rings are in the amino form in the predominant state and only rarely assume the imino configuration. Likewise, the oxygen atoms attached to the guanine and thymine normally have the keto form and only rarely take on the enol configuration. As examples, Figure 5 shows tautomerization of cytosine into the imino form (a) and guanine into the enol form (b). As we shall see, the capacity to form an alternative tautomer is a frequent source of errors during DNA synthesis.

The Two Strands of the Double Helix Are Held Together by Base Pairing in an Anti-Parallel Orientation

The double helix is composed of two polynucleotide chains that are held together by weak, non-covalent bonds between pairs of bases, as shown in Figure 3. Adenine on one chain is always paired with thymine on the other chain and, likewise, guanine is always paired with cytosine.

The two strands have the same helical geometry but base pairing holds them together with the opposite polarity. That is, the base at the 5 end of one strand is paired with the base at the 3 end of the other strand. The strands are said to have an anti-parallel orientation. This anti-parallel orientation is a stereochemical consequence of the way that adenine and thymine and guanine and cytosine pair with each together (see Figure 6).

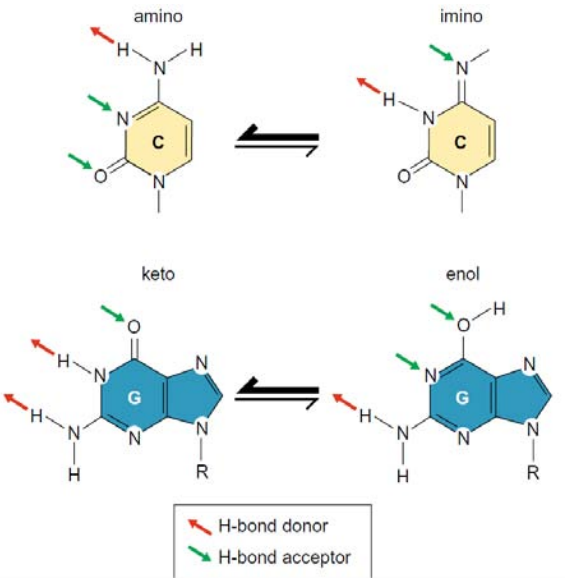


Figure 5: Base Tautomers. Amino ⇌ imino and keto ⇌ enol tautomerism. (a) Cytosine is usually in the amino form but rarely forms the imino configuration. (b) Guanine is usually in the keto form but is rarely found in the enol configuration.

The Two Chains of the Double Helix Have Complementary Sequences

The pairing between adenine and thymine and between guanine and cytosine results in a complementary relationship between the sequence of bases on the two intertwined chains and gives DNA its self-encoding character. For example, if we have the sequence 5-ATGTC-3 on one chain, the opposite chain must have the complementary sequence 3-TACAG-5.

The strictness of the rules for this "Watson-Crick" pairing derives from the complementarity both of shape and of hydrogen bonding properties between adenine and thymine and between guanine and cytosine (Figure 6). Adenine and thymine match up so that a hydrogen bond can form between the exocyclic amino group at C6 on adenine and the carbonyl at C4 in thymine; and likewise, a hydrogen bond can form between N1 of adenine and N3 of thymine. A corresponding arrangement can be drawn between a guanine and a cytosine, so that there is both hydrogen bonding and shape complementarity in this base pair as well. A G:C base pair has three hydrogen bonds, because the exocyclic NH_2 at C2 on guanine lies opposite to, and can hydrogen bond with, a carbonyl at C2 on cytosine. Likewise, a hydrogen bond can form between N1 of guanine and N3 of cytosine and between the carbonyl at C6 of guanine and the exocyclic NH2 at C4 of cytosine. Watson-Crick base pairing requires that the bases are in their preferred tautomeric states.

An important feature of the double helix is that the two base pairs have exactly the same geometry; having an A:T base pair or a G:C base pair between the two sugars does not perturb the arrangement of the sugars. Neither does T:A or C:G. In other

words, there is an approximately twofold axis of symmetry that relates the two sugars and all four base pairs can be accommodated within the same arrangement without any distortion of the overall structure of the DNA.

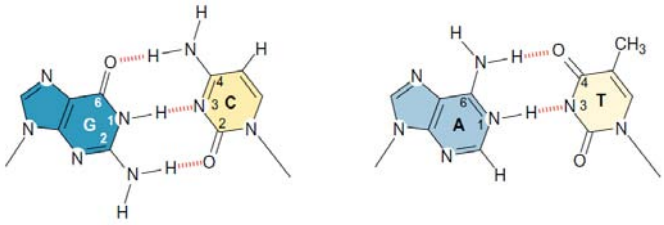


Figure 6: A:T and G:C Base Pairs. The figure shows hydrogen bonding between the bases.

Hydrogen Bonding Is Important for the Specificity of Base Pairing

The hydrogen bonds between complementary bases are a fundamental feature of the double helix, contributing to the thermodynamic stability of the helix and providing the information content and specificity of base pairing. Hydrogen bonding might not at first glance appear to contribute importantly to the stability of DNA for the following reason. An organic molecule in aqueous solution has all of its hydrogen bonding properties satisfied by water molecules that come on and off very rapidly. As a result, for every hydrogen bond that is made when a base pair forms, a hydrogen bond with water is broken that was there before the base pair formed. Thus, the net energetic contribution of hydrogen bonds to the stability of the double helix would appear to be modest. However, when polynucleotide strands are separate, water molecules are lined up on the bases. When strands come together in the double helix, the water molecules are displaced from the bases. This creates disorder and increases entropy, thereby stabilizing the double helix. Hydrogen bonds are not the only force that stabilizes the double helix. A second important contribution comes from stacking interactions between the bases. The bases are flat, relatively water insoluble molecules, and they tend to stack above each other roughly perpendicular to the direction of the helical axis. Electron cloud interactions ($\pi - \pi$) between bases in the helical stacks contribute significantly to the stability of the double helix.

An A:C base pair would be unstable because water would have to be stripped off the donor and acceptor groups without restoring the hydrogen bond formed within the base pair.

Hydrogen bonding is also important for the specificity of base pairing. Suppose we tried to pair an adenine with a cytosine. Then we would have a hydrogen bond acceptor (N1 of adenine) lying opposite a hydrogen bond acceptor (N3 of cytosine) with no room to put a water molecule in between to satisfy the two acceptors (Figure 7). Likewise, two hydrogen bond donors, the NH_2 groups at C6 of adenine and C4 of cytosine, would lie opposite each other.

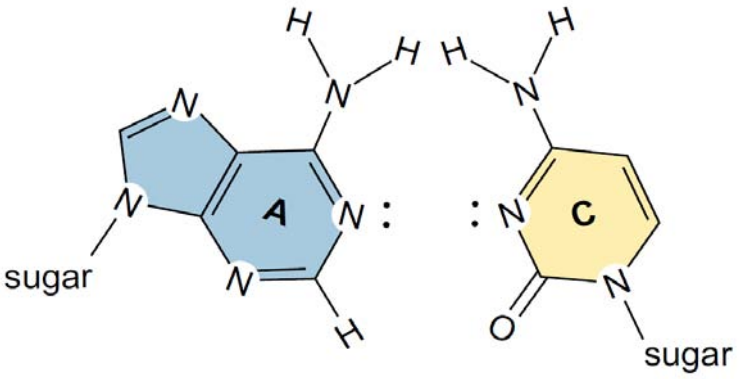


Figure 7: A:C Incompatibility. The structure shows the inability of adenine to form the proper hydrogen bonds with cytosine. The base pair is therefore unstable.

Bases Can Flip Out from the Double Helix

As we have seen, the energetics of the double helix favor the pairing of each base on one polynucleotide strand with the complementary base on the other strand. Sometimes, however, individual bases can protrude from the double helix in a remarkable phenomenon known as base flipping shown in Figure 8. Certain enzymes that methylate bases or remove damaged bases do so with the base in an extra helical configuration in which it is flipped out from the double helix, enabling the base to sit in the catalytic cavity of the enzyme. Furthermore, enzymes involved in homologous recombination and DNA repair are believed to scan DNA for homology or lesions by flipping out one base after another. This is not energetically expensive because only one base is flipped out at a time. Clearly, DNA is more flexible than might be assumed at first glance.

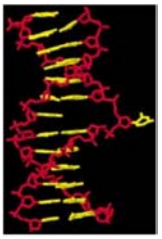


Figure 8: Base Flipping. Structure of isolated DNA, showing the flipped cytosine residue and the small distortions to the adjacent base pairs.

DNA Is Usually a Right-Handed Double Helix

Applying the handedness rule from physics, we can see that each of the **polynucleotide** chains in the double helix is right-handed. In your mind's eye, hold your right hand up to the DNA molecule in Figure 9 with your thumb pointing up and along the long axis of the helix and your fingers following the grooves in the helix. Trace along one strand of the helix in the direction in which your thumb is pointing. Notice that you

go around the helix in the same direction as your fingers are pointing. This does not work if you use your left hand.

Polynucleotide molecule is a biopolymer composed of 13 or more nucleotide monomers covalently bonded in a chain.

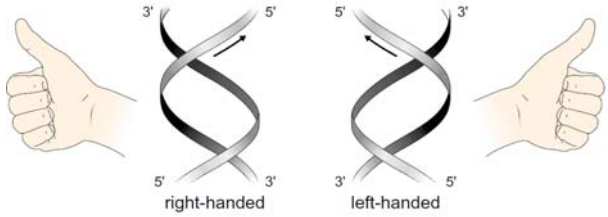


Figure 9: Left- and Right-Handed Helices.

A consequence of the helical nature of DNA is its periodicity. Each base pair is displaced (twisted) from the previous one by about 36°. Thus, in the X-ray crystal structure of DNA it takes a stack of about 10 base pairs to go completely around the helix (360°) (see Figure 1a). That is, the helical periodicity is generally 10 base pairs per turn of the helix.

The Double Helix Has Minor and Major Grooves

As a result of the double-helical structure of the two chains, the DNA molecule is a long extended polymer with two grooves that are not equal in size to each other. Why are there a minor groove and a major groove? It is a simple consequence of the geometry of the base pair. The angle at which the two sugars protrude from the base pairs (that is, the angle between the glycosidic bonds) is about 120° (for the narrow angle or 240° for the wide angle) (see Figures 1b and 6). As a result, as more and more base pairs stack on top of each other, the narrow angle between the sugars on one edge of the base pairs generates a minor groove and the large angle on the other edge generates a major groove.

The Major Groove is Rich in Chemical Information

The edges of each base pair are exposed in the major and minor grooves, creating a pattern of hydrogen bond donors and acceptors and of van der Waals surfaces that identifies the base pair (see Figure 10). The edge of an A:T base pair displays the following chemical groups in the following order in the major groove: a hydrogen bond acceptor (the N7 of adenine), a hydrogen bond donor (the exocyclic amino group on C6 of adenine), a hydrogen bond acceptor (the carbonyl group on C4

of thymine) and a bulky hydrophobic surface (the methyl group on C5 of thymine). Similarly, the edge of a G:C base pair displays the following groups in the major groove: a hydrogen bond acceptor (at N7 of guanine), a hydrogen bond acceptor (the carbonyl on C6 of guanine), a hydrogen bond donor (the exocyclic amino group on C4 of cytosine), a small non-polar hydrogen (the hydrogen at C5 of cytosine).

Thus, there are characteristic patterns of hydrogen bonding and of overall shape that are exposed in the major groove that distinguish an A:T base pair from a G:C base pair, and, for that matter, A:T from T:A, and G:C from C:G. We can think of these features as a code in which A represents a hydrogen bond acceptor, D a hydrogen bond donor, M a methyl group, and H a nonpolar hydrogen. In such a code, A D A M in the major groove signifies an A:T base pair, and A A D H stands for a G:C base pair. Likewise, M A D A stands for a T:A base pair and H D A A is characteristic of a C:G base pair. In all cases, this code of chemical groups in the major groove specifies the identity of the base pair. These patterns are important because they allow proteins to unambiguously recognize DNA sequences without having to open and thereby disrupt the double helix. Indeed, as we shall see, a principal decoding mechanism relies upon the ability of amino acid side chains to protrude into the major groove and to recognize and bind to specific DNA sequences.

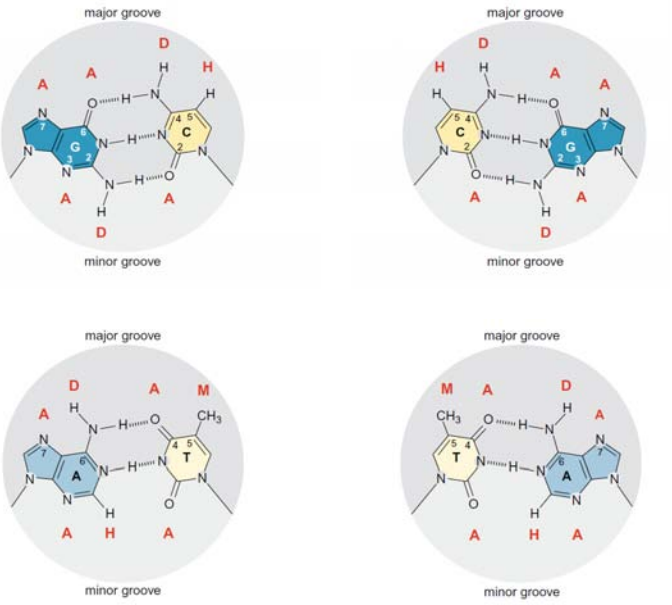


Figure 10: Chemical Groups Exposed in the Major and Minor Grooves from the Edges of the Base Pairs. The letters in red identify hydrogen bond acceptors (A), hydrogen bond donors (D), nonpolar hydrogens (H), and methyl groups (M).

The minor groove is not as rich in chemical information and what information is available is less useful for distinguishing between base pairs. The small size of the minor groove is less able to accommodate amino acid side chains. Also, A:T and T:A base pairs and G:C and C:G pairs look similar to one another in the minor groove. An A:T base pair has a hydrogen bond acceptor (at N3 of adenine), a nonpolar hydrogen (at N2 of adenine) and a hydrogen bond acceptor (the carbonyl on C2 of thymine).

Thus, its code is A H A. But this code is the same if read in the opposite direction, and hence an A:T base pair does not look very different from a T:A base pair from the point of view of the hydrogenbonding properties of a protein poking its side chains into the minor groove. Likewise, a G:C base pair exhibits a hydrogen bond acceptor (at N3 of guanine), a hydrogen bond donor (the exocyclic amino group on C2 of guanine), and a hydrogen bond acceptor (the carbonyl on C2 of cytosine), representing the code A D A. Thus, from the point of view of hydrogen bonding, C:G and G:C base pairs do not look very different from each other either. The minor groove does look different when comparing an A:T base pair with a G:C base pair, but G:C and C:G, or A:T and T:A, cannot be easily distinguished (see Figure 10).

The Double Helix Exists in Multiple Conformations

Early X-ray diffraction studies of DNA, which were carried out using concentrated solutions of DNA that had been drawn out into thin fibers, revealed two kinds of structures, the B and the A forms of DNA (Figure 11). The B form, which is observed at high humidity, most closely corresponds to the average structure of DNA under physiological conditions. It has 10 base pairs per turn, and a wide major groove and a narrow minor groove. The A form, which is observed under conditions of low humidity, has 11 base pairs per turn. Its major groove is narrower and much deeper than that of the B form, and its minor groove is broader and shallower. The vast majority of the DNA in the cell is in the B form, but DNA does adopt the A structure in certain DNA-protein complexes. Also, as we shall see, the A form is similar to the structure that RNA adopts when double helical.

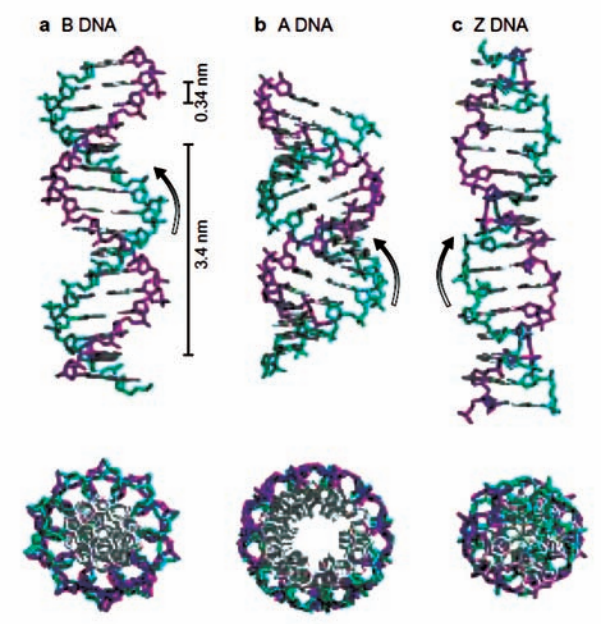


Figure 11: Models of the B, A, and Z Forms of DNA.

The B form of DNA represents an ideal structure that deviates in two respects from the DNA in cells. First, DNA in solution, as we have seen, is somewhat more twisted on average than the B form, having on average 10.5 base pairs per turn of the helix. Second, the B form is an average structure whereas real DNA is not perfectly regular. Rather, it exhibits variations in its precise structure from base pair to base pair. This was revealed by comparison of the crystal structures of individual DNAs of different sequences. For example, the two members of each base pair do not always lie exactly in the same plane. Rather, they can display a "propeller twist" arrangement in which the two flat bases counter rotate relative to each other along the long axis of the base pair, giving the base pair a propeller-like character (Figure 12). Moreover, the precise rotation per base pair is not a constant. As a result, the width of the major and minor grooves varies locally. Thus, DNA molecules are never perfectly regular double helices. Instead, their exact conformation depends on which base pair (A:T, T:A, G:C, or C:G) is present at each position along the double helix and on the identity of neighboring base pairs. Still, the B form is for many purposes a good first approximation of the structure of DNA in cells.

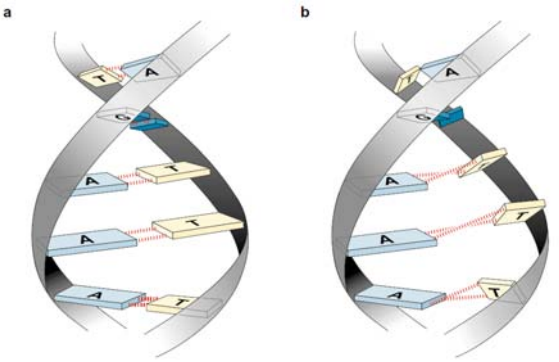


Figure 12: The Propeller Twist between the Purine and Pyrimidine Base Pairs of a Right-Handed Helix.

A **glycosidic bond** or glycosidic linkage is a type of covalent bond that joins a carbohydrate (sugar) molecule to another group, which may or may not be another carbohydrate.

DNA Can Sometimes Form a Left-Handed Helix

DNA containing alternative purine and pyrimidine residues can fold into left-handed as well as right-handed helices. To understand how DNA can form a left-handed helix, we need to consider the glycosidic bond that connects the base to the 1 position of 2-deoxyribose. This bond can be in one of two conformations called syn and anti (Figure 13). In right-handed

DNA, the **glycosidic bond** is always in the anti-conformation. In the left-handed helix, the fundamental repeating unit usually is a purine-pyrimidine dinucleotide, with the glycosidic bond in the anti-conformation at pyrimidine residues and in the syn conformation at purine residues. It is this syn conformation at the purine nucleotides that is responsible for the left-handedness of the helix. The change to the syn position in the purine residues to alternating anti–syn conformations gives the backbone of left-handed DNA a zigzag look (hence its designation of Z DNA; see Figure 11), which distinguishes it from right-handed forms. The rotation that effects the change from anti to syn also causes the ribose group to undergo a change in its pucker. Note, as shown in Figure 13, that C3 and C2 can switch locations. In solution alternating purine–pyrimidine residues assume the left-handed conformation only in the presence of high concentrations of positively charged ions (e.g., Na^+) that shield the negatively charged phosphate groups. At lower salt concentrations, they form typical right-handed conformations. The physiological significance of Z DNA is uncertain and left-handed helices probably account at most for only a small of proportion of a cell's DNA.

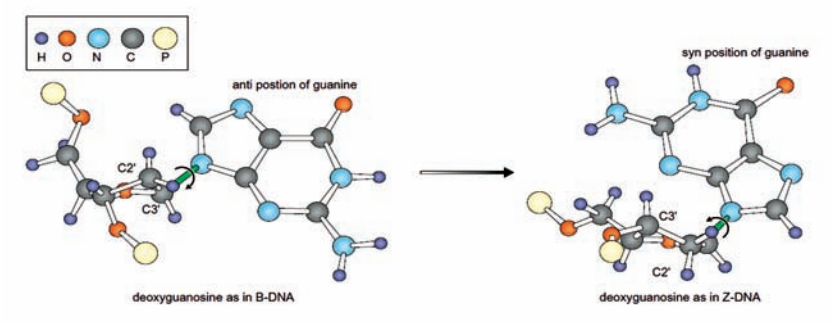


Figure 13: Syn and Anti Positions of Guanine in B and Z DNA.



CORE CONCEPT: DNA FUNCTIONS

DNA stores the information needed to build and control the cell. The transmission of this information from mother to daughter cells is called vertical gene transfer and it occurs through the process of DNA replication. DNA is replicated when a cell makes a duplicate copy of its DNA, then the cell divides, resulting in the correct distribution of one DNA copy to each resulting cell.

The functions of DNA are vital for inheritance, coding for proteins and the genetic blueprint of life. Given the enormity of DNA's functions in the human body and its responsibility for the growth and maintenance of life, it is not surprising that the discovery of DNA has led to such a great number of developments in treating disease.

DNA holds the instructions for an organism's development and reproduction - ultimately, its survival.

Coding for Proteins

DNA holds the code for proteins, which are complex molecules that do huge amounts of work around our body. Information in DNA is initially 'read' and then it is transcribed into a messenger molecule. After, the information held in this messenger molecule is translated into a 'language' that the body can understand. This language is one of amino acids, which are also known as the building blocks of proteins. It is this specific language that dictates how the amino acids should produce a particular protein. If you think about the twenty different kinds of amino acids, you can see that the ordering can produce an enormous variety of proteins.

DNA Replication

DNA replication is vital for a virtually endless list of functions, from reproduction to maintenance and growth of cells, tissues and body systems. To copy itself, a DNA molecule essentially 'unzips,' thus resulting in a series of bases without pairs along the backbone of the molecule. DNA has four bases - all part of a nucleotide that also consists of a sugar and phosphate. The four bases in DNA are very specific about which base they will attach to, which means that adenine only pairs with thymine and guanine will only pair with cytosine. As the nucleotides connect with unpaired bases on the backbone of the DNA molecule, they build a new strand that complements - or matches - the original sequence. The end result is a strand that is a perfect match to the original one prior to it unzipping.

Blood cell, also called a haematopoietic cell, hemocyte, or hematocyte, is a cell produced through hematopoiesis and found mainly in the blood.

Cells in your body replicate for purposes such as making new skin or **blood cells**. When mistakes occur, there are repair systems in place to remedy the mistake or alternately, a cell has a marker for destruction. If a cell survives a mutation, there are still benefits to an organism. In fact, this concept is essentially the basis for evolution.

Genetic Code

DNA is important in terms of our genetic code, in the sense that it transfers genetic messages to all of the cells in your body. If you think about DNA in a reproductive sense, consider that the joining of an egg and sperm to create your first cell provided

your completed genetic code that your body would use for all of your life. Within that initial cell, half of your chromosomes - containing your DNA - came from your father and half came from your mother.

DNA clearly plays important roles in the human body and is one of the most significant discoveries of the twentieth century. Our continued research and knowledge of DNA functions will likely help us to learn even more about this important molecule.

Packaging of DNA

Packaging of DNA within the Cells

The length of DNA molecules existing in a particular cell is much longer than the long dimensions of the cell or the organelle where they exist. The contour length (i.e., its helix length) of a double stranded DNA can be calculated from the molecular weight, presuming the average molecular weight of each nucleotide pair to be about 650 Da and there is one nucleotide pair for every 0.36 nm of the duplex.

Accordingly the length of the smallest DNA is 1938.96 nm, belonging to the X174 virus (in duplex form), whose particle long dimension is 25 nm. On the other hand the total contour length of the entire DNA in a single human cell is about 2 metres and the cell nucleus is just 5-10 nm in diameter. These long DNA molecules are very tightly compacted, so as to fit into the cell. This packing is possible due to DNA getting further coiled into different fashions. The linear double helical DNA called the relaxed DNA, bends or twists upon its own axis, which is called as DNA coiling. This coiled DNA further coils upon itself to form the DNA supercoil, just like the telephone cord wire from the base of the phone to the receiver.

The degree of DNA supercoiling depends upon the type of the cell/organelle to be packed and is coiled in such a manner that DNA can easily be accessible to the enzymes/proteins for all of its functions like replication and transcription.

There are possibilities of two types of DNA supercoiling:

(a) Solenoidal, wherein the DNA coils in a spherical fashion upon itself and

(b) Plectonemic, wherein the DNA coils back upon its reverse length in the form of pleats. Further the supercoiling and packing of DNA differ in the prokaryotes (i.e., those lacking a true nuclear envelope) and in the eukaryotes (i.e., those having a nuclear membrane).

Packaging of viral DNA

Though the viral DNA is much smaller than a bacterial or eukaryotic DNA, its contour length is much bigger than the long dimensions of the viral particle in which they are found. The DNA of bacteriophage T2 is 3500 times its particle diameter. The long dimensions of different viral particle/the contour length of their DNAs respectively in nanometers is T2-210/65520; Lamdaphage-190/17460; Ti-78/14376 and φX174 (in duplex form)-25/1938. In order to get themselves packed within the particle, most

of these viral DNAs are linked covalently by the ends and, therefore, form an endless belt and thus become circular. Some of the single stranded DNA (φX174) becomes double stranded and circular.

Packaging of bacterial DNA

The length of E.coli cell is 2 micro-meters and its complete DNA molecule (the complete DNA molecule of an organism is called the chromosome) is 1.7 mm long, which exists as a single, covalently closed double stranded circular molecule, coiled and packed within nucleon of the cell.

This circular chromosome is organized in a scaffold-like structure, which folds the chromosome into looped domains. These domains are further coiled around some basic proteins, called Hu proteins (Mw = 10 000). In addition to the nucleosomal DNA the bacteria contains some small circular supercoiled non-chromosomal DNA-called plasmids.

Eukaryotic cells also contain other membrane-bound organelles such as mitochondria and the Golgi apparatus, and in addition, some cells of plants and algae contain chloroplasts.

Organization and packing of eukaryotic DNA

All the chromosomal DNA of an **eukaryotic cell** is embedded in a membranous cellular organelle called the nucleus. The eukaryotic DNA, in the nucleus is linear and not circular. In the non-dividing resting cell all the DNA of the cell forms a fine filamentous network in the nucleus called the chromatin. During cell division the chromatin network is subdivided into defined number and shaped chromosomes, their diploid number (pairs) depends upon the species of organism.

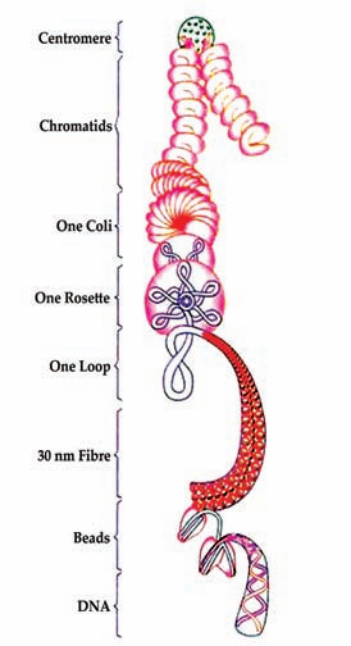
The normal chromosome number in humans is 46 (23 pairs). Each chromosome has a central axis called the centromere, from which two arms of DNA project out and each is referred to as chromatid. Each chromosome differs in size and shape within a given organism.

The 2 meter long eukaryotic human cell DNA is to be packed in the cell of about 5-10 micrometer in diameter. In order to facilitate its package, the helical DNA molecule is bound, tightly around beads of basic proteins called histones, which are spaced at regular intervals. The complexes of histones and DNA are called nucleosomes. Each nucleosome contains eight histone molecules, two copies each of H_2A, H_2B, H_3 and H_4 winding 146 base pairs of DNA.

Between the two nucleosomes there is a spacer DNA of 54 base pairs with a single histone molecule (H_1). Wrapping

DNA around a nucleosome compacts it to about seven-fold. These nucleosomes are organized very close together to form a structure, simply called 30 nm fibre. This provides 100-fold compaction of DNA. The 30 nm fibre then forms plectonemic pleats, called loops.

Six such loops are bounded by scaffold attachment to proteins (histone- H_1/topoisomerase-II) to give rise to a cluster of loops called rosette. 30 such rosettes bunch together to form a single coil, 10 such coils (like phone cord) forms a chromatid and the two chromatids are linked together by a highly repetitive base sequence rich in AT base pairs called satellite DNA, which is the centromere. Two chromatids with a centromere form a chromosome.



Packaging of eukaryotic DNA

The Importance of DNA

It has been argued that the discovery of DNA as well as our understanding of its structure and functioning may well be the most important discovery of the last century. The effect of the discovery of DNA on scientific and medical progress has been enormous, whether it involves the identification of our genes that trigger major diseases or the creation and manufacture of drugs to treat these devastating diseases. In fact, the identification of these genes and their subsequent analysis in terms of therapeutic treatment has ultimately influenced science and will continue to do so in the future.

Disease Diagnosis and Treatment

One important area of DNA research is that of genetics and medical research. Due to our discovery of DNA, our ability to actually diagnose diseases early on has been

vastly improved. In addition, we have been able to better assess a person's genetic susceptibility to specific diseases. In doing so, we have also paved the pathway to formulate brand new drugs to treat these diseases. In fact, drugs can essentially be custom made to complement a person's personal biochemistry and genetic makeup. For those diseases that were previously considered lethal and where treatment was either non-existent or largely unsuccessful, the discovery of DNA has essentially led to breakthrough drugs and treatments for patients with serious illnesses.

Paternity and Legal Impact

While the discovery of DNA has perhaps impacted medicine the most, its contribution to other areas is still similarly significant. Paternity cases have an enormous impact on families and children around the world. Through the assessment of DNA, the paternity of a child can be identified, which has a significant effect on the child's upbringing and his or her life.

Forensics and DNA

DNA has been notably important to the field of forensic science. The discovery of DNA has meant that the guilt or innocence of a person who is investigated for a crime can be determined. It also means that scarce evidence can still yield vital clues regarding the perpetrator of a crime. Also important is that the identification of victims can occur, particularly in cases where the victim's condition is unrecognisable to family or friends. In this sense, DNA has been important in revolutionising the entire field of forensic science. This impact is felt within the criminal justice system and contributes to the accurate safeguarding of society.

Agriculture and DNA

The impact of DNA on agriculture has been a very important one because it has allowed breeders to facilitate the breeding of animals that have a better resistance to diseases. It also allows farmers to produce more nutritious produce, which has particularly important consequences in developing countries where the population subsists on a small range of staple foods that have little variety. This means that micronutrient deficiencies can be addressed in these countries.

DNA Importance Continues

While the discovery of DNA has been a significant one in the twentieth century, it will continue to revolutionise medicine, agriculture, forensics, paternity and many other important fields in society today. DNA research encompasses an evolving area of progress and continued funding and interest in its relevance will likely fuel new discoveries in the future.

CONCEPT OF DNA REPLICATION

DNA replication is a crucial process that ensures the accurate duplication of the genetic information in a cell before cell division occurs. This fundamental biological process is essential for the transmission of genetic material from one generation to the next, ensuring genetic continuity and stability in all living organisms. DNA replication takes place during the interphase of the cell cycle, specifically in the S phase. It begins with the unwinding of the double helix structure of DNA by specialized enzymes called helicases, creating two separated strands. Each strand then serves as a template for the synthesis of a new complementary strand by DNA polymerases. As the enzymes move along the DNA molecule, they add nucleotides to the growing new strands according to the base-pairing rules (adenine with thymine and cytosine with guanine). This process results in the formation of two identical DNA molecules, each with one original and one newly synthesized strand. DNA replication is a highly precise and controlled process, ensuring the faithful transmission of genetic information from one generation of cells to the next and providing the foundation for the continuity of life.

SUMMARY

- DNA sequencing is another critical method that allows scientists to determine the exact order of nucleotides in a DNA molecule, revealing the genetic code and providing insights into an organism's genetic makeup.
- DNA serves as the blueprint of life, encoding the genetic information that shapes the characteristics and functions of all living organisms. The analysis of DNA has revolutionized numerous fields, from medicine and forensics to evolutionary biology and genetics.
- A long polymer, DNA is made up of smaller units called nucleotides. In turn, each nucleotide consists of a phosphate group, a sugar and a nitrogenous base.
- DNA can exist in different conformations and these are important for a range of DNA mechanisms. These conformations interact with enzymes in your body and are also involved in aspects such as DNA repair.
- The pairing between adenine and thymine and between guanine and cytosine results in a complementary relationship between the sequence of bases on the two intertwined chains and gives DNA its self-encoding character.
- The hydrogen bonds between complementary bases are a fundamental feature of the double helix, contributing to the thermodynamic stability of the helix and providing the information content and specificity of base pairing.
- DNA replication is vital for a virtually endless list of functions, from reproduction to maintenance and growth of cells, tissues and body systems. To copy itself, a DNA molecule essentially 'unzips,' thus resulting in a series of bases without pairs along the backbone of the molecule.

MULTIPLE CHOICE QUESTIONS

1. DNA fingerprinting recognizes the differences in

a. satellite DNA
b. bulk DNA
c. Repetitive DNA
d. both (a) and (c)

2. This force can stabilize a DNA double-helix

a. Hydrophilic sugar-phosphate groups are found on the exterior of the helix where interaction with water occurs
b. Hydrophobic bases are present in the interior of the helix, each base-pair is stabilized by the same number of hydrogen bonds
c. covalent base stacking interactions may take place between neighbouring bases within the same strand in the helix
d. non-covalent N-glycosidic bonds may form between nitrogenous bases in opposite strands in the helix

3. In this type of DNA replication, of the two newly formed molecules, one is purely a new one and the other is an old one

a. dispersive
b. conservative
c. semiconservative
d. both (b) and (c)

4. The process of DNA replication is affected by an enzyme known as

a. Mutase
b. Ligase
c. Polymerase I
d. Ribonuclease

5. A DNA molecule in which both strands have radioactive thymidine is permitted to replicate in an environment that contains non-radioactive thymidine. What is the right number of DNA molecules which possess some radioactive thymidine post three duplications?

a. one such molecule
b. two such molecules
c. four such molecules
d. eight such molecules

Answer

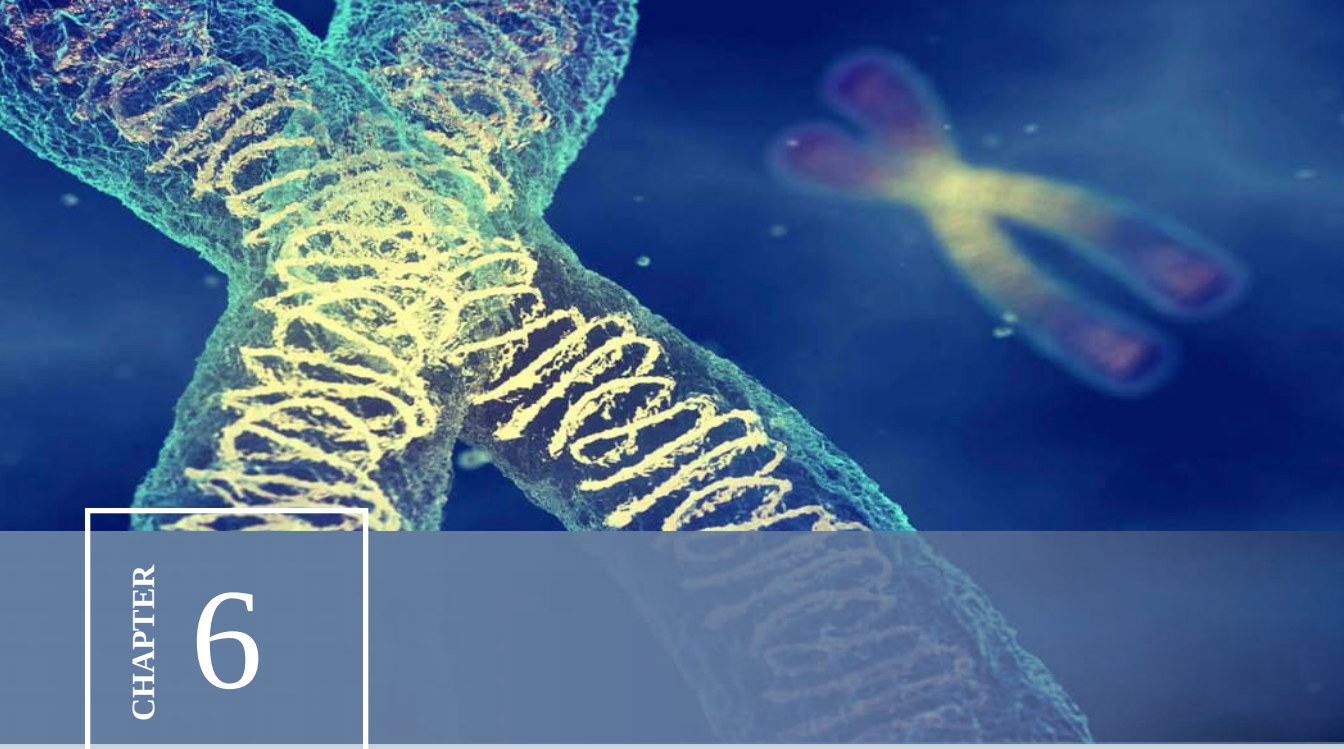
1. (d) 2. (a) 3. (c) 4. (c) 5. (b)

REVIEW QUESTIONS

1. Discuss the properties of DNA.
2. Why are hydrogen bonding is important for the specificity of base pairing?
3. What is the role of phosphodiester bonds within the sugar-phosphate backbone of DNA?
4. In considering the structure of the DNA double helix, how would you expect the structure to differ if there was base pairing between two purines? Between two pyrimidines?
5. What do you understand by genetic code?

REFERENCES

1. CrashCourse (2012). DNA Structure and Replication: Crash Course Biology #10. [video] Available at: https://www.youtube.com/watch?v=8kK2zwjRV0M [Accessed 11 Apr. 2017].
2. Dahm, R. (2005). Friedrich Miescher and the discovery of DNA. [online] ScienceDirect. Available at: http://www.sciencedirect.com/science/article/pii/S0012160604008231 [Accessed 11 Apr. 2017].
3. Doherty, E. A. and Doudna, J. A. 2001. Ribozyme structures and mechanisms. Ann. Rev. Biophys. Biomol. Struct. 30:457–475.
4. Krude, Torsten, ed. DNA: Changing Science and Society. Cambridge: Cambridge University Press, 2004.
5. Rennie, John. "Celebrating the Genetic Jubilee: A Conversation with James D. Watson." Scientific American, (April 2003): 66-69.
6. Stahl, Franklin W. (ed). We Can Sleep Later. Alfred D. Hershey and the Origins of Molecular Biology. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2000.
7. Wang, J. C. 2002. Cellular roles of DNA topoisomerases: A molecular perspective. Nat. Rev. Mol. Cell Biol. 3:430–440.



CHAPTER 6

LINKAGE AND CHROMOSOME MAPPING



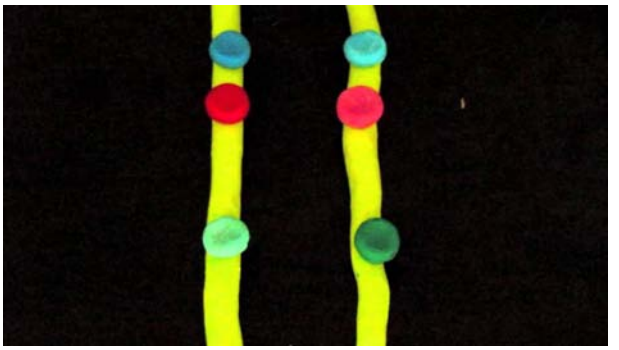
Linkage refers to the presence of two different genes on the same chromosome. Two genes that occur on the same chromosome are said to be linked, and those that occur very close together are tightly linked. Study of linkage provides information about the relative position of genes on chromosomes, allowing the construction of chromosome maps.

After studying this chapter, you will understand the core concepts of:

- Genetic Linkage
- Chromosome Mapping

INTRODUCTION

Genetic linkage is the tendency of DNA sequences that are close together on a chromosome to be inherited together during the meiosis phase of sexual reproduction. Two genetic markers that are physically near to each other are unlikely to be separated onto different chromatids during chromosomal crossover, and are therefore said to be more linked than markers that are far apart. In other words, the nearer two genes are on a chromosome, the lower the chance of recombination between them, and the more likely they are to be inherited together. Markers on different chromosomes are perfectly unlinked, although the penetrance of potentially deleterious alleles may be influenced by the presence of other alleles, and these other alleles may be located on other chromosomes than that on which a particular potentially deleterious allele is located. Chromosome mapping, on the other hand, is the process of determining the positions of genes on chromosomes. By analyzing the patterns of genetic linkage and recombination, scientists can create genetic maps that show the order and distances between genes on a chromosome. These maps provide valuable insights into the organization of genetic material and are crucial for understanding the inheritance patterns of traits, as well as for identifying the genetic basis of diseases and conducting genetic research. Linkage and chromosome mapping are indispensable tools in genetics, contributing to our understanding of the genetic basis of traits and diseases and facilitating advancements in various fields, from agriculture and biotechnology to medicine and evolutionary biology.



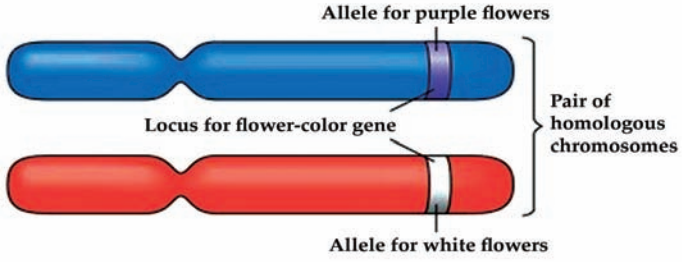


CORE CONCEPT: GENETIC LINKAGE

Genetic linkage is the coupling of two genes' patterns of inheritance because they are located on the same chromosome. Linked genes sit close together on a chromosome, making them likely to be inherited together (left). Genes on separate chromosomes are never linked (center). But not all genes on a chromosome are linked. Genes that are farther away from each other are more likely to be separated during a process called homologous recombination (right).

Two genes are said to be linked if they are located on the same chromosome. We assume that different chromosomes segregate independently during meiosis. Mendel's second law, the law of independent assortment, concerns the random separation of different genes carried on different chromosomes. During meiosis, each chromosome in a homologous pair makes its way into a gamete. Which chromosome from each homologous pair is sorted together into the same gamete is determined by pure chance, so characters on different chromosomes are inherited independently? This is the basis of the law of independent assortment. But what happens to genes that are on the same chromosome? These genes are physically linked, so they kind of have to go together into the same **gamete**. They are not inherited independently of one another like genes on different chromosomes, so as well as being physically linked, they are also said to be genetically linked.

Gamete is an egg (female gamete) or a sperm (male gamete). This is an example of anisogamy or heterogamy, the condition in which females and males produce gametes of different sizes.



Different forms of the same gene, called alleles, are present on matching, or homologous, chromosomes in similar positions, or loci. For instance, in Gregor Mendel's

experiments with peas, green and yellow are two alleles for pod color. In a heterozygote, which has both alleles, the two alleles occupy the same loci on homologous chromosomes. Similarly, round and wrinkled are alleles for seed texture. In the pea, these two genes—pod color and seed texture—are on different pairs of homologs and are therefore not linked. When gametes form in double heterozygotes (for example, a green/yellow–round/wrinkled plant), these genes assort independently, because the two chromosomes that bear them assort independently. Therefore, meiosis will create equal numbers of green-round, green-wrinkled, yellow-round, and yellow-wrinkled gametes. Mating between double heterozygotes (called a dihybrid cross) will give a characteristic ratio of the different possible plant types.

DNA sequencing has established the order of variants on a chromosome, so genetic linkage mapping essentially aims to measure distances in each interval (between two adjacent variants) on a chromosome.

However, if the two traits were located close to one another on the same chromosome—in other words, if they were linked—the observed ratio will be quite different from that seen for unlinked traits. Allele combinations that began together (for instance, round-green) will tend to stay together, and the offspring will show a skewed ratio reflecting the original combinations.



Despite being on the same chromosome, the round and green alleles could become separated during meiosis by crossing over, a form of genetic recombination. During crossing over, homologous chromosomes exchange segments. This could allow the yellow allele to switch places with the green allele and lead to a round-yellow gamete. If the loci for the two genes are very close, crossing over is unlikely to separate alleles, whereas if they are far apart, crossing over is much more likely to separate them. Therefore, the frequency of crossing over is related to the physical distance between the loci for the two genes. The particular combination of alleles on the homologous chromosomes in the dihybrid parent (for example, round-green) is known as linkage phase. Separation of this combination by crossing over is said to be a change in phase. The two alleles of a particular gene are said to be markers for that site of the chromosome.

Linkage in Fruit Flies

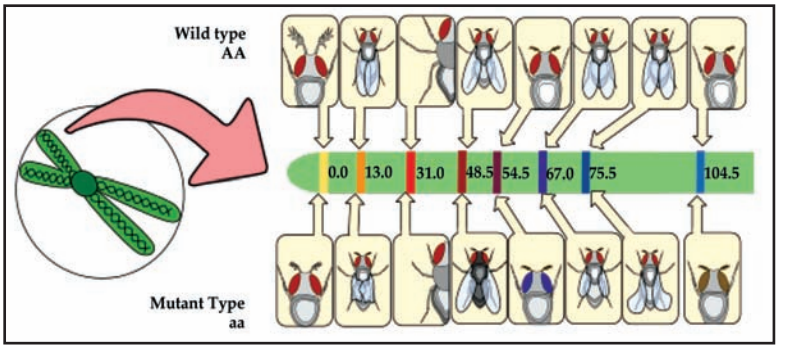
An example of using linkage to explore gene position is provided by inheritance of eye color and body color in fruit flies, both of which are located on the X chromosome. This example begins with purebred (homozygous) parents, one yellow-bodied and red-eyed, the other grey-bodied and white-eyed. They mate to produce all heterozygous daughters, who carry the yellow-red combination on one homologous chromosome and the grey-white combination on the other. When the heterozygotes create gametes, the eye-color alleles cannot assort independently from the body-color alleles because they are linked. Some crossing over can occur, though. As in humans, male fruit flies carry only one X chromosome, and so will show exactly what alleles are present on their X. When one counts the male offspring, approximately 49.5 percent are yellow-bodied and red-eyed, 49.5 percent are grey-bodied and white-eyed, 0.5 percent are yellow-bodied and white-eyed, and 0.5 percent are grey-bodied and red-eyed. This indicates very tight linkage—close proximity—of the two genes.



In this example, the yellow-body allele and the white-eye allele are said to be "out of phase" in the parental strains. The most frequent pair of gamete types are described as "parental types" because they retain the alleles for the two genes as transmitted by the original parent strains. The two gamete types that are less frequent are the "recombinant types," which results only from an exchange or crossover of homologous chromosomes in the interval between the genes.

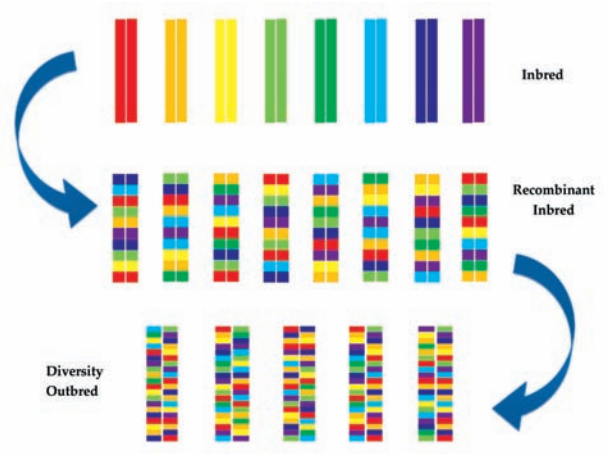
Linkage Map

A genetic linkage map is a tool used in genetic research to help scientists map the order of particular traits as they appear in a strand of deoxyribonucleic acid (DNA). It does not map distance between the traits on the DNA strand. Instead, it maps how frequently they will recombine. A genetic linkage map can help geneticists to find the genes that cause genetic disorders. Normally, each cell in a human, plant, or animal carries two copies of each gene. Sex cells, such as eggs, sperm, and spores, undergo a process called meiosis, or cell division, which splits the cell in half. This leaves only one thread of DNA in each cell.



Genetic recombination happens before the cell splits. First the chromosomes form two lines along the middle of the cell, creating pairs of genes. Sometimes the chromosomes break in half. Broken pieces are then combined to produce new molecules. This process, known as genetic crossover, happens to each human chromosome an average of 1.5 times for each sex cell that is formed. This exchange of genetic material is what allows offspring to have different traits from their parents. It is why two brown-haired parents can produce a blond child. The process of genetic recombination allows a species to genetically adapt to its environment over time. It can also lead to genetic diseases.

Geneticists can use a genetic linkage map to find where certain traits appear. These traits are called DNA markers. Any inherited trait can be a DNA marker as long as it differs from person to person and can be detected easily in a lab. Mapping known traits, such as hair and eye color, can help scientists extrapolate where other traits might appear. The physical distance between traits on a genetic linkage map is determined by the frequency of recombination. DNA markers with a higher frequency of recombination are shown as far apart. Markers with a lower frequency are close together.

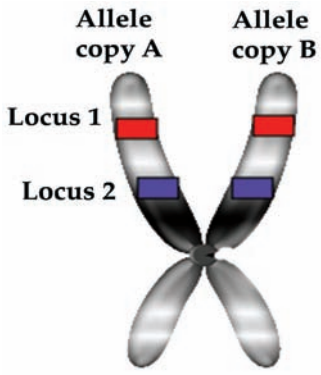


This works because genes that are close together are less likely to be separated by recombination. For instance, if the gene for blue eyes is next to the gene for blond hair on the genetic linkage map, then children with blue eyes are likely to also have blond hair. If, however, there are a lot of other genes between blue eyes and blond hair, the

confusion of genetic crossover may separate these traits, causing blue eyes and brown hair. These are only assumptions, however. Genetic linkage maps are often used as a framework for physical maps, which are detailed maps of the DNA sequence. They allow scientists to rapidly identify particular genes.

Linkage Disequilibrium

In the field of genetics, population genetics studies how often specific copies of genes — or alleles — appear and how they interact with other alleles from other genes. It also looks at how processes of evolution such as natural selection and mutations affect alleles. The distribution of alleles can be either independent of other alleles at a different location, or it may be affected by other alleles. When the distribution of an allele, or alleles, is not independent of other alleles, it is called linkage disequilibrium.



Evolution of an organism occurs through genetic variation. Different combinations of genes, and different combinations of alleles of different genes, lead to different phenotypes of organisms. This means that genetic variation produces organisms of a species that are similar but can look different and function differently. Genetic variation can make an organism more or less successful in terms of survival and reproduction. The theory of natural selection states that evolution occurs through survival of the fittest, or survival of those that are more successful at passing on their genes.

Genes are not independent units that are passed on by themselves to offspring. Instead, genes are affected by the environment as well as by other genes. Genetic coadaption is the term used to refer to how well genes interact with other genes. Natural selection favors those alleles that interact well with alleles of other genes at different locations within the DNA. Genetic coadaption can exist between the alleles of some genes but not others.

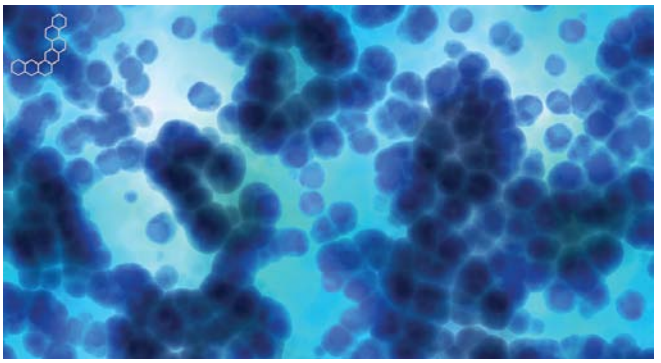
For example, gene A and gene B each have two alleles, which are A1 and A2 and B1 and B2, respectively. If an organism inherits A1 and B1 or A2 and B2, it can be presumed in this example to be better adapted to survive than if it inherits A1 and B2 or A2 and B1. This means that through natural selection, the genetic combinations of A1B1 and A2B2 would be favored as the organism with these alleles is more likely to survive and reproduce. As such, these alleles are said to be in linkage disequilibrium as they would not be passed on randomly to offspring, but instead, certain pairs are

Zygote is a eukaryotic cell formed by a fertilization event between two gametes.

favored. Linkage disequilibrium is measured by comparing the frequencies of alleles within a population, but it is not permanent. Random mating between organisms can lead to a decrease in the occurrence of the linked alleles. Permanent linkage disequilibrium can occur if one combination results in a higher fitness level of an organism, such as when the other combination results in a lethal mutation in a **zygote**. The further apart the locations of the alleles, the more difficult it is to maintain linkage disequilibrium.

Linkage Analysis

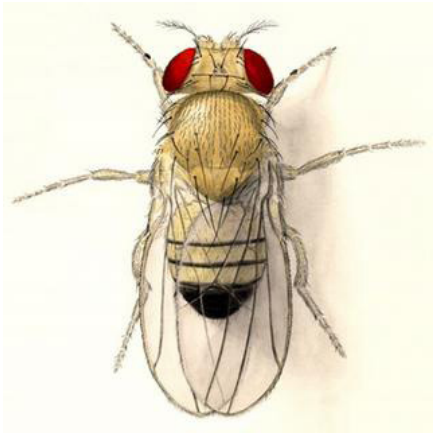
Linkage analysis is a well-established statistical method for mapping heritable trait genes to their chromosome locations. Genome-wide markers are tested in pedigrees segregating a trait. The statistical method of linkage analysis combines these data to identify chromosome regions likely to harbor genes for the trait. Parametric linkage analysis is used for traits with a Mendelian form of inheritance. LOD scores and recombination fractions are used to test the gene locations. Model-free linkage analysis is used for complex traits, where the model of inheritance is not known. Linkage analysis can be used to map genes for both binary and quantitative traits.



Linkage Analysis in Drosophila

Drosophila also known as fruit fly were most widely used and genetically best-known of all eukaryotic organisms for genetic research to Thomas Hunt Morgan in 1909, because of many factors like, fruit flies had a generation time of 10 days (from egg to a mature adult). The organisms are very small so a large number of them could be kept and bread for experimental purposes. The easiest way to generate large numbers is by

growing them on a medium made from instant mashed potatoes with a little yeast. Drosophila grows between the temperatures of 18 - 28°C. It grows best at 22 -26°C. At this temperature, the entire life cycle requires 8 - 14 days for one **generation**.



Generation is "all of the people born and living at about the same time, regarded collectively." It can also be described as, "The average period, generally considered to be about thirty years, during which children are born and grow up, become adults, and begin to have children of their own."

By 1915, Morgan and his students had discovered about 85 different mutants, which when isolated were found to belong to four different linkage groups, one of which contained very few mutant genes. This finding related to the fact that drosophila cell have four pairs of homologous chromosomes one of which is very small. Still the linkage between the alleles on the same chromosome were found to be incomplete, i.e. the alleles of two different genes say for example short wings and black body, which were originally present on one given chromosome were not found to be present together during the production of gametes. The conclusion that was drawn out by this observation was that two characteristics that were present together on the same chromosome could get separated from one another and go to separate gamete. Conversely, it could be said that the maternal and parental characteristics inherited by an individual on separate homologous chromosomes could get rearranged and might land upon the same chromosome of a gamete.



In 1909, F. A. Janssens observed the same phenomenon where the homologous chromosomes of the bivalents got wrapped around each other during the stage of meiosis. Janssens then proposed interaction between the maternal and paternal chromosomes which resulted in the breakage and exchange of characteristics. This result in 1911, made Morgan offer an explanation of the "breakdown" in the linkage which could account for the appearance of unexpected combinations of genetic traits in the offspring. Morgan termed this phenomenon as the Crossing Over (or genetic recombination**)**. Analysis of offspring from a large number of crosses between adults with a variety of alleles on the same chromosome indicated that:

- The percentage of recombination between a given pair of genes on a chromosome such as eye color and wing length, was essentially constant from one experiment to another.
- The percentage of recombination between different pairs of genes such as between eye color and wing length, with eye color and body color, could be very different.

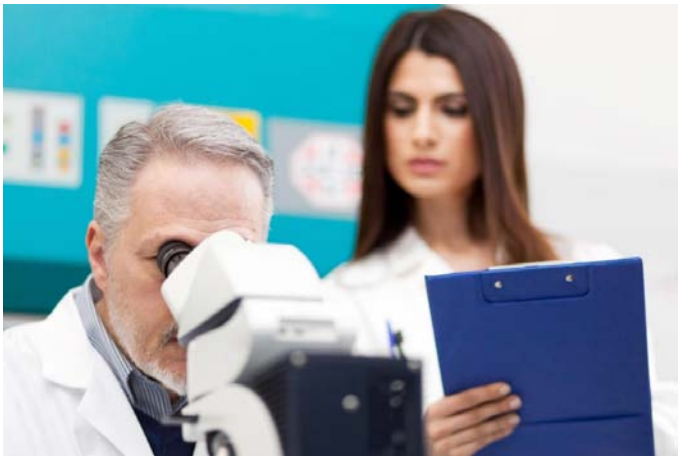
The indication that a given pair of genes gave the same approximate frequency of reconstruction in every cross suggested that the positions of the genes along the chromosome (their loci) were fixed and that their positions did not vary from one fly to another. Thus if the locus of each gene is fixed, then the frequency of recombination between two genes will provide a measure of distance that separates those two genes.

LOD Score

A LOD score is a representation of the likelihood of a linkage between two genetic traits. If the score is high, it means that the traits are closely linked, and therefore usually inherited together. Low scores, on the other hand, indicate a low linkage. Understanding these numbers is important to geneticists for a number of reasons, ranging from wanting to understand particular genetic conditions to a desire to figure out where a gene is located, and using information about known genes. To understand the importance of the LOD score, it helps to know that the closer two gene loci are to each other, the more likely it is that they will be inherited together. If genes are on different chromosomes, for example, they are usually inherited separately, because a great deal of crossover can occur during meiosis. Genes that are on the same chromosome but far apart are also usually inherited independently. When loci are in close proximity on the same chromosome, however, they tend to be inherited together. This information is at the root of the scoring system.

LOD stands for "logarithm of the odds," reflecting the fact that the number is expressed logarithmically. As a general rule, if the score is 3 or higher, it means that there is a high probability of genetic linkage. Lower numbers indicate less likelihood of genetic linkage, although they can still be useful in the process of elimination. To determine a LOD score, a data sample is needed, and in this type of statistical analysis, the bigger the sample, the better. Researchers working with organisms like fruit flies may actively breed a population to bring out or suppress desired traits for the purpose of studying the linkage between specific traits, while with populations

such as humans, researchers must confine themselves to data collected through observation. Large samples tend to yield more statistically meaningful information because they reduce the risk of flukes and clusters which skew the data.

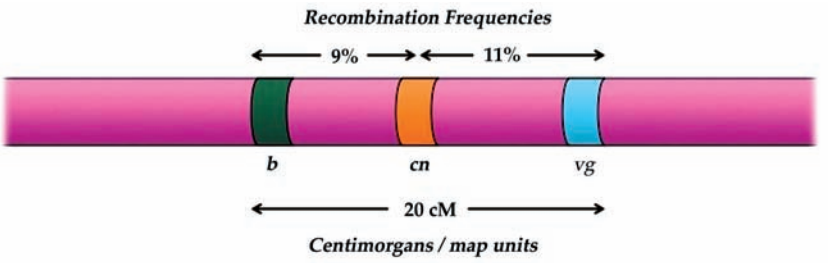


Using data from a population of interest, researchers look at how often certain traits occur together, and how often they emerge separately. This information is used to arrive at the recombination frequency, looking at how often genetic traits recombine. To be on the same chromosome, the recombination frequency must be less than 50%, and it must decline even more to indicate that loci are close to each other. Researchers analyze this information to determine how likely it is that there is a genetic linkage or not. The LOD score is determined by dividing the likelihood of linkage by the likelihood of no linkage. Typically, scores are calculated for multiple loci.

Recombination Mapping

Recombination, in genetics, regrouping of the maternal and paternal genes during the formation of gametes (sex cells). Recombination occurs randomly in nature as a normal event of meiosis, the process by which gametes are produced. Recombination is enhanced by the phenomenon of crossing over, in which gene sequences called linkage groups are disrupted, resulting in an exchange of segments between paired chromosomes that are undergoing separation. Thus, although a normal daughter cell produced in meiosis always receives half of the genetic material contained in the parent cell (i.e., is haploid), recombination acts to ensure constant variability: no two daughter cells are identical, nor are any identical in genetic content to the **parent cell**.

Parent cell is the cell that divides to give rise to two daughter cells. In mitosis, the two daughter cells contain the same genetic content as the parent cell. In meiosis, the daughter cells have different genetic content and half the number of chromosomes of the parent cell.



Laboratory study of recombination has contributed significantly to the understanding of genetic mechanisms, allowing scientists to map chromosomes, identify linkage groups, isolate the causes of certain genetic anomalies, and manipulate recombination itself by transplantation of genes from one chromosome to another. Because of its potential for creating new—and possibly pathogenic—organisms, experimental recombination is viewed by some scientists as both dangerous and unethical.

- Genetic Linkage and Independent Assortment

  The position of the genes in a chromosome exerts vital effect in its inheritance. Genes on different chromosomes or located farer on a same chromosome are said to be unlinked and assort independent of each other, known as independent assortment. On the contrary, location of genes on the same chromosome causes them to be linked and inherited together, known as genetic linkage. The lesser the distance between 2 genes, the lesser the chances of recombination of them. Hence, the proportion of recombinant offspring gives an idea of the distance between the 2 genes.

- Recombination Frequency and Linkage Mapping

  The frequency of a single chromosomal crossover between 2 genes at the stage of meiosis is known as recombination frequency. In meiosis, crossover occurs at random positions; thus, the closely located genes show less chances of crossover. Hence, on the basis of recombinant frequency, a genetic map showing the position of genes on a chromosome can be drawn known as linkage mapping.

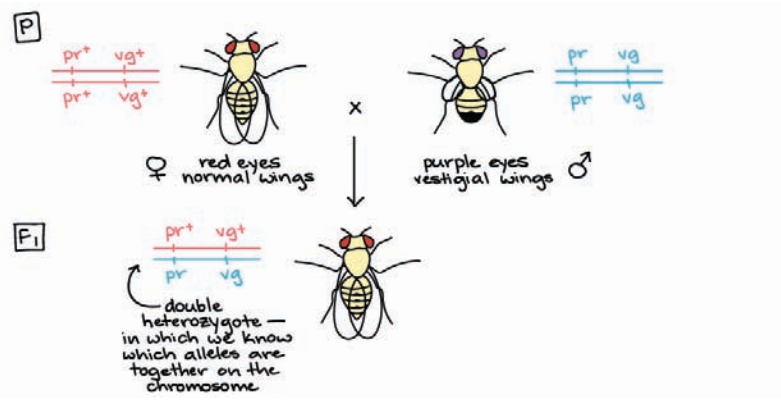
Finding Recombination Frequency

Let's suppose we are interested in seeing whether two genes in the fruit fly (*Drosophila*) are linked to each other, and if so, how tightly linked they are. In our example, the genes are:

- The *purple* gene, with a dominant pr^+ allele that specifies normal, red eyes and a recessive *pr* allele that specifies purple eyes.
- The *vestigial* gene, with a dominant vg^+ allele that specifies normal, long wings and a recessive *vg* allele that specifies short, «vestigial» wings.

If we want to measure recombination frequency between these genes, we first need to construct a fly in which we can observe recombination. That is, we need to make a fly that is not just heterozygous for both genes, but where we know exactly

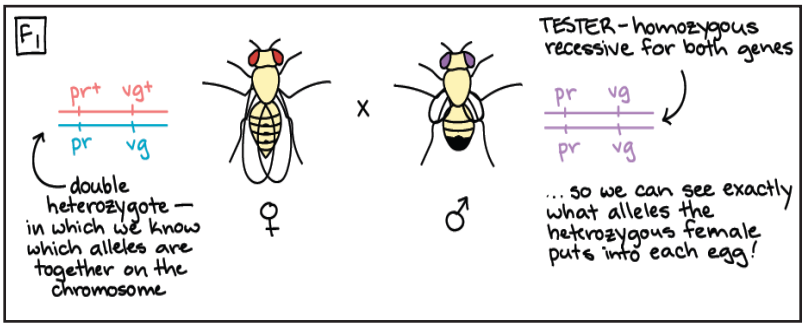
which genes are together on the chromosome. To do so, we can start by crossing two homozygous flies as shown below:



This cross gives us exactly what we need to observe recombination: a fly that's heterozygous for the *purple* and *vestigial* genes, in which we know clearly which alleles are together on a single chromosome. Now, we need a way to "see" recombination events. The most direct approach would be to look into the gametes made by the heterozygous fly and see what alleles they had on their chromosomes. Practically, though, it is much simpler to use those gametes in a cross and see what the offspring look like.

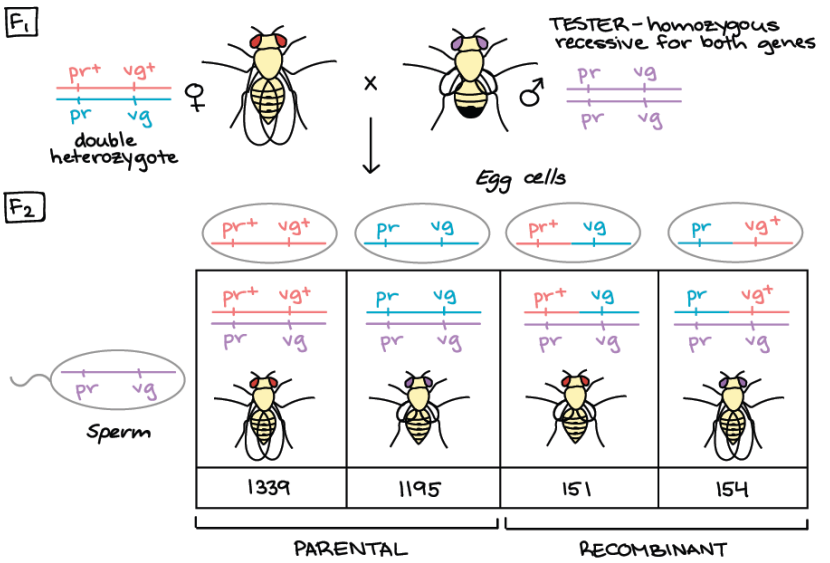
To do so, we can cross a double heterozygous fly with a tester, a fly that's homozygous recessive for all the genes of interest (in this case, the *pr* and *vg* alleles). The purpose of using a tester is to ensure that the alleles provided by the non-tester parent fully determine the phenotype, or appearance, of the offspring. When we cross our fly of interest to a tester, we can directly "read" the genotype of each gamete from the physical appearance of the **offspring**.

Offspring are the young born of living organisms, produced either by a single organism or, in the case of sexual reproduction, two organisms.



Below, we can see a modified Punnett square showing the results of the cross between our double heterozygous fly and the tester fly. Four different types of eggs are produced by a double heterozygous female fly, each of which combines with

a sperm from the male tester fly. Four different phenotypic (appearance-based) classes of offspring are produced in this cross, each corresponding to a particular gamete from the female parent:



The four classes of offspring are *not* produced in equal numbers, which tells us that the *purple* and *vestigial* genes are linked. As we expect for linked genes, the parental chromosome configurations are over-represented in the offspring, while the recombinant chromosome configurations are under-represented. To measure linkage quantitatively, we can calculate the recombination frequency (RF) between the *purple* and *vestigial* genes:

$$\text{Recombination frequency (RF)} = \frac{\text{Recombinants}}{\text{Total offspring}} \times 100\%$$

In our case, the recombinant progeny classes are the red-eyed, vestigial-winged flies and the purple-eyed, long-winged flies. We can identify these flies as the recombinant classes for two reasons: one, we know from the series of crosses we performed that they must have inherited a chromosome from their mother that had undergone a recombination event; and two, they are the underrepresented classes (relative to the overrepresented, parental classes). So, for the cross above, we can write our equation as follows:

$$\text{RF} = \frac{151 + 154}{1339 + 1195 + 151 + 154} \times 100\% = 10.7\%$$

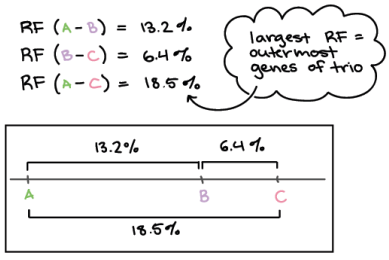
The recombination frequency between the purple and vestigial genes is 10.7%.

Recombination frequency and linkage maps

What is the benefit of calculating recombination frequency? One way that recombination frequencies have been used historically is to build linkage maps, chromosomal maps based on recombination frequencies. In fact, studying linkage

helped early geneticists establish that chromosomes were in fact linear, and that each gene had its own specific place on a chromosome. Recombination frequency is not a direct measure of how physically far apart genes are on chromosomes. However, it provides an estimate or approximation of physical distance. So, we can say that a pair of genes with a larger recombination frequency are likely farther apart, while a pair with a smaller recombination frequency are likely closer together.

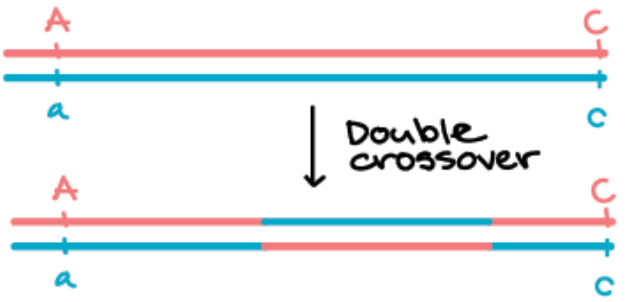
Importantly, recombination frequency "maxes out" at 50%, percent (which corresponds to genes being unlinked, or assorting independently). That is, 50%, percent is the largest recombination frequency we will ever directly measure between genes. So, if we want to figure out the map distance between genes further apart than this, we must do so by adding the recombination frequencies of multiple pairs of genes, "building up" a map that extends between the two distant genes. Comparison of recombination frequencies can also be used to figure out the order of genes on a chromosome. For example, let's suppose we have three genes, *A*, *B*, and *C*, and we want to know their order on the chromosome (*ABC*? *ACB*? *CAB*?) If we look at recombination frequencies among all three possible pairs of genes (*AC*, *AB*, *BC*), we can figure out which genes lie furthest apart, and which other gene lies in the middle. Specifically, the pair of genes with the largest recombination frequency must flank the third gene:



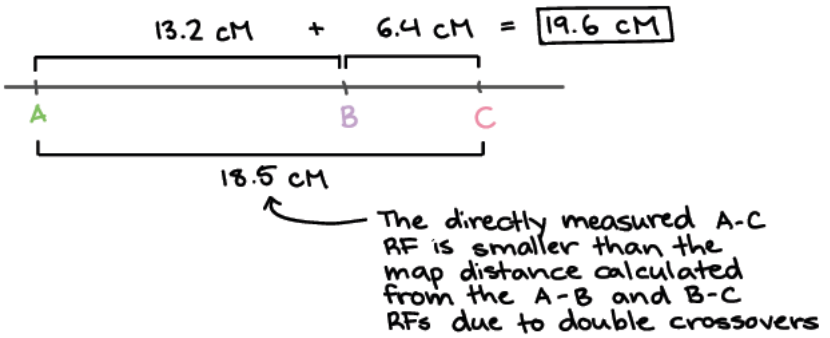
By doing this type of analysis with more and more genes (e.g., adding in genes *D*, *E*, and *F* and figuring out their relationships to *A*, *B*, and *C*) we can build up linkage maps of entire chromosomes. In linkage maps, you may see distances expressed as centimorgans or map units rather than recombination frequencies. Luckily, there's a direct relationship among these values: a 1%, percent recombination frequency is equivalent to 1 **centimorgan** or 1 map unit. Is map distance always the same as recombination frequency? Sometimes, the directly measured recombination frequency between two genes is not the most accurate measure of their map distance. That's

Centimorgan is a unit for measuring genetic linkage. It is defined as the distance between chromosome positions for which the expected average number of intervening chromosomal crossovers in a single generation is 0.01.

because, in addition to the single crossovers we have discussed in this section, double crossovers (two separate crossovers between the two genes) can also occur:



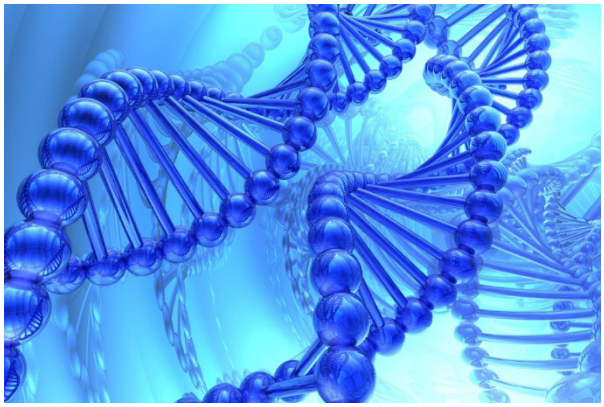
Double crossovers are "invisible" if we are only monitoring two genes, in that they put the original two genes back on the same chromosome (but with a swapped-out bit in the middle). For example, the double crossover shown above would not be detectable if we were just looking at genes *A* and *C*, since these genes end up back in their original configuration. Because of this, double crossovers are not counted in the directly measured recombination frequency, resulting a slight underestimate of the actual number of recombination events. This is why, in the example below, the recombination frequency directly measured between *A* and *C* is a bit smaller than the sum of the recombination frequencies between *A*-*B* and *B*-*C*. When *B* is included, double crossovers between *A* and *C* can be detected and accounted for.



By measuring recombination frequencies for closer-together gene pairs and adding them up, we can minimize "invisible" double crossovers and get more accurate map distances.

Modern Applications

Sturtevant's discovery led to the golden age of chromosome transmission genetics, with an emphasis on identifying genes through alleles with visible phenotypes, and using them as markers for determining their position on the linkage map. Since then the emphasis in genetics has shifted to understanding the functions of genes. Linkage and gene mapping studies have progressed to being a critical tool in cloning genes and providing more description of their roles in the organism.



These approaches include:

- Using map locations to distinguish different genes with similar sequences, mutant phenotypes, or functions. Examples are the cell division cycle mutants of the yeast *Saccharomyces cercvisiae* or the uncoordinated mutants of the roundworm *C. elegans.* In some cases mutants with different phenotypes have been shown to be done to different mutations in the same gene, as is the case with the *Drosophila* circadian rhythm period mutants termed short, long, and none (per[S], per[L] and per[0]).
- Using map locations to track down genes to clone their deoxyribonucleic acid (DNA) by chromosome position. Examples are the human cystic fibrosis transmembrane regulator gene mutated in cystic fibrosis, or the polyglutamine repeat gene that is mutated in Huntington's disease. With genome sequences available on databases, mapping mutant phenotypes points to candidate loci for the gene at the chromosome position.

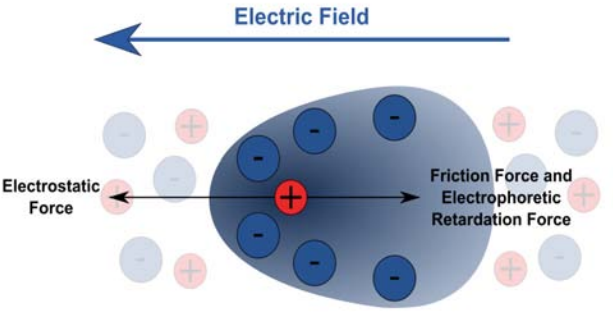
New classes of markers in linkage analysis are based on naturally occurring DNA variation in the genome, and have many advantages. These variations are usually harmless and do not interrupt a gene, so there is no selection against them, meaning they persist over many generations. They are quite numerous and are distinguished throughout in the genome. Individuals are likely to be heterozygous from many of them and therefore the markers are informative for linkage. If the DNA variant is present heterozygously, can be detected, and shows Mendelian segregation, it is as good a linkage marker as yellow bodies or white eyes. The disadvantage is that analysis to detect the variant is sometimes more laborious and requires the techniques of molecular biology.

The common types of DNA markers and the molecular techniques used to follow their inheritance are:

- Restriction fragment length polymorphisms (RFLPs) are derived from sequence variation that results in the loss of a restriction enzyme digestion site. The result is a longer fragment of the DNA from that location following digestion with that enzyme. A heterozygous parent will transmit either the allele specifying the long fragment or the allele specifying the short fragment to each child. After size separation of DNA fragments by

Electrophoresis is used in laboratories to separate macromolecules based on size. The technique applies a negative charge so proteins move towards a positive charge.

gel **electrophoresis** and transfer to a Southern blot, these DNA fragments of interest can be identified with a specific DNA or ribonucleic acid (RNA) probe that also comes from that location. If the long fragment, for example, is linked to a disease gene, the child's DNA can reveal if he or she is likely to develop the disease.



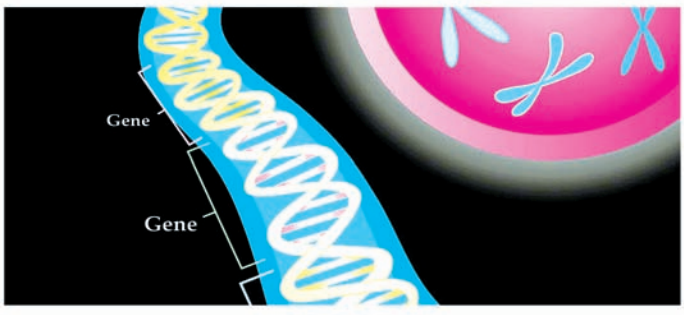
- Randomly amplified polymorphic DNAs (RAPDs) are derived from sequence variation that results in the loss of the complementary site to a primer necessary to initiate chain amplification by polymerase chain reaction (PCR). If the DNA used as template contains complementary sites for both primers, a PCR product is obtained that can be detected by gel electrophoresis. If either site is absent or changed in the template no product will be obtained from the reaction.

Genetic Marker

Genetic marker, any alteration in a sequence of nucleic acids or other genetic trait that can be readily detected and used to identify individuals, populations, or species or to identify genes involved in inherited disease. Genetic markers consist primarily of polymorphisms, which are discontinuous genetic variations that divide individuals of a population into distinct forms (e.g., AB versus ABO blood type or blond hair versus red hair). Genetic markers play a key role in genetic mapping, specifically in identifying the positions of different alleles that are located close to one another on the same chromosome and tend to be inherited together. Such linkage groups can be used to identify unknown genes that influence disease risk. Technological advances, especially in DNA sequencing, have greatly increased the catalogue of variable sites in the human genome.

Multiple types of polymorphisms serve as genetic markers, including single nucleotide polymorphisms (SNPs), simple sequence length polymorphisms (SSLPs), and restriction fragment length polymorphisms (RFLPs). SSLPs include repeat sequences, variations known as minisatellites (variable number of tandem repeats, or VNTRs) and microsatellites (simple tandem repeats, STRs). Insertions/deletions (indels) are another example of a genetic marker.

In the human genome, the most common types of markers are SNPs, STRs, and indels. SNPs affect only one of the basic building blocks—adenine (A), guanine (G), thymine (T), or cytosine (C)—in a DNA segment. For example, at a genomic location with the sequence ACCTGA in most individuals, some persons may contain ACGTGA instead. The third position in this example would be considered an SNP, since there is a possibility of either a C or a G allele occurring in the variable position. Because every individual inherits one copy of DNA from each parent, every person has two complementary copies of DNA. As a result, in the above example, three genotypes are possible: homozygous CC (two copies of the C allele at the variable position), heterozygous CT (one C and one T allele), and homozygous TT (two T alleles). The three genotype groups can be used as "exposure" categories to assess associations with an outcome of interest in a genetic epidemiology setting. Should such an association be identified, researchers may investigate the marked genomic region further to identify the particular DNA sequence in that region that has a direct biological effect on the outcome of interest.



STRs are markers in which a piece of sequence is repeated several times in a row, and the number of repeats (considered an allele) is variable within and across individuals. For example, a CCT pattern may be repeated up to 10 times, such that individuals in the population may have genotypes at that locus (chromosomal location) representing any combination of two repeat alleles of sizes 1 to 10 repeats (e.g., $10(10 + 1)/2 = 55$ different possible genotypes). Indels are polymorphisms in which a piece of DNA sequence exists in some versions (insertion allele) and is deleted in others (deletion allele) in the population.

Estimation of the recombination fraction

Recombination fractions are estimated from the proportion of recombinant gametes. This is relatively easy to determine if we know

- Linkage phase in parents

- The haplotype of the gamete that was transmitted from parent to offspring

If the linkage phase is known in parents, we know can know which gametes are recombinants, and which ones are non-recombinant. However, in practice, linkage phases are not always known. This is especially the case in animals, as it is hard to create inbred lines. And markers are often in linkage equilibrium, even across breeds. If the linkage phase is not known, we can usually infer the parental linkage phase, as the number of recombinants is expected to be smaller than the number of non-recombinants. However, there is some chance that by chance there are more recombinants. Maximum Likelihood is used to determine the most likely phase, and therefore, to determine the most likely **recombination fraction**.

Recombination fraction is the number of offspring that inherit different alleles of a trait from each parent, rather than inheriting all alleles from the same parent.



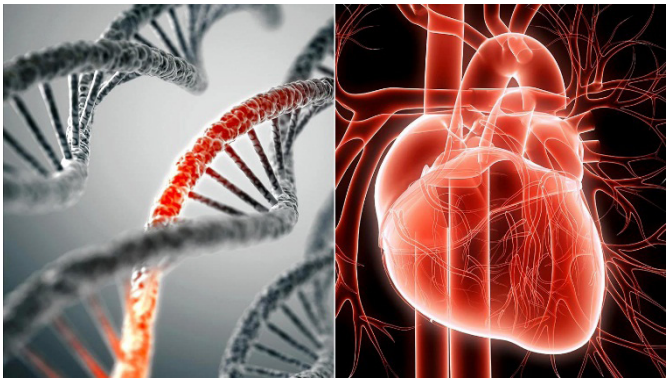
Information about the gamete that was received by an offspring depends on the genotypes on offspring, parents. If parents and offspring are all heterozygous (e.g. Aa), then we do not know which allele was paternal and which was maternal. If marker genotypes of parents are not heterozygous, we have no information about recombination events during their meiosis. For example, if the sire has genotype AB/Ab we cannot distinguish between recombinant gametes. However, if one parent is homozygous, it increases the chance of having informative meiosis on the other parent.

Maximum likelihood estimation of linkage (recombination fraction)

The likelihood is equal to the probability of observing a certain data set for given parameter values. In linkage studies, the most important parameter involved is recombination fraction. Other parameters can be population allele frequencies, but these are not needed if all parents are genotyped.

Human Disease Genes

Human families pose some of the greatest challenges to linkage analysis. Human families are small, and matings are not designed by the needs of genetic analysis. Mapping a mutation that causes a disease usually requires assembling enough families that transmit the mutation in hopes that some of them will be heterozygous, or informative, at some RFLP, RAPD, or other markers that are near enough to the disease gene to show linkage. Instead of determining linkage by counting crossover numbers as Sturtevant did, human genetics uses an alternative means to estimate whether linkage is present between marker and disease gene. This approach is called LOD score analysis, after Log of the Odds for or against linkage. Each child from informative parents is scored as recombinant (R) or parental (P). The total number of R and P results for each family is used to calculate "scores" for the odds that the results are due to linkage at a table of recombination frequencies from 1 cM, 10 cM, 20 cM, etc., relative to the chance that the results came from independent assortment.



The logs of the odds scores for each family are added to the log scores of other families to increase the number of independent observations. A LOD score value of 3, representing no more than a 5 percent chance of mistakenly declaring linkage, is the minimum acceptable score for assumption of true linkage between marker and disease gene. The recombination value that gives the highest LOD score over all the families is the presumptive linkage distance of the disease gene mutation from the adjacent markers. The first human disease gene mapped this way was Huntington's disease, which had a LOD score of over 6 for a recombination distance from its marker of between 5 and 10 cM. Once a marker has been found, it can be used to predict whether any particular family member has inherited the marker and therefore is likely to have inherited the disease gene.

Disease gene mapping hypothesizes the presence of a disease locus, for which genetic distances to one after another genetic variant are estimated. A significantly small such distance (recombination fraction significantly smaller than 50%) establishes genetic linkage between the disease locus and a given variant.



CORE CONCEPT: CHROMOSOME MAPPING

Chromosome mapping is physically locating and "marking" a gene or a specific sequence of DNA on the chromosome. The chromosome map is as useful as or more so than any other map. Having a road map shows how far towns are from each other and in which sequence they lie. A chromosome map is the same; genes and sequences are located along the chromosome and one knows the distance between markers and the sequence in which they lie.

Chromosome mapping is the process of determining the locations of specific genes on a chromosome. Chromosomes are collections of deoxyribonucleic acid, or DNA that contain genes, although it is not well understood where exactly these genes lie within chromosomes. Creating a chromosome map is useful in understanding how genes give rise to physical traits in living things. Chromosome mapping can also be important in the treatment of disease by allowing doctors to identify disease-causing genes. There are different types of chromosome maps, each with varying levels of precision and completeness.



Every living organism contains a set of operating instructions inside long molecules called DNA. These molecules can be found inside every cell of the organism and are grouped together in chromosomes. Specific units of these molecules, called genes, have been found to be responsible for physical traits in the organism, such as eye color or blood type. Understanding genes, therefore, allows scientists to understand physical traits in living things. Some genes on a chromosome tend to be inherited together with adjacent genes. This tendency, called genetic linkage, was discovered in the early years of the twentieth century. During **human reproduction**, a child's chromosomes are formed from a complicated combination of the parents' chromosomes. It is known that genes from each parents' chromosomes can be transmitted in clusters. Discovering the locations of genes on chromosomes, therefore, could facilitate the understanding of how genes are inherited and expressed in offspring.

It is important to distinguish between two different procedures related to chromosome mapping. Genetic mapping, or sequencing, is obtaining an ordered list of genes on a strand of DNA. This is similar to making a numbered list of every exit along a highway. Physical mapping, on the other hand, is determining the physical location of genes on a particular chromosome. This is more like identifying the landmarks in the area near each exit on a highway.

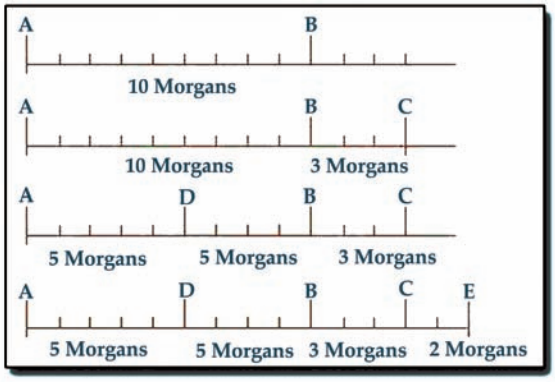
Human reproduction is any form of sexual reproduction resulting in human fertilization, typically involving sexual intercourse between a man and a woman.



There are several methods used in chromosome mapping. Historically, relative gene positions were inferred from breeding experiments with organisms. This was done extensively with the fruit fly species Drosophila melanogaster. If certain traits were frequently inherited together, the responsible genes were assumed to be located near each other. This method is useful in chromosome mapping for determining the relative location, rather than the absolute location, of genes. Physical maps can be made using modern microbiology techniques. DNA strands are sometimes separated and broken into fragments for further study. These maps can vary widely in resolution. Resolution in physical genetic mapping is the degree of separation that can be discerned in closely related genetic elements.

Chromosome Map Unit

A frequency of one per cent crossing over between two genes is known as one unit of map distance between these genes and it is termed a Morgan after the name of T. H. Morgan.



How to prepare chromosome maps

Certain examples are quoted here to explain the technique of preparing a chromosome map. Suppose the five genes A, B, C, D and E are located in a single chromosome. If the crossing over percentage between A and B is 10 Morgan's, these genes are plotted 10 arbitrary units apart in a linear scale. If the crossing over percentage between B and C is three, the crossing over percentage between A and C will determine the position of C, since it is three units from B towards A or away from A. It is assumed that A-C cross over rate is 13 per cent, and therefore, C will be plotted 3 Morgan's on the right side of B. As per above mentioned examples, the location of the genes D and E can be plotted. Let A- D cross over rate be 5 per cent and B-D cross over rate also 5 per cent. D will then be plotted between A and B, i.e., 5 Morgan's from each. If cross over rate between C and E happens to be two per cent and between B and E 5 per cent. E will be plotted two Morgan's on the right side of C.

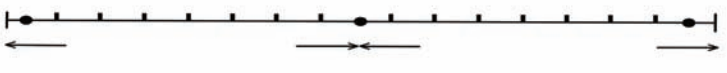


Figure 1: Morgan and Sturtevant mapping genes.

Morgan and Sturtevant got recombination frequencies of (i) 18 between the genes B for black body and for vestigial wings, (ii) 9 between the genes b for black body and cn cinnabar eye and (iii) 9.5 between the genes cn for cinnabar eye and vg for vestigial wings. If the gene b is plotted at one end of the chromosome, the gene vg will be 18 Morgan's away from it (Figure 1a). The gene cn being 9 Morgan's from b and 9.5 Morgan's from vg will be plotted approximately in the middle between the genes b and vg. Thus, the sequence of genes on the chromosome is b, cn and vg. The recombination frequency of 9.5 was obtained between cn and vg. This revealed the sequence to be b – cn – vg with a total distance of 9 + 9.5 = 18.5 map units which is close to the observed 18 map units for b — vg. If it were cn – b – vg, c – vg, distance should be 9 + 18 = 27 map units apart. Using this method the relative distances and position of hundreds of mutant genes of Drosophila have been mapped.

All these mutants cluster together to form four groups of linked genes called linkage groups. Drosophila possesses four types of chromosomes, i.e., I, II, III and IV. Each linkage group remains associated with a particular chromosome. No gene has been identified on the Y chromosome. Chromosome maps were first prepared for Drosophila, maize and Neurospora. Now the linkage maps have also been made for certain other organisms. By a somewhat different technique the chromosome maps of certain viruses and for the circular DNA of some bacteria have also been prepared. A map of chromosome II of Drosophila is given in Figure 2.

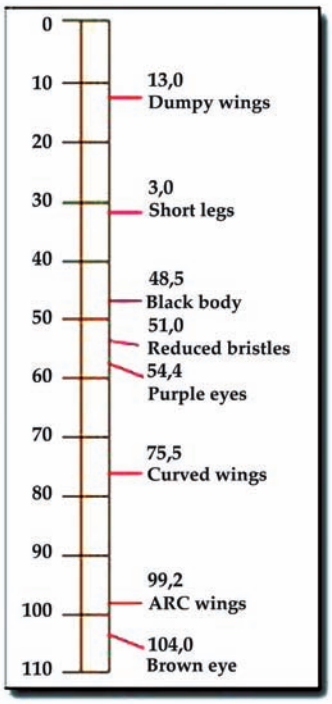
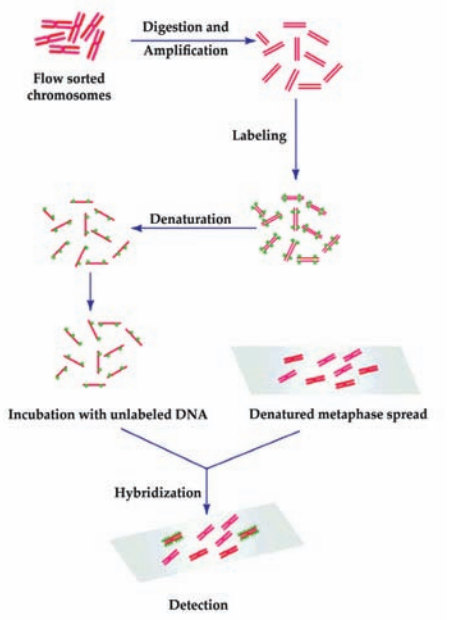


Figure 2: A map of Drosophila Chromosome II.

General Methodology of Chromosome Painting

Using these competitive hybridization techniques, whole chromosomes can be visualized in metaphase spreads and interphase nuclei. Homologies between the chromosomes of different species can be detected by chromosome painting.

- The suspensions of chromosomes from dividing cells are sorted by flow cytometry method. The flow sorted chromosomes contain sequences which are shared with other chromosomes.
- Flow sorted chromosomes are digested and amplified with the help of Polymerase Chain Reaction (PCR)
- Then the amplified DNA from one chromosome is labelled with a fluorescent dye using a technique called fluorescence in situ hybridization. This labelled DNA will paint and help in the identification of homologous regions of DNA of other species. We can identify the homologous regions from apes to mice. Homologous regions show up in the same color on the chromosome charts.
- The labeled fragments are denatured again and then incubated with unlabeled DNA.
- The fluorescent-labelled DNAs will hybridize with similar chromosomes from which they were derived. Because DNA fragments with the same base sequences have the characteristic of hybridizing with each other.



Technique of Chromosome Painting

Chromosome painting method is actually a modification of Fluorescence In-Situ Hybridization (FISH) and is also called Zoo-FISH. Human chromosome-6 consists of hundreds of genes in major histocompatibility complex (MHC). With the help of chromosome painting researchers proved that large segments of human chromosome-6 have their homologous genes in,

- Chromosome 5 of chimpanzee
- Chromosome B2 of domestic cat
- Chromosome 7 of pig
- Chromosome 23 of cow etc.

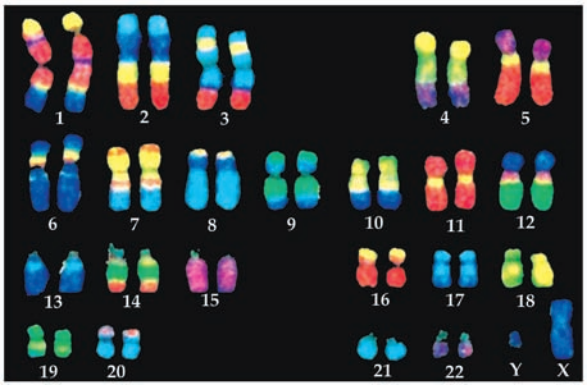
With this technique, we can also visualize chromosome aberrations if any by using fluorescently labeled DNA probes. These labeled probes hybridize to chromosomal DNA. Even multiple fluorochromes can be attached to the probes. Upon hybridization, this produces a multicolor painted effect with a unique color at each site of hybridization. We can also detect cross-species homology by labeling probes from one species for **hybridization** with chromosomes from another species.

Hybridization is a phenomenon in which single-stranded deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) molecules anneal to complementary DNA or RNA.

Probes used in Chromosome Painting

Day by day with the progress in technology, chromosome painting probes are also available for huge number of species.

Chromosome painting probes for mouse and rat are most common and are very helpful as animal models to analyses human diseases.



Chromosome painting helps in visualization of individual chromosomes in metaphase or interphase stages. With the help of this technique we can identify both numerical and structural chromosomal aberrations with high sensitivity and specificity. As already mentioned, the synchronized hybridization of multiple chromosome painting probes, each tagged with a specific light-emitting fluorochrome is called color or spectral karyotyping. It helps in differential color display of human and mouse chromosomes. Chromosome painting is also called M-FISH or multicolor FISH because modified fluorescent in situ hybridization (FISH) has been used to detect the location of specific genomic targets using probes that are labelled with specific fluorochromes. This technique allows detection of simple and complex chromosomal rearrangements.

Benefits of Genetic Mapping

The techniques developed for genetic mapping have had great impact on the life sciences, and particularly in medicine. But genetic mapping technologies also have useful applications in other fields. Commercialization of the fruits of genomics research promises immense opportunities for industry. A round-up of genetic mapping applications would include (but not be limited to) the areas below.

Gene therapy is the therapeutic delivery of nucleic acid into a patient's cells as a drug to treat disease.

Medicine

Scientists have become more proficient in genetic sequencing - the detailed genetic maps that help locate the risk genes for a host of genetic diseases. The ability to investigate the root cause of diseases may one day allow medical researchers to

develop strategies to avoid the environmental conditions that serve as triggers to disease, formulate customized drugs, and techniques for **gene therapy**.

Agricultural Applications

Knowledge of the genetic maps of plants and animals leads to the development of agricultural crops and animal breeds that are more nutritious, productive and can better resist diseases, insects and drought. Researchers can breed special plants that help clean up wastes that are difficult to break down.

Energy and the Environment

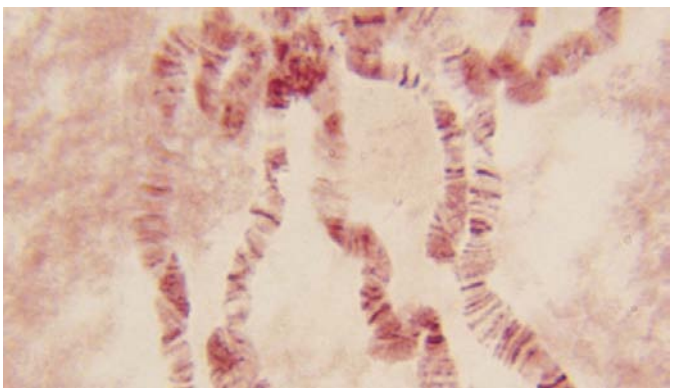
Genetic maps of microbes enable researchers to harness the power of bacteria for producing energy from bio-fuels, reducing toxic waste, and developing environment-friendly products and industrial processes.

Forensics

You are already familiar with the use of genetic mapping in crime investigations, paternity tests, and identification. The technique can also be used in organ transplants to achieve better matches between recipients and donors, thus minimizing the risks of complications and maximizing the use of donated healthy organs, a scarce resource. For more delectable applications, genetic mapping can authenticate the origins of consumer goods like caviar, fruits, and wine or the pedigree of livestock and animal breeds.

Difference between Linkage Mapping & Chromosome Mapping

Mapping of genes allows us to understand the arrangement of genes, to identify genes responsible for the expression of a particular phenotype and to identify the mutated gene responsible for a specific variant. They are useful in the field of molecular biology for performing map-based cloning and marker-aided selection. Genetic mapping gives us an idea on relative location of a gene on a chromosome which is calculated based on the recombination frequencies. The distance on genetic maps is denoted in centimorgans (cM). Physical map indicates the physical distance of the loci based on the number of base pairs present between the loci.



In 2003, scientists working on the Human Genome Project declared the sequencing of human DNA complete. However, the sequence alone does not give us the full story. Chromosome and linkage maps are useful because we do not know what all of these sequenced genes actually do. Linkage, or how often traits are passed to offspring together, is used to map, identify and organize genes according to the phenotype -- or how the genes are expressed. Chromosome mapping identifies the physical location of genes on a chromosome.

Mapping Chromosomes

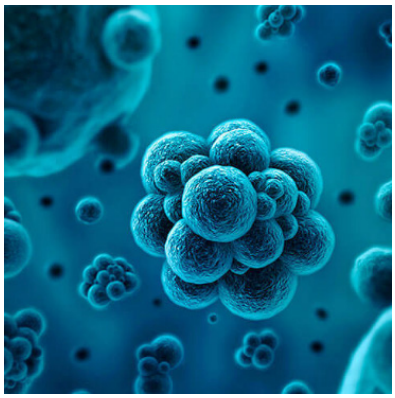
Chromosome mapping refers to the process of locating and identifying genes on a specific chromosome. A diagram used to display chromosome maps is called an ideogram. Ideograms help scientists identify chromosomes based on their size, shape and banding patterns, as well as pinpoint regions of a given chromosome. By comparing the genetic map, or ideogram, of an individual's chromosome with a standard ideogram, a geneticist can identify abnormalities in an individual's genetic code.

Linked Genes

Genes that are close neighbors on the chromosome are more likely to be linked, meaning they are often inherited together. This is referred to as genetic linkage. For example, flower color and pollen shape of the sweet pea plant are inherited traits that are linked together in offspring. In general, the closer two genes are located, the greater chance that they will be inherited together. The further apart two genes are on the chromosome, the greater the chance that they will be separated when DNA recombines during the cell division process of meiosis, resulting in a greater recombination frequency. Instead of a specific physical distance along a chromosome, a linkage map shows position of known genes or genetic markers according to their recombination frequency.

Applications of chromosome Painting

Generally, a Karyotype is the featured chromosome complement of each and every eukaryotic species.



Following are some of the applications of chromosome painting.

- With the help of chromosome painting, we can get information about the divergence of lineages which has occurred over several million tears. Also the rearrangements which took place between species can be tracked.
- Chromosome painting helps in visualization of individual chromosomes in metaphase or interphase stages.
- Chromosome painting can identify both numerical and structural chromosomal aberrations with high sensitivity and specificity
- Chromosome painting is highly sensitive and specific.
- The application of chromosome painting in the field of comparative cytogenetic helps in understanding and determining the chromosomal changes that have occurred during the evolution of species.
- Chromosome painting identifies homologous chromosome segments in different species.
- This technique can map probes of different complexities and chromosome rearrangements in one shot-single experiment.
- In recent years, the complete **karyotypes** of various mammals including primates, carnivores and artiodactyls have been analyzed by chromosome painting.

CONCEPT OF LINKAGE DISEQUILIBRIUM

Linkage disequilibrium is a concept in population genetics that describes the non-random association of alleles at different loci within a given population. It occurs when certain combinations of alleles at different genetic loci are found together more frequently than expected by chance. Linkage disequilibrium arises when genetic loci are physically close to each other on the same chromosome or when there are selective or historical factors that maintain the association between certain alleles.

Linkage disequilibrium is a powerful tool in population genetics as it provides insights into the genetic history and evolutionary relationships of populations. The extent of linkage disequilibrium between loci can be used to study the genetic structure of populations, investigate the effects of natural selection, and understand the patterns of genetic variation within and between populations.

In medical genetics, linkage disequilibrium plays a significant role in identifying genetic risk factors for diseases. When a genetic variant is associated with a particular disease, nearby variants may also be in linkage disequilibrium with the disease-causing variant. This information can be leveraged to identify candidate genes or regions associated with the disease and guide the development of personalized medicine and targeted therapies.

Additionally, linkage disequilibrium is a key consideration in genome-wide association studies (GWAS), where researchers examine the relationship between genetic variations and specific traits or diseases. By identifying regions of the genome with high linkage disequilibrium, GWAS can uncover genetic loci that are associated with complex traits and diseases.

Linkage disequilibrium is a valuable concept in population genetics, providing insights into the genetic associations between different loci and their implications for evolution, disease, and genetic research. Its study continues to play a crucial role in our understanding of the genetic basis of complex traits and the complex interactions between genes in populations.

SUMMARY

- Genetic linkage is the tendency of DNA sequences that are close together on a chromosome to be inherited together during the meiosis phase of sexual reproduction.
- A genetic linkage map is a tool used in genetic research to help scientists map the order of particular traits as they appear in a strand of deoxyribonucleic acid (DNA).
- Linkage analysis is a well-established statistical method for mapping heritable trait genes to their chromosome locations. Genome-wide markers are tested in pedigrees segregating a trait.
- Recombination is enhanced by the phenomenon of crossing over, in which gene sequences called linkage groups are disrupted, resulting in an exchange of segments between paired chromosomes that are undergoing separation.
- Genetic marker, any alteration in a sequence of nucleic acids or other genetic trait that can be readily detected and used to identify individuals, populations, or species or to identify genes involved in inherited disease.
- Mapping a mutation that causes a disease usually requires assembling enough families that transmit the mutation in hopes that some of them will be heterozygous, or informative, at some RFLP, RAPD, or other markers that are near enough to the disease gene to show linkage.
- Chromosome mapping is the process of determining the locations of specific genes on a chromosome. Chromosomes are collections of deoxyribonucleic acid, or DNA that contain genes, although it is not well understood where exactly these genes lie within chromosomes.
- Chromosome painting method is actually a modification of Fluorescence In-Situ Hybridization (FISH) and is also called Zoo-FISH.
- Mapping of genes allows us to understand the arrangement of genes, to identify genes responsible for the expression of a particular phenotype and to identify the mutated gene responsible for a specific variant.

MULTIPLE CHOICE QUESTIONS

1. Who coined the term linkage?

a. Correns
b. Mendel
c. Morgan
d. de Vries

2. A Drosophila has four pairs of chromosomes. It contains the following connection groups:

a. one more than the pair of chromosomes
b. one less than the pair of chromosomes
c. four
d. eight

3. If a proportion of two genes crosses over, the distance between them is 10.

a. 5 morganoid
b. 10 centimorgans
c. 20 centimorgans
d. 40 map units

4. Linkage is defined as the distance between two genes

a. Decreases, decreases, decreases, decreases, decreases, decreases, decreases,
b. Has no effect, but decreases
c. Decreases and Gains
d. Increases, Increases, Increases, Increases, Increases, Increases, Increases, Increases

5. Crossing over is more frequent in

a. males
b. females
c. both
d. None of these

Answer

1. (c) 2. (c) 3. (c) 4. (c) 5. (b)

REVIEW QUESTIONS

1. What do you understand by linkage map?
2. Discuss about linkage disequilibrium.
3. Define the linkage analysis in drosophila.
4. What is recombination mapping?
5. Write a short note on genetic marker.
6. What do you mean by chromosome map unit? And how to prepare chromosome maps.
7. Differentiate between linkage mapping and chromosome mapping.
8. What are the various applications of chromosome painting?

REFERENCES

1. Alberts, Bruce, et al. Molecular Biology of the Cell, 4th ed. New York: Garland Publishing, 2000.
2. Cantor, Rita M. (2013), "Analysis of Genetic Linkage", in Rimoin, David; Pyeritz, Reed; Korf, Bruce (eds.), Emery and Rimoin's Principles and Practice of Medical Genetics (6th ed.), Academic Press, pp. 1–9
3. Cooper, DN; Krawczak, M; Polychronakos, C; Tyler-Smith, C; Kehrer-Sawatzki, H (October 2013). "Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease". Human genetics. 132 (10): 1077–130.
4. Ferreira, Manuel A. R. (2004-10-01). "Linkage Analysis: Principles and Methods for the Analysis of Human Quantitative Traits". Twin Research and Human Genetics. 7 (5): 513–530.
5. Fisher, RA; Balmukand, B. "The estimation of linkage from the offspring of selfed heterozygotes". Journal of Genetics. 20 (1): 79–92
6. Griffiths, AJF; Miller, JH; Suzuki, DT (2000). "Accurate calculation of large map distances, Figure 6-4". An Introduction to Genetic Analysis (7th ed.). New York: W. H. Freeman.
7. Lawson, N. D., & Weinstein, B. M. In vivo imaging of embryonic vascular development using transgenic zebrafish. Developmental Biology 248, 307–318 (2002).
8. Mark J. Daly; Hirschhorn, Joel N. (2005-02-01). "Genome-wide association studies for common diseases and complex traits". Nature Reviews Genetics. 6 (2): 95–108.
9. McKee, Bruce D. (2004-03-15). "Homologous pairing and chromosome dynamics in meiosis and mitosis". Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression. 1677 (1–3): 165–180.
10. Patton, E. E., & Zon, L. I. The art and design of genetic screens: Zebrafish. Nature Reviews Genetics 2, 956–966 (2001).
11. Wilson, G.N. Clinical Genetics: A Short Course. New York: Wiley- Liss, Inc., 2000.



CHAPTER 7

GENETIC LABORATORY TECHNIQUES



Genetic laboratory techniques play a crucial role in advancing our understanding of genetics and molecular biology. These techniques encompass a diverse range of methodologies used to analyze and manipulate DNA, RNA, and proteins. One fundamental technique is polymerase chain reaction (PCR), which allows for the amplification of specific DNA sequences, enabling researchers to study genes and genetic variations. Gel electrophoresis is another essential method, facilitating the separation and visualization of DNA fragments based on their size. DNA sequencing techniques, such as Sanger sequencing and Next-Generation Sequencing (NGS), enable the determination of the exact order of nucleotides in a DNA molecule. Genetic engineering techniques, like CRISPR-Cas9, have revolutionized the field by enabling precise genome editing. Additionally, gene expression analysis techniques, like quantitative PCR (qPCR) and microarrays, provide valuable insights into gene activity levels.

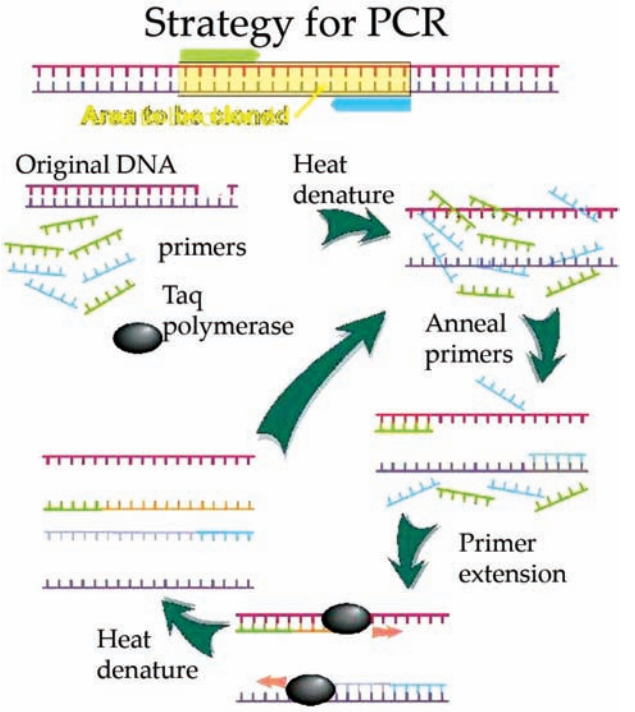
After studying this chapter, you will understand the core concepts of:

- Culturing Drosophila Stock
- Concept of Recombinant DNA Techniques Cloning
- Blood Typing
- DNA Fingerprinting
- Cryogenic Preservation of Animal Germplasm

INTRODUCTION

In order to analyze DNA, it is necessary to isolate it from its source. Often times, this source will be whole cells from organisms, which in addition to DNA, will also contain lipids, proteins, signaling molecules, and other types of compounds. Pure DNA is obtained through extraction of it from the rest of the cell, separating it from other things in the cell and allowing for analysis of the pure DNA. Traditionally, this has been done using a process known as Phenol extraction, which lyses the cell and precipitates DNA and nothing else in the cell, allowing us to obtain DNA through a series of mixing and fractioning. In recent years, however, companies such as Qiagen have created kits that allow for simple extraction of DNA without the long tedious process normally required by Phenol extraction. This allows geneticists to easily and quickly obtain pure DNA for subsequent analysis.

These subsequent steps for analysis of the pure DNA, however, requires the DNA to be in large amounts. While organisms typically have many cells from which to extract DNA, the extraction process is not fully accurate, and some cells have different sets of DNA due to processes such as mitotic recombination. Furthermore, our current techniques of analysis allow us to analyze specified lengths of DNA (<1000 base pairs) at any given location. Thus, it would be convenient to have a large, pure pool of a specific location in the genome of interest. This is obtained through a process known as Polymerase Chain Reaction (PCR). PCR involves the amplification of a portion of the DNA (<1000 base pairs) on an exponential scale. It involves repeated processes of melting the double stranded helix, annealing of primers, synthesizing the complement strand using polymerase, and yielding two strands of DNA identical to the first. Thus, with each cycle of this, the number of strands of DNA available doubles.





CORE CONCEPT: CULTURING DROSOPHILA STOCK

Culturing Drosophila stock involves maintaining a laboratory population of fruit flies (Drosophila melanogaster) for research purposes. The process typically starts by establishing a breeding population with carefully selected flies. These flies are kept in vials or bottles containing a nutrient-rich medium, which serves as their food source. The culture environment is controlled for temperature, humidity, and light conditions to optimize breeding success. To prevent genetic drift, regular passages or transfers of flies to fresh culture media are performed. Researchers use this established Drosophila stock to study genetics, developmental biology, behavior, and other fundamental biological processes due to the fruit fly's convenient characteristics as a model organism.

Drosophila was one of the first organisms to be studied genetically: its small size, short life cycle (10 ~14 days at 25°C), high reproductive rate (an adult female can lay 400-500 eggs in 10 days), and ease of culture and genetic manipulation have made it perhaps the best understood animal genetic system. Many different species, and a large number and wide variety of naturally-occurring and artificially induced genetic variants are available.

Virgin Females

All female flies used in controlled genetic crosses must be "virgins". Female flies are capable of mating as early as 8 hours after emerging from the pupae stage and are polyandrous, that is, capable of mating with several males. Once mated, females can retain viable sperm for several days and this will confuse the results of a subsequent controlled mating. To prevent this, all adult flies are removed from the culture bottle about 7 hours prior to lab time, so that all newly hatched flies will remain virgin.

Basic Genetics

The karyotype of **Drosophila** comprises four pairs of chromosomes, of which three pairs are autosomes and one pair are sex chromosomes. Female Drosophila are XX, and males XY.

Drosophila is a genus of small flies, belonging to the family Drosophilidae, whose members are often called "fruit flies" or (less frequently) pomace flies, vinegar flies, or wine flies, a reference to the characteristic of many species to linger around overripe or rotting fruit.

A gene is a heritable factor that controls the expression of some trait, which may be morphological, behavioral, molecular, etc. Each such gene occupies a specific physical locus (pl. loci) on a particular chromosome. Variant forms of these loci are termed alleles. Gene, locus, and allele are often used more or less interchangeably, and this can lead to confusion. Gene is the popular and most general term, and is most appropriate when the inherited basis of a trait is emphasized, e.g., a "gene" for eye color. Locus is most appropriate when the physical nature or position of a gene, especially with respect to other genes, is emphasized, as for example in gene mapping and linkage studies. Allele is most appropriate when the particular form(s) of a gene found in any particular individual or chromosome is (are) emphasized: e.g., there are "brown" and "blue" alleles of the eye color gene It is therefore inaccurate to say, for example, "He has the gene for sickle-cell anemia," and more accurate to say "He has two HbS alleles at the beta-globin locus on Chromosome six." We all have the "gene" for every genetic condition, some of us have the particular allele(s) that result in the condition being expressed. In the technical literature, "locus" and "allele" are probably more common than "gene".

If the gene occurs on a sex chromosome, females may be either homozygous or heterozygous, but a male fly with only one allele at a locus will be a hemizygote and should be described as hemizygous.

Drosophila, like most species we will deal with in this course, are diploid, with two sets of chromosomes and therefore two alleles at each autosomal locus. If both alleles are identical, the individual is a homozygote and is described as homozygous. If the alleles differ from each other, the individual is a heterozygote and is described as heterozygous.

Drosophila of typical appearance are said to show the "wild-type" forms (phenotypes) of genetically-controlled traits for body color, eye color, wing shape, etc. Naturally-occurring or artificially-induced genetic variants (mutations) of the alleles that control these traits produce flies with different morphologies, according to the dominant or recessive nature of the alleles involved in the genotype. Such mutant alleles are designated by symbols that are typically abbreviations of the mutant name. For example, the typical body color phenotype is grey. One mutant produces an ebony (shiny black) body color. Because this allele is recessive, it is symbolized by a lower-case letter, e. The wild-type allele is symbolized by a "+" sign, used either alone (if there is no ambiguity) or in combination with the mutant allele symbol, in this case e^+. Thus, the genotype of a wild-type homozygote would be designated e^+ e^+ (or ++), a mutant homozygote ee, and a heterozygote e^+ e or e^+.

It is important to remember that not all mutants are recessive. A mutation that is dominant to the wild-type is symbolized by a capital letter. For example, the typical eye shape is round. One mutant produces a narrow "bar eye": the allele is dominant, symbolized by a capital letter B, and the wild-type (round) eye is B^+.

Genetic Crosses

An "X" is used to indicate that two individuals have been mated together. The parents are designated as P (for parental) and the offspring as F (for filial). When several generations are involved, subscripts are added to designate the generations.

P_1 give rise to F_1 (first filial) progeny.

If the F_1 are crossed together they become P_2 and their progeny F_2.

A cross between members of the F_1 and members of the P_1 is a backcross.

A cross between members of the F_1 and the true breeding recessive P_1 is a test cross.

Monohybrid Cross

The simplest form of a cross is a monohybrid cross, which analyses a single trait and its associated variations. The diagram below shows the progression of a pair of alternative alleles for a single gene through two generations.

P_1 **ee** (ebony body) x **e^+e^+** (wild-type body) [homozygous parents]

gametes **e** **e^+**

F_1 **e^+e** (wild-type body) [heterozygous offspring]

gametes **e^+** & **e**

♀ \ ♂	**e^+**	**e**
e^+	**e^+e^+**	**e^+e**
e	**e^+e**	**ee**

F_2 Phenotype ratio 3 wild-type: 1 ebony body

Genotype ratio 1 **e^+e^+** : 2 **e^+e** : 1 **ee**

During gamete formation, the members of a pair of alleles are duplicated and then segregated from one cell into four separate gametes so that each contains only one member of the pair (Law of Segregation).

A Punnet Square diagram can be used to calculate the various combinations. The gametes from one parent are written across the top, and those from the other go down the side. Each one of these gametes has an equal chance of combining with either of the gametes from the other parent (Random Union of Gametes).

The F_2 phenotype ratio of 3:1 indicates a case of complete dominance. That is, one allele completely masks the expression of the other (recessive) allele. In cases of incomplete dominance, on the other hand, neither allele masks the other, nor do heterozygous individuals express new phenotypes that are intermediate between

the homozygous parents. This may arise for example if the dominant homozygous phenotype results from the expression of a double-dose of gene product, and the heterozygous phenotype from a single dose. The F_2 phenotype ratio of 1:2:1 is characteristic. A non-Drosophila example of this is seen in red- and white-flowered snap dragons:

P_1 **RR** (red) x **rr** (white)
F_1 **Rr** (pink)
F_2 1 **RR** (red) : 2 **Rr** (pink) : 1 **rr** (white)

When both alleles are expressed the effect is known as codominance. Heterozygous individuals express gene products from both alleles: unlike incomplete dominance, the phenotype need not be intermediate. This sort of interaction is seen in the ABO blood group system of humans. One allele controls the production of A antigen while the other controls the B antigen (a third allele O produces no antigen). Heterozygotes carrying the allele for antigen A and the allele for antigen B have blood type AB in which both proteins are present in equal quantities. The F_2 shows a ratio of 1:2:1 , as in the case of incomplete dominance.

Dihybrid Cross

Dihybrid crosses involve manipulation and analysis of two traits controlled by pairs of alleles at different loci. For example, in the cross ebony body x vestigial wing

e is ebony body colour
e$^+$ is wild-type body colour
vg is vestigial wing shape
vg$^+$ is wild-type wing shape:

Genotype is the part of the genetic makeup of a cell, and therefore of an organism or individual, which determines a specific characteristic (phenotype) of that cell/organism/individual.

where the loci for ebony body color and vestigial wing are on separate autosomes. Therefore the **genotypes** and gametes are the same for male and female.

P_1 ebony body x vestigial wing

ee vg^+vg^+ e^+e^+ vgvg

gametes e vg^+ e^+vg

F_1 e^+e vg^+ vg (all wild-type)

gametes e^+ vg^+ , e^+ vg, evg^+ , e vgF^2 genotype combinations:

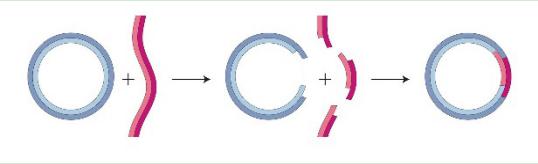
♀ / ♂	e^+vg^+	e^+vg	$e\ vg^+$	$e\ vg$
e^+vg^+	$e^+e^+vg^+vg^+$	$e^+e^+vg^+vg$	$e^+e\ vg^+vg^+$	$e^+e\ vg^+vg$
e^+vg	$e^+e^+vg^+vg$	e^+e^+vgvg	$e^+e\ vg^+vg$	$e^+e\ vgvg$
evg^+	$e^+e\ vg^+\ vg^+$	$e^+e\ vg^+\ vg$	$ee\ vg^+\ vg^+$	$ee\ vg^+\ vg$
evg	$e^+e\ vg^+\ vg$	$e^+e\ vg\ vg$	$ee\ vg^+\ vg$	$ee\ vg\ vg$

F_2 Phenotype ratio: 9 wild-type: 3 ebony: 3 vestigial: 1 ebony vestigial

In a dihybrid cross, each of the F_1 parents can produce four different gamete types, so there are 16 (= 4 x 4) possible offspring combinations. Because the two traits show complete dominance and separate independently of each other (Law of Independent Assortment), the expected genotypic and phenotypic ratios from an analysis of these 16 possibilities can be calculated.

Phenotype	Genotype
(9:3:3:1)	(1:2:1:2:4:2:1:2:1)

These ratios can be derived from the results of a monohybrid ratio. A basic principle of probability theory is that the probability of two independent events occurring together is equal to the product of the two independent probabilities.



CORE CONCEPT: RECOMBINANT DNA TECHNIQUES CLONING

Recombinant DNA techniques, specifically cloning, involve the artificial manipulation and replication of DNA molecules from different sources. The process starts by isolating a DNA fragment of interest and inserting it into a vector, typically a plasmid or a viral genome. The recombinant DNA molecule is then introduced into host cells, such as bacteria or yeast, using various methods like transformation or transfection. Once inside the host cells, the recombinant DNA replicates alongside the host's genome, resulting in the production of multiple copies of the inserted DNA fragment.

Recombinant DNA (or rDNA) is made by combining DNA from two or more sources. The process often involves combining the DNA of different organisms. The process depends on the ability to cut and re-join DNA molecules at points which are identified

by specific sequences of nucleotide bases called restriction sites. DNA fragments are cut out of their normal position in the chromosome using restriction enzymes (also called restriction endonucleases) and then inserted into other chromosomes or DNA molecules using enzymes called ligases.

Gene Cloning

Polymerase chain reaction (PCR) is a technique used in molecular biology to amplify a single copy or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence.

The process of copying fragments of DNA which can then be used for many different purposes, such as creating GM crops, or finding a cure for disease. There are two types of gene cloning: in vivo, which involves the use of restriction enzymes and ligases using vectors and cloning the fragments into host cells (as can be seen in the image above). The other type is in vitro which is using the **polymerase chain reaction (PCR)** method to create copies of fragments of DNA.

For in vivo cloning a fragment of DNA, containing a single gene or a number of genes, is inserted into a vector that can be amplified within another host cell. A vector is a section of DNA that can incorporate another DNA fragment without losing the capacity for self-replication, and a vector containing an additional DNA fragment is known as a hybrid vector. If the fragment of DNA includes one or more genes the process is referred to as gene cloning.

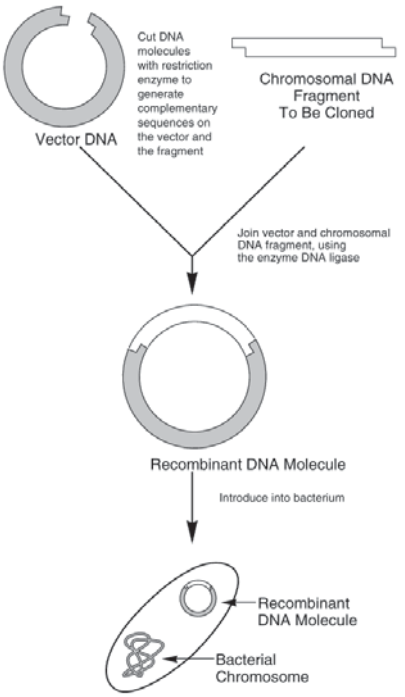


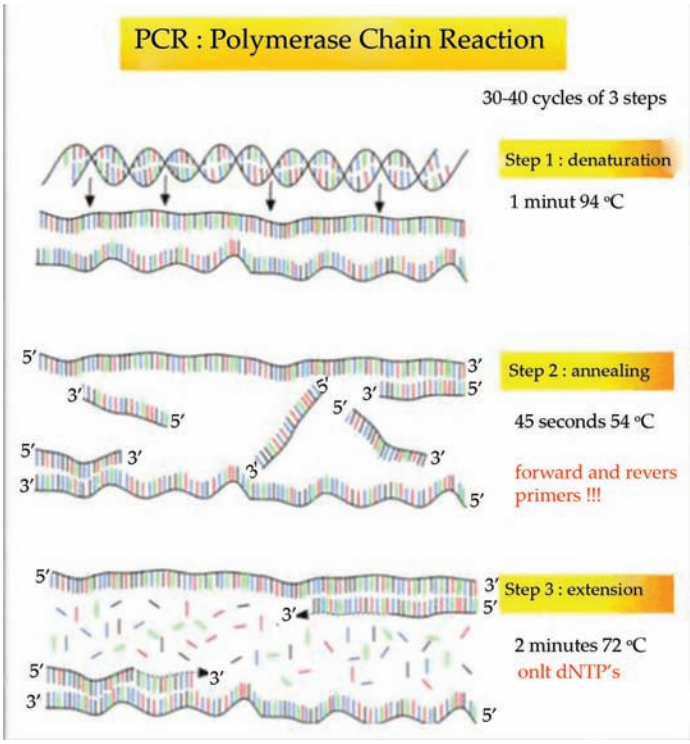
Figure 1: Cloning DNA in Plasmids.

There are 4 different type of vectors:

- Plasmid vectors
- Lamda (λ) phage vectors
- Cosmids
- Expression vectors

The host cell copies the cloned DNA using its own replication mechanisms. A variety of cell types are used as hosts, including bacteria, yeast cells and mammalian cells.

Polymerase Chain Reaction (PCR)



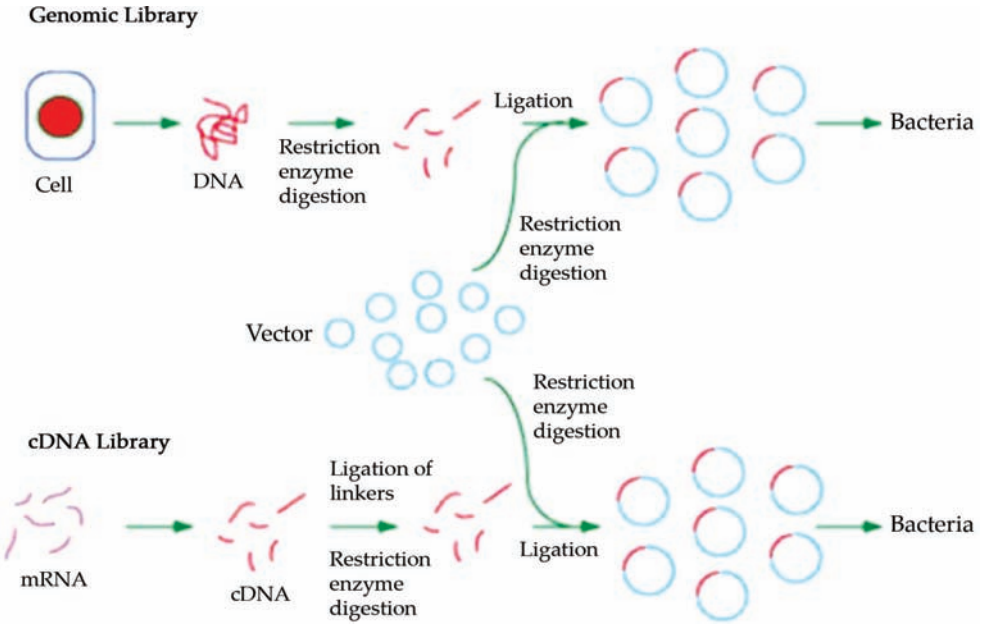
This is an in vitro method for making many copies of a specific section of DNA, without the need for vectors or host cells. The DNA to be copied – the template DNA – is mixed with forward and reverse primers complementary to the end of the template DNA, nucleotides, and a version of DNA polymerase known as Taq polymerase. (This enzyme is stable under high temperatures, and is obtained from the thermophilic bacterium Thermus aquaticus.) The process involves the repetition of three steps:

- denaturation, which separates the two nucleotide strands of the DNA molecule
- primer annealing, in which the primers bind to the single-stranded DNA
- extension, in which nucleotides are added to the primers – in the 5' to 3' direction – to form a double-stranded copy of the target DNA.

Gene Libraries

A gene library is a large collection of cloned DNA sequences from a single genome. A genomic library in theory, would contain at least one copy of every sequence in an organism's genome. These are used to investigate the structure of a given chromosome, or to clone specific genes. These types of libraries may be prepared from a subset of the entire genome (for example, a single chromosome). The first step in creating a genomic library is to break up, or 'fractionate', the genome using physical methods or restriction enzymes. The fragments are then linked to appropriate vectors and cloned in a suitable host cell population.

A cDNA library (complementary DNA) contains DNA present in a given cell population which is prepared from the mRNA (messenger RNA) using the enzyme reverse transcriptase. The resulting cDNA represents the genes expressed in the cell population as a subset of the entire genome, and can be cloned using a vector and suitable host cell. The cDNA will not include introns or regulatory sequences as these are removed from the RNA during processing, and this makes a cDNA library easier to maintain. A cDNA library can also be prepared using reverse transcriptase PCR (RT-PCR).

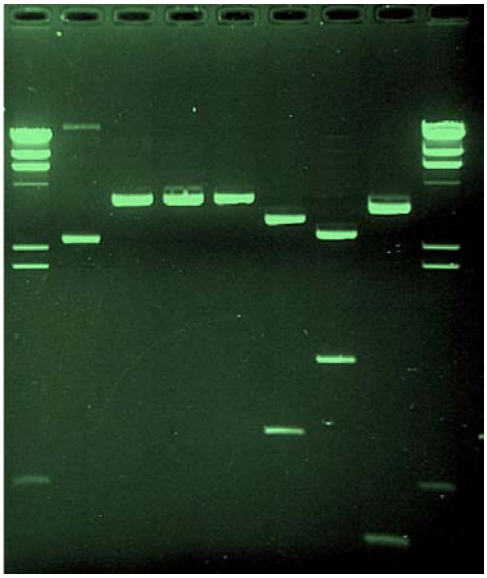


The Identification of Gene Products in a Gene Library

Restriction enzymes (to cut the DNA) and gel electrophoresis (to separate the resulting fragments) can be used to produce a physical map of DNA segments in a process known as restriction mapping.

There are also a number of techniques that can be used to identify specific genes or gene products within a gene library and these are: Southern blotting, Northern blotting and Western blotting. However, the most powerful experimental technique

for investigating genetics at the molecular level is DNA sequencing, which allows the nucleotide sequences of genes – even whole chromosomes – to be determined. Automated sequencing technologies are now allowing us to sequence the entire genomes of organisms from bacteria to human beings.

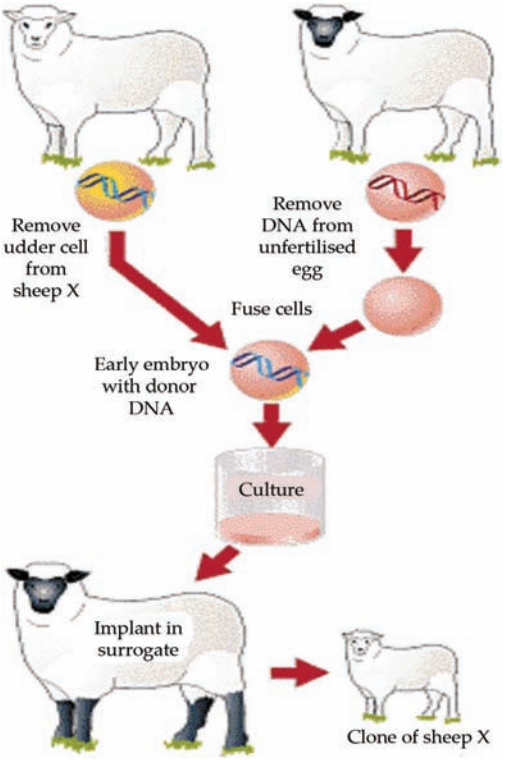


Molecular Genetics and Biotechnology

The new techniques of molecular genetics, combined with developments in associated biotechnologies, have led to advances in a number of different fields. We can now analyses the genomes of species that make an important contribution to agriculture, fuel production or drug development. We can move specific genes from one organism to another to create transgenic plants and animals, and use animal cloning techniques to produce animals that are genetically identical, such as Dolly the sheep, and more recently, cloned pets such as cats and dogs.

The process of cloning is straightforward, but the results are not always predictable. It took many hundreds of attempts to get it to work and produce one live sheep, and cloning in itself raises many questions not only about benefits and risks but also many ethical questions.

The technique of genetic fingerprinting has found many applications, including the identification of individuals and the relationships between individuals. Research into gene therapy examines the possibility of introducing cloned genes to compensate for defective, mutant genes.



Karyotyping

Karyotyping is a laboratory procedure that allows a physician to examine a patient's set of chromosomes. "Karyotype" also refers to the actual collection of chromosomes being examined. Examining these chromosomes through karyotyping allows your physician to determine whether there are any abnormalities or structural problems.

Chromosomes, which are in almost every cell of your body, contain the genetic material inherited from your parents. They are composed of deoxyribonucleic acid (DNA) and determine the way that every human being develops.

Why the Test Is Useful

If you have an unusual number of chromosomes, if they are arranged incorrectly, or if they are malformed, this can be a sign of a genetic condition, such as Down or Turner syndrome.

Karyotyping can be used to detect myriad genetic disorders. For instance, if a woman has premature ovarian failure, she may have a chromosomal defect that karyotyping can distinguish. The test is also useful for identifying the Philadelphia chromosome, the presence of which can signal chronic myelogenous leukemia.

Babies can be karyotype tested before they are born to diagnose genetic abnormalities that indicate serious birth defects, such as Klinefelter syndrome, in which a boy is born with an extra X chromosome.

How the Test Is Performed

The first step in karyotyping is to take a sample of your cells. The sample cells can come from a number of different tissues, including bone marrow, blood, amniotic fluid, or placenta. Sampling can be done using various methods, depending on which area of your body is being tested. If, for instance, amniotic fluid needs to be tested, the doctor will use amniocentesis to collect the sample.

After the sample has been taken, it is placed in a laboratory dish that allows the cells to grow. A lab technician will take cells from the sample and stain them so that the chromosomes they contain can be viewed under a microscope.

These stained cells are examined under a microscope for potential abnormalities, including:

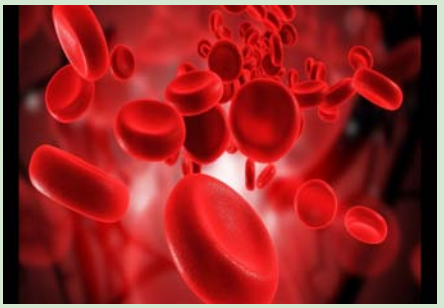
- extra chromosomes
- missing chromosomes
- missing portions of a chromosome
- extra portions of a chromosome
- portions that have broken off of one chromosome and reattached to another

The laboratory technician can see the chromosomes' shape, size, and number. This information is instrumental in determining whether or not any genetic abnormalities are present.

What the Test Results Mean

A normal test result will show 46 chromosomes, 44 of which are autosomes (which are unrelated to determining gender) and two of which are sex chromosomes. These sex chromosomes determine one's gender: females have two X chromosomes, while males have one X chromosome and one Y chromosome.

Abnormalities that appear in a test sample could be the result of any number of genetic syndromes or conditions. Sometimes an abnormality will occur in the lab sample that is not reflected in the patient's actual body. In order to confirm that an abnormality is present in the patient, the karyotype test can be repeated.



CORE CONCEPT: BLOOD TYPING

Blood typing is a common test when blood transfusions and tissue transplants are required, as well as during pregnancy. A blood type lab test identifies certain inherited substances (antigens) that may be present on the surface of red blood cells and classifies them into four common groups: A, B, AB, or O, and is known as the ABO system.

Blood types (or groups) are determined by specific antigens found on the surface of erythrocytes. In humans, there is the ABO system of blood types, whereas animals have a variety of different blood types. Knowledge of blood types in the different species is important as transfusion of incompatible blood (the donor animal has a different blood type from the recipient animal) can result in severe hemolytic transfusion reactions and even death, in some instances.

There are two types of antibodies to blood group antigens; naturally occurring antibodies and antibodies acquired after exposure to the blood group antigen. Naturally occurring antibodies occur in most species and vary in their pathological significance, i.e. some will not produce a transfusion reaction. Acquired antibodies are produced after exposure to an incompatible blood type, which is from exposure to blood or products containing erythrocytes or their antigens. The most common route of exposure is from previous blood transfusions, however there are less obvious sources of exposure, such as vaccinations that contain foreign red blood cell antigens. Antibodies that are pathogenic (i.e. induce a hemolytic reaction) can cause agglutination and/or **hemolysis** of red cells.

Hemolysis is the rupturing (lysis) of red blood cells (erythrocytes) and the release of their contents (cytoplasm) into surrounding fluid (e.g. blood plasma).

Blood typing (for the most common blood groups) is offered by a few specialized veterinary diagnostic laboratories. Blood type compatibility (or incompatibility) is determined in the laboratory using cross matching procedures. Since administration of typed negative blood will not prevent a transfusion reaction to less well-characterized red cell antigens, cross matching should always be performed in an individual that has been previously exposed to blood group antigens.

Canine Blood Types

Any animal that is routinely used as a blood donor should be blood typed for the most common antigens that produce a hemolytic reaction and (ideally) should be negative for these antigens.

There are 8 major blood groups in the dog, labeled as DEA (dog erythrocyte antigen) 1 to 8. These are illustrated in the table below. The major antigens are DEA 1.1 and DEA 1.2. Dogs can be positive for either (not both) DEA 1.1 or 1.2 or are negative for both. Naturally occurring antibodies occur in 20% of DEA 3-negative, 10% of DEA 5-negative, and 20-50% of DEA 7-negative dogs. New blood groups in dogs are being detected, including those in Dalmations (Dal) (Blias et al 2007) and Kai-1 (IgM) and 2 (IgG) (Euler et al 2016)

Acute hemolytic transfusion reactions only occur in DEA 1.1 and 1.2 negative dogs. As these dogs do not have naturally occurring antibodies, a reaction will only be seen after sensitization of the dog through exposure to DEA 1.1 or 1.2 positive blood (antibody production takes 7-10 days after exposure). The normal lifespan of compatible transfused erythrocytes in dogs is approximately 21 days. In an acute hemolytic transfusion reaction, the lifespan of incompatible transfused erythrocytes ranges from minutes to 12 hours. Although DEA 3-, 5- and 7-negative dogs do have naturally occurring antibodies to DEA 3, 5 and 7 positive red cells, these blood groups do not incite severe hemolytic reactions. Rather, transfusion of incompatible blood is hemolysed more rapidly (within 4 to 5 days) than compatible blood would be (so-called delayed hemolytic reaction). Therefore, crossmatching in dogs does not need to be done on the first transfusion. Neonatal isoerythrolysis has been reported in DEA 1 negative female dogs mated to DEA 1.1 positive male dogs.

DEA group	"old" name	Population incidence	Natural antibody	Transfusion significance
1.1	A_1	40-60%	No	Acute hemolytic reaction
1.2	A_2	10-20%	No	Acute hemolytic reaction
3	B	5-20%	Yes	Delayed hemolysis
4	C	85-100%	No	None
5	D	10-25%	Yes	Delayed hemolysis
6	F	98-99%	No	Unknown
7	Tr	10-45%	Yes	Delayed hemolysis
8	He	40%	No	Unknown
Dal		100%	No	Acute hemolysis in some Dalmations
Kai-1		94%	Not as yet	Unknown
Kai-2		1%	Not as yet	Unknown

Feline Blood Types

Only 1 blood group system, the AB system, has been identified in cats. In this system,

there are 3 blood types; A, B and AB. Similar to humans, the blood group antigens are defined by specific carbohydrates on erythrocyte membranes. A N-glycolyl-neuraminic acid determines the A antigen and a N-acetyl-neuraminic acid determines the B antigen with equal amounts of both acids being found on AB erythrocytes. Cats with blood group B lack a hydroxylase enzyme that converts N-acetyl-neuraminic acid to N-glycolyl-neuraminic acid. The blood group antigens are inherited as a simple autosomal trait with A being dominant over B. The inheritance of the AB allele is, as yet, unknown (it is not due to codominance of A and B). The percentage of cats that are A or B positive is breed-dependent. The overall incidence of A positive DSH and DLH cats varies between countries, with a higher incidence in the USA (94 to 99%) than in the United Kingdom (87%) and Australia (73%). AB cats are quite uncommon (5% in the United Kingdom and <1% in the USA and Australia).

Type B frequency	**Breeds**
None	Siamese and related breeds, Burmese, Tonkinese, Russian Blue
1-10%	Maine Coone, Norwegian Forest, DSH, DLH
11-20%	Abyssinian, Birman, Himalayan, Persian, Somali, Sphinx, Scottish Fold
20-45%	Exotic and British Shorthair cats, Cornish and Devon Rex
Type AB	DSH, Scottish Fold, Birman, British Shorthair, Somali, Bengal, Abyssinian

Cats have naturally occurring antibodies (alloantibodies) which are responsible for potentially life-threatening transfusion reactions. In B cats, the anti-A antibodies are strong agglutinins and hemolysins, especially of the IgM class. In contrast, anti-B antibodies in type A cats are weaker agglutinins and hemolysins (and are of the IgG and IgM class). Type AB cats lack naturally occurring antibodies and can safely receive blood from either type A or B cats (universal recipients).

A	Most DSH and DLH (98%) cats, all Siamese are type A.	In theory, do not need to crossmatch type A cats (or breeds with high likelihood of being cat A) if transfusing with blood from another type A cat.
B	Exotic breeds like Himalayan, Abyssinian, Somali, Birman, British shorthair, Devon Rex, and Persian.	Agglutinating antibodies are predominantly IgM. These should be crossmatched before a first transfusion because they are likely to die if transfuse with A blood.
AB	Scottish fold, Birman, British Shorthair and DSH.	AB kittens born to a B queen are in danger of neonatal isoerythrolysis, but can safely receive A or B blood.
Mik	Was discovered in a cat after a transfusion reaction of compatible A blood.	Pre-existing antibodies exist and will not be detected with typing (maybe with crossmatching if sufficient antibody titer).

The half-life of transfused erythrocytes in matched feline transfusions (i.e. type A blood to a type A cat or type B blood to a type B cat) is 29 to 39 days. Transfusion

of A blood into a B cat results in rapid destruction of the donated type A blood (mean half-life of 1.3 hours) with severe clinical signs (hypotension, defecation, vomiting, hemoglobinemia, neurologic depression) and even death. In contrast, transfusion of type B blood into A cats produces milder clinical signs and the transfused erythrocytes have a mean half-life of 2.1 days. Due to the presence of these naturally occurring antibodies, cats must be crossmatched before their first transfusion (especially in those breeds with a high incidence of type B or AB blood). In addition, neonatal isoerythrolysis can occur in kittens bearing the A or AB blood group antigen from a mating of B queens to an A or AB tom.

Equine Blood Types

There are over 30 blood groups in horses, of which only 8 are major systems. Of these 8, 7 are internationally recognized (A, C, D, K, P, Q and U), whilst the T system is primarily of research interest. Of these, the Aa and Qa are most important for hemolytic reactions, especially neonatal isoerythrolysis (NI). Other blood groups can occasionally give NI reactions, including Dc, Ua, Ab and Pa. In addition, all horses lack a unique red cell antigen to donkeys, so they will produce antibodies (and NI) when exposed to donkey blood (such as in mule pregnancies). Natural antibodies do exist, particularly to Ca antigens, which cause weak agglutination and hemolytic cross-match reactions, however the antibodies to Ca do not appear to produce a significant hemolytic reaction in vivo. The incidence of Aa and Qa is breed-dependent. The table below gives the percentage of animals in the listed breed that are negative for the factor.

System	Thoroughbred	Arabian	Standardbred	Quarterhorse	Morgan
Aa-	15%	18%	44%	51%	43%
Qa-	39%	79%	100%	83%	99%

Ruminant Blood Types

- **Cattle**: There are 11 major blood group systems in cattle, A, B, C, F, J, L, M, R, S, T and Z. The B group has over 60 different antigens, making it difficult to closely match donor and recipient. The J antigen is a lipid that is found in body fluids and is adsorbed onto erythrocytes (therefore, it is not a "true" antigen). Newborn calves lack this antigen, acquiring it in the first 6 months of life. Some animals have only a small amount of J antigen on erythrocytes and none in serum; these so-called "J-negative" animals can develop antibodies against the J-antigen and develop transfusion reactions if transfused with J-positive blood. Neonatal isoerythrolysis is not a naturally occurring phenomenon in cattle. Bouts of NI have occurred secondary to blood-derived vaccines (e.g. against anaplasmosis, babesiosis). The most common antigens that cattle were sensitized to were the A and F systems.
- **Sheep**: Seven blood group systems have been identified in sheep (A, B, C, D, M, R and X). Similar to cattle, the B system is highly polymorphic.

The R system is similar to the J system in cattle, in that the antigen is soluble. The M-L system is involved in active red cell potassium transport and polymorphisms in this system result in breeds of sheep with varying erythrocyte potassium content. Neonatal isoerythrolysis has been reported in lambs administered bovine colostrum. This is due to the presence of antibodies to sheep erythrocytes in bovine colostrum (called "heterophile" antibodies), which is a common occurrence. They are antibodies produced to common cross-reactive antigens present on the surface of bacteria and protozoa that are identical to epitopes on blood group antigens.

- **Goats**: Blood group antigens in goats are similar to those in sheep and the same reagents are used to type both species. Five major systems have been identified in goats; A, B, C, M and J (the latter is also a soluble antigen like in cattle).



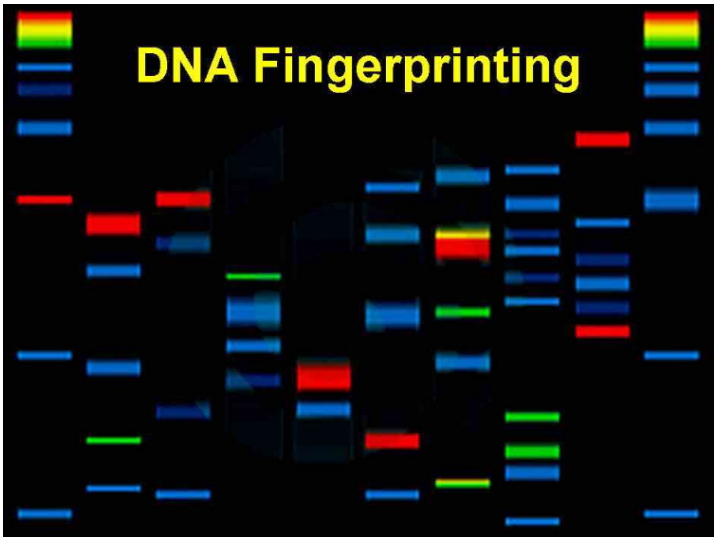
CORE CONCEPT: DNA FINGERPRINTING

DNA fingerprinting, also known as DNA profiling or genetic fingerprinting, is a forensic technique used to identify individuals based on their unique DNA patterns. Short tandem repeat (STR) regions in the genome are analyzed to create a distinctive genetic profile, making it highly accurate for personal identification, criminal investigations, and paternity testing.

DNA fingerprinting, also called DNA typing, DNA profiling, genetic fingerprinting, genotyping, or identity testing, in genetics, method of isolating and identifying variable elements within the base-pair sequence of DNA (deoxyribonucleic acid). The technique was developed in 1984 by British geneticist Alec Jeffreys, after he noticed that certain sequences of highly variable DNA (known as minisatellites), which do not contribute to the functions of genes, are repeated within genes. Jeffreys recognized that each individual has a unique pattern of minisatellites.

The procedure for creating a DNA fingerprint consists of first obtaining a sample of cells, such as skin, hair, or blood cells, which contain DNA. The DNA is extracted from the cells and purified. In Jeffreys's original approach, which was based on restriction fragment length polymorphism (RFLP) technology, the DNA was then cut at specific points along the strand with proteins known as restriction enzymes. The enzymes produced fragments of varying lengths that were sorted by placing them on a gel and then subjecting the gel to an electric current (electrophoresis): the shorter

the fragment, the more quickly it moved toward the positive pole (anode). The sorted double-stranded DNA fragments were then subjected to a blotting technique in which they were split into single strands and transferred to a nylon sheet. The fragments underwent autoradiography in which they were exposed to DNA probes—pieces of synthetic DNA that were made radioactive and that bound to the minisatellites. A piece of X-ray film was then exposed to the fragments, and a dark mark was produced at any point where a radioactive probe had become attached. The resultant pattern of marks could then be analyzed.

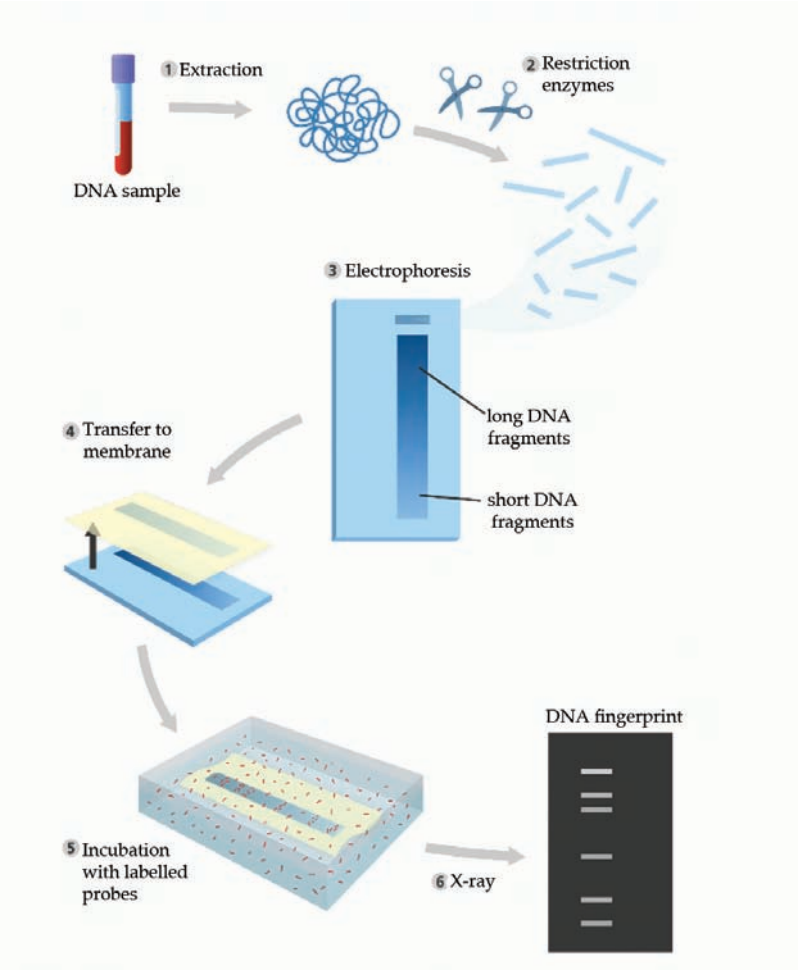


The assay developed by Jeffreys has been supplanted by approaches that are based on the use of the polymerase chain reaction (PCR) and so-called microsatellites (or short tandem repeats, STRs), which have shorter repeat units (typically 2 to 4 base pairs in length) than minisatellites (10 to more than 100 base pairs in length). PCR amplifies the desired fragment of DNA (e.g., a specific STR) many times over, creating thousands of copies of the fragment. It is an automated procedure that requires only small amounts of DNA as starting material and works even with partially degraded DNA. Once an adequate amount of DNA has been produced with PCR, the exact sequence of nucleotide pairs in a segment of DNA can be determined by using one of several biomolecular sequencing methods. Automated equipment has greatly increased the speed of DNA sequencing and has made available many new practical applications, including pinpointing segments of genes that cause genetic diseases, mapping the human genome, engineering drought-resistant plants, and producing biological drugs from genetically altered bacteria.

How was the first DNA fingerprint produced?

- The first step of DNA fingerprinting was to extract DNA from a sample of human material, usually blood.
- Molecular 'scissors', called restriction enzymes[?], were used to cut the DNA. This resulted in thousands of pieces of DNA with a variety of different lengths.

- These pieces of DNA were then separated according to size by a process called gel electrophoresis:
  - The DNA was loaded into wells at one end of a porous gel, which acted a bit like a sieve.
  - An electric current was applied which pulled the negatively-charged DNA through the gel.
  - The shorter pieces of DNA moved through the gel easiest and therefore fastest. It is more difficult for the longer pieces of DNA to move through the gel so they travelled slower.
  - As a result, by the time the electric current was switched off, the DNA pieces had been separated in order of size. The smallest DNA molecules were furthest away from where the original sample was loaded on to the gel.

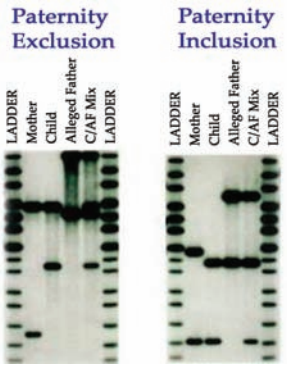


- Once the DNA had been sorted, the pieces of DNA were transferred or 'blotted' out of the fragile gel on to a robust piece of nylon membrane and then 'unzipped' to produce single strands of DNA.
- Next the nylon membrane was incubated with radioactive probes.

  - Probes are small fragments of minisatellite DNA tagged with radioactive phosphorous.
  - The probes only attach to the pieces of DNA that they are complementary? To – in this case they attach to the minisatellites in the genome.
- The minisatellites that the probes have attached to were then visualized by exposing the nylon membrane to X-ray film.
  - When exposed to radioactivity a pattern of more than 30 dark bands appeared on the film where the labelled DNA was. This pattern was the DNA fingerprint.
  - To compare two or more different DNA fingerprints the different DNA samples were run side-by-side on the same electrophoresis gel.

Benefits of DNA Fingerprinting

The most important benefit of DNA fingerprinting is that there is strong similarities shown between genetic fingerprints of parents and children. This is a benefit because a child's genetic fingerprint is made up of half the father's genetic information and half of the mother's information. This means that the bands of a child's genetic fingerprint will match the bands on both of their parents, making it possible to establish paternity and maternity tests.



Limitations of DNA Fingerprinting

One of the main problems with the process of DNA fingerprinting is that the sample can be easily ruined. The tiniest pieces of genetic junk can contaminate DNA samples, causing them to be useless. Although DNA fingerprinting requires a good sample to work with, this problem can be solved by using the newer technique called PCR. PCR can use extremely small samples of DNA and produce a much faster result. But this also means the DNA samples that PCR uses are even more likely to be contaminated because of their size, as it is harder to find a small sample with hardly any contamination. Another limitation of fingerprinting is that the procedure is so complex and hard to read the DNA patterns, that sometimes the juror finds the evidence almost invisible.

Although DNA Fingerprinting is a highly advanced process, there are still some things that it is unable to do. In dogs for example, a fingerprint does not make it possible to determine if the animal is a carrier of a disease causing allele. Also, a DNA fingerprint is unable to show a crossbreed in animals. This is because second or third generation crosses cannot be seen by working backwards in a pedigree. It may soon become possible to discover the crossbreed of dogs, although right now this is not possible.



CORE CONCEPT: CRYOGENIC PRESERVATION OF ANIMAL GERMPLASM

Cryogenic preservation of animal germplasm involves the long-term storage of reproductive cells, such as sperm and embryos, at extremely low temperatures. The process uses liquid nitrogen to freeze and maintain these cells in suspended animation, preserving their genetic material. This technique ensures the conservation of valuable genetic resources, genetic diversity, and endangered species, allowing for future use in artificial insemination, assisted reproduction, and genetic research.

Cryoprotectant is a substance used to protect biological tissue from freezing damage (i.e. that due to ice formation)

Spermatozoa were the first mammalian cells to be cryopreserved successfully. This success was due to the serendipitous discovery by Polge and co-workers of the **cryoprotective** effect of glycerol. Since then, many methods have been developed for various types of cells, tissues and organs. Much progress in the field has come from empirical work as well as from fundamental cryobiology. Increased understanding of the causes of cryo-injury has continually helped to improve cryopreservation methods.

The two most commonly used cryopreservation methods for animal germplasm are slow-freezing and vitrification. These are quite different methods, but relate to the same physico-chemical relationships. The differences between the two can be explained by first describing what happens during slow freezing.

Slow Freezing

In slow-freezing, cells in a medium are cooled to below freezing point. At some stage, ice masses containing pure crystalline water will form. What remains between the growing ice masses is the so-called unfrozen fraction, in which all cells and all solutes are confined. The concentrations of sugars, salts and cryoprotectant (e.g. glycerol) increase, while the volume of the unfrozen fraction decreases. The increase in osmotic strength causes an efflux of water from the cells. Slow cooling is needed in order to allow sufficient efflux of water to minimize the chance of intracellular ice formation. As cooling continues, the viscosity of the unfrozen fraction ultimately becomes too high for any further crystallization. The remaining unfrozen fraction turns into an amorphous solid that contains no ice crystals.

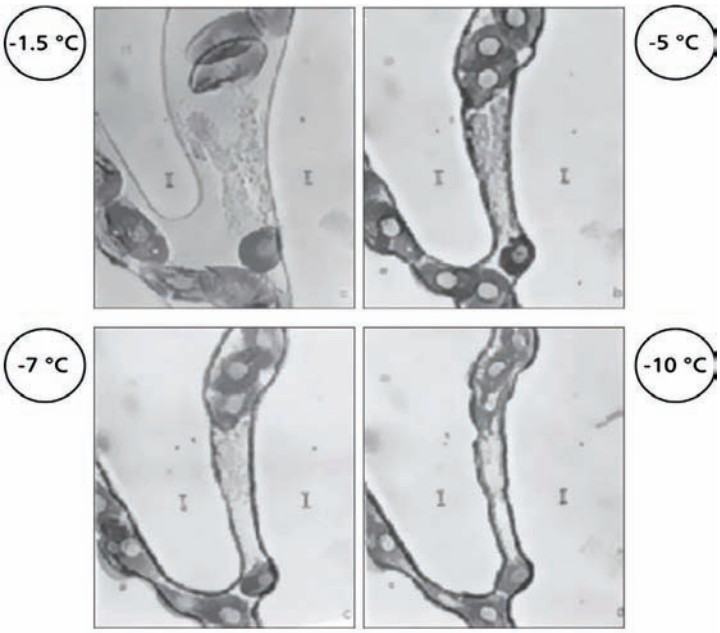


Figure 2: Frog erythrocytes in the "unfrozen fraction", which is enclosed by growing masses of ice

Chilling Injury and Cold Shock

The first challenge in cryopreserving cells from homeotherm (warm-blooded) animals is in cooling the cells below body temperature. Cells may be damaged by very rapid cooling (cold shock) or be damaged by low temperature per se (chilling injury). Behavior and function of membrane lipids and proteins may be affected by temperature. For example, membrane lipids that are normally in a liquid crystalline state may solidify at non-physiological temperatures, which can change their function and begin processes such as cryocapacitation of the production of reactive oxygen species that increase damage to membranes. Decreasing the temperature may cause an imbalance in cellular processes because the rate of one process may be affected more strongly than that of another. One example is the disintegration of the metaphase

spindle of oocytes caused by a change in the dynamic equilibrium of the association/dissociation of the tubulin filaments.

Supercooling

In slow-freezing methods cells are brought into a suitable freezing medium and cooling is continued below the freezing point of the medium. Ice formation does not necessarily start at the freezing point. Small ice crystals have a lower melting/freezing point than "bulk" ice, due to their large surface tension. Spontaneous ice nucleation will in most cases occur after the solution is supercooled to a temperature between -5 and -15 °C. Thereafter, ice will grow rapidly in all directions, and the release of the latent heat of fusion will cause the sample to warm up abruptly until the freezing/melting temperature of the solution (i.e. of the remaining unfrozen fraction) is reached. At this point, the ice formation will stop, or will proceed at a rate governed by the rate at which the heat of fusion is transported from the sample. Finally, the sample can "catch up" again with the lower temperature in the freezing apparatus. From a practical perspective, this means that the cells undergoing cryopreservation in a typical semen straw have to withstand a series of large and abrupt temperature changes.

Conditions in the Unfrozen Fraction

Cells are faced with very high concentrations of solutes in the unfrozen fraction. Dehydration and high salt concentration may result in loss of stability in the membranes or denaturation of proteins. Moreover, high salt concentrations may cause extracellular salts to enter the cells, a process known as "solute loading". The fast efflux of water causes a rapid decrease in the volume of the cells to approximately 50 % of their original volume. This leads to structural deformation of the cells. Further mechanical stress may be caused by cells being confined in very narrow channels of unfrozen solution and squeezed between growing masses of ice.

The Influence of Cryoprotectants

At all practical cooling rates, the total solute concentration (which is measured in moles per kg water) is determined only by the subzero temperature (Figure 3). When the initial freezing medium contains only salts (electrolytes), salt concentrations in the unfrozen fraction will reach extremely high levels as the temperature decreases. In contrast, in a medium that contains a large proportion of non-electrolytes, the total solute concentration at each subzero temperature will be the same as that found at the equivalent temperature in a medium containing only salts; however, the salt concentration will be much lower.

Sugars can be used as non-electrolyte solutes, but they will only affect the extracellular salt concentration. Moreover, high concentrations of impermeable solutes impose osmotic stress on the cells already before freezing. This is much less the case when a membrane permeable solute, such as glycerol, is used rather than a non-permeable solute. When cells are brought into a hypertonic glycerol medium,

water will leave the cells because of the osmotic pressure difference. However, at the same time, glycerol will enter the cells. After a short period of equilibration, the cells will have regained their original volume. The osmotic stress imposed by a hypertonic glycerol solution is therefore much smaller than that imposed by a hypertonic sugar solution. Hence, glycerol can be used at greater concentrations than sugars without damaging the cells. A substantial initial glycerol concentration in the medium means that part of the extracellular and intracellular water is replaced by the glycerol. Hence, the amount of ice formed is lower, the unfrozen fraction remains larger, the degree of shrinkage of the cells is limited, and the electrolyte concentration in the unfrozen solution and in the cells will be relatively small (see Figure 3). The mechanisms through which other membrane permeable substances, such as ethylene glycol and dimethyl-sulfoxide (DMSO), provide cryoprotection are similar to those involving glycerol.

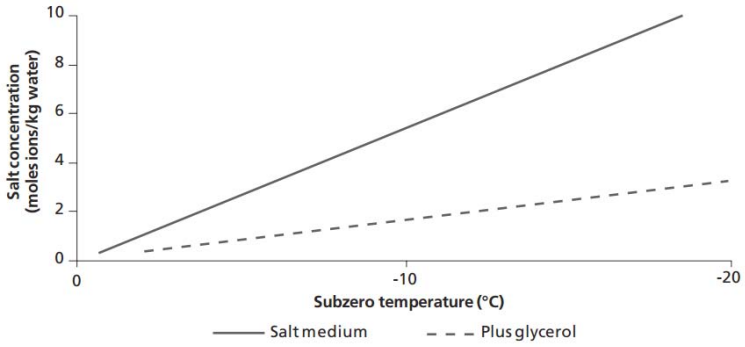


Figure 3: Effect of subzero temperature on salt concentration in the presence and absence of glycerol.

There are additional mechanisms through which polyols, such as like glycerol and several sugars, provide cryoprotection. These substances can stabilize lipid membranes by hydrogen bonding with the polar head groups of membrane lipids, which is especially important under severely dehydrated conditions. In addition, these substances may affect the mechanical properties of the unfrozen fraction, especially its viscosity and glass-forming tendency. The degree to which cells shrink and re-swell after addition of a membrane-permeable cryoprotectant depends on the concentration of the cryoprotectant and the relative permeability of the membrane to water and to the cryoprotectant. For instance, bull sperm shrink very little when brought into a freezing medium with glycerol, whereas bovine embryos react much more strongly. Upon thawing, removal of the cryoprotectant has the opposite effect on cells: they first swell and then they shrink again. This may lead to damage if the cells expand too much. Damage due to over-swelling of cells can be prevented by stepwise removal of the cryoprotectant.

The Influence of Cooling Rate

A general observation in the cryopreservation of cells and other biological systems is that each system has a specific optimal cooling rate, with decreased survival at cooling

rates that are too low (slow-cooling damage) or too high (fast-cooling damage). Ice growth is a rapid process, but transport of water through the cell membrane is relatively slow, because the membrane acts as a resistance barrier. Therefore, as cooling and extracellular ice growth continue, the liquid water of the unfrozen fraction remains very close to equilibrium with the ice, but the intracellular water lags behind. This means that the water concentration (i.e. the chemical potential of water) is too high for thermodynamic equilibrium, and there may be a risk of intracellular ice formation.

The optimal cooling rate falls in a range that is neither too fast nor too slow. When cells are cooled very slowly, the intracellular water lags behind only a little, and the risk of intracellular ice formation is minimal. However, it also means that the dehydration of the cells is maximal, which is not desired. At higher cooling rates, intracellular dehydration, intracellular solute concentration and shrinkage of the cells is less excessive. Moreover, the cells are exposed to the unfavorable conditions for a shorter period of time. However, when cooling rates are increased too much, the dehydration may not be fast enough to prevent intracellular ice nucleation. Fast-cooling damage can also be caused by other factors. For instance, it has been proposed that rapid water flow through membrane pores could lead to an uneven distribution of pressure on the membrane. Fast-cooling damage could also result from the very sudden changes in size, shape and ultrastructure, caused by the rapid efflux of water.

Different cells or other biological materials (embryos, tissue pieces) may have different optimal cooling rates. The optimal cooling rate of cells is largely determined by their volume and their membrane surface area (volume to surface area ratio), and by the permeability of the membrane to water and to cryoprotectant.

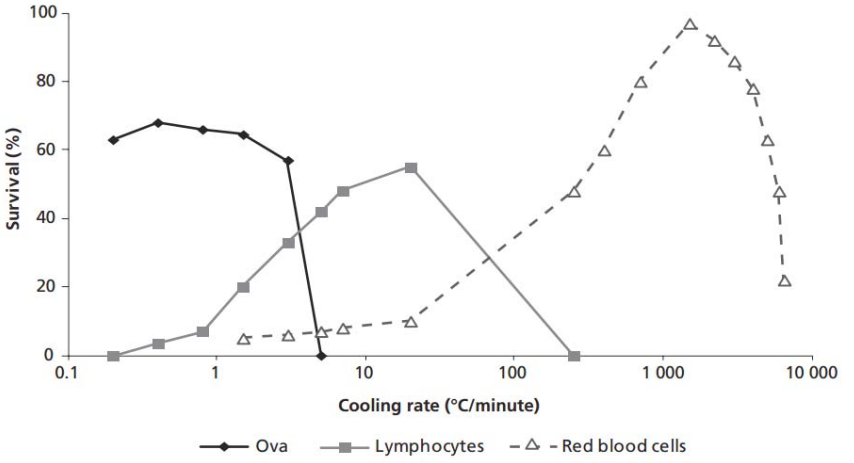


Figure 4: Effect of cooling rate on survival of different kinds of cells.

Programmable and Non-programmable Freezers

Biological material can either be frozen using quite simple, non-programmable, freezers or using more sophisticated, programmable, freezers (see Figure 5). Although programmable freezers are more expensive, they do not necessarily yield

more satisfactory results, especially for experienced technicians and cryobiologists. Therefore, the choice between programmable and non-programmable systems will depend on the financial resources available and the experience of the technicians. In some cases, even the most experienced technicians prefer the operating simplicity of programmable models.

In most programmable freezers, the straws or vials are cooled by cold nitrogen vapor. The temperature inside the cooling chamber can be accurately controlled and the time course of the temperature can be programmed. However, the time course of temperature inside the straws may be different due to the generation of heat of fusion (Figure 6). In non-programmable freezers, the straws may be cooled by being exposed to vapor (or a cold surface) at a constant low temperature. An example of a simple system is the freezing of straws placed on a rack in a Styrofoam box partially filled with liquid nitrogen without ventilation. The height of the straws above the liquid nitrogen determines the rate of heat exchange. Alternatively, straws can be placed on a piece of Styrofoam that floats on the liquid nitrogen. The thickness of the Styrofoam piece determines the rate of heat exchange.



Figure 5: Example of a programmable freezer.

Generally in such systems, the rate of heat exchange is governed by the temperature difference between the inside and the outside of the straw, and by the extent of heat conduction. The latter depends strongly on the volume to surface ratio of the straw or vial and the rate of (forced) ventilation. Therefore, it is difficult to compare one type of non-programmable freezer with another, or to know the actual freezing rate obtained with any given non-programmable apparatus. Optimal conditions have to be determined by experimentation.

Non-programmable systems do, however, have an advantage. The cooling curve (the time course of cooling and freezing) is, by default, of the form theoretically predicted to be optimal for slow freezing, with relatively low cooling rates directly after ice formation begins and higher cooling rates later.. Thus, the heat of fusion liberated during ice formation slows the cooling exactly at the point when cells need extra time to export intracellular water. The overall steepness of the freezing curve can be adjusted in non-programmable systems by choosing the height of the straws above the liquid nitrogen, which is proportional to the temperature of the vapor

around the straws (Figure 6). In more sophisticated systems with forced ventilation and adjustable preset vapor temperatures, the rate of heat exchange can be adjusted by choosing the preset vapor temperature.

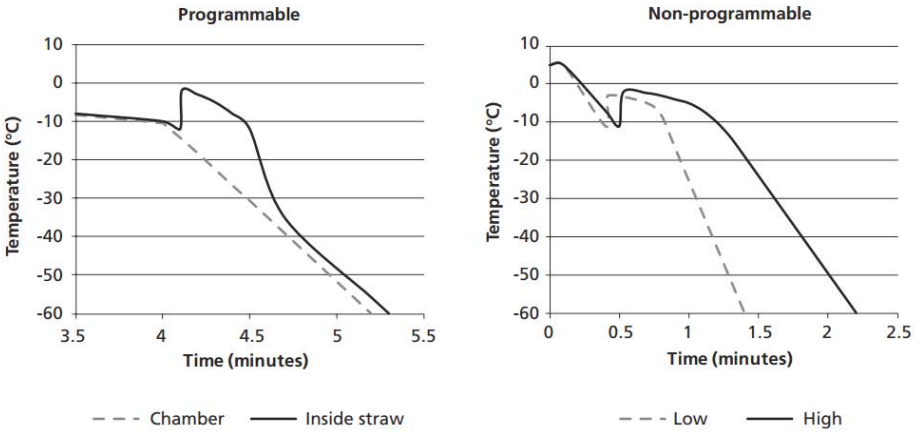


Figure 6: Examples of freezing curves in programmable and non-programmable freezers.

Vitrification

The term "vitrification" refers to any process resulting in "glass formation", the transformation from a liquid to a solid in the absence of crystallization. According to this definition, cells that are properly slow frozen become "vitrified".

If, in slow-cooling methods, cells ultimately become vitrified, how do so-called vitrification methods differ? Vitrification methods involve the use of a medium that has a very high solute concentration to begin with. Thus, ice cannot form in any part of the sample. As no ice forms, cooling does not have to be slow. In fact, it may be beneficial to cool very rapidly. The vitrified state and the associated physico-chemical conditions obtained using vitrification methods, are to some extent similar to those obtained by slow cooling, but the way of reaching this point is quite different.

Chilling Injury and cold Shock

As in the case of slow-freezing methods, vitrification methods can damage cells or tissues through cold shock and chilling injury. Depending on the material and the protocol used, however, cells or tissues may be rapidly cooled from a temperature at which chilling injury and cold shock play no role (e.g. room temperature). Extremely high rates of cooling from such a temperature to the vitrified state seem to be able to "outrun" cold shock and chilling injury. For example, rapid cooling seems to prevent disintegration of the metaphase spindle of oocytes.

The bulk of the ice formation happens in the temperature range between the freezing point and -10 °C, and consequently most of the water efflux from the cells must also take place in this temperature range.

Cryoprotective Agents

In vitrification methods, cells or tissues are brought into a medium that has a very high concentration of cryoprotective agents, also known as cryoprotectants. If the concentration of solutes is high enough, vitrification solutions will solidify to a glass without any risk of intracellular or extracellular ice formation during cooling or warming, independently of the cooling and warming rates used. However, the very high concentrations of cryoprotective agent needed for vitrification may cause damage due to abrupt osmotic changes, extremely low water potential or chemical toxicity. The embryos are first equilibrated with 25 percent vitrification solution at room temperature. Then the embryos are cooled to 4 °C and transferred to 50 percent vitrification solution and then to 100 percent vitrification solution. They are then rapidly packed and transferred into liquid nitrogen. The stepwise increase of cryoprotective agent concentration reduces osmotic effects, while the low temperature and rapid transfer help prevent damage by chemical toxicity. In addition, chemical toxicity may be reduced by using mixtures of various permeant CPAs, or addition of non-permeant CPAs (60 g/litre polyethylene glycol) or 60 g/litre bovine serum albumin (BSA).

Reduction of Cryoprotective Agent Concentration at High Cooling Rates

Solutions that have a solute concentration lower than that of classical vitrification solutions have freezing points below which there is a significant tendency to form ice crystals. But when the solution is cooled very rapidly, there is simply no time for ice formation. Below a certain temperature, the solution becomes so viscous and stiff that ice formation becomes impossible, and the solution turns into “metastable” glass. The solute concentration needed for metastable vitrification decreases as a function of increasing cooling rate. The most recent vitrification procedures, therefore, make use of high cooling rates in order to reduce the concentration of CPAs and thereby decrease the damage caused by osmotic stress and chemical toxicity.

CONCEPT OF GENE PRODUCTS IN GENE LIBRARY

A gene library, also known as a genomic library or DNA library, is a collection of cloned DNA fragments representing the entire genetic information (genome) of an organism. In a gene library, each cloned DNA fragment corresponds to a specific gene or part of a gene. The concept of gene products in a gene library refers to the proteins or functional RNA molecules that are encoded by the genes represented in the library.

When constructing a gene library, researchers extract genomic DNA from the organism of interest and enzymatically cut it into smaller fragments. These fragments are then individually inserted into a vector, such as a plasmid or a viral genome, which acts as a carrier to replicate the DNA in host cells. The host cells are usually bacteria or yeast.

Once the gene library is created, it becomes a valuable resource for studying the genetic content of an organism comprehensively. It allows researchers to analyze and manipulate the genes of interest, leading to a better understanding of gene structure, function, and regulation.



To identify gene products within the library, various screening methods can be employed. One common approach involves using labeled gene probes, which are complementary DNA or RNA sequences that bind specifically to the target genes. By screening the library with these probes, researchers can identify clones that carry the genes of interest.

Furthermore, gene products in a gene library can also be studied using expression systems. For example, the genes can be inserted into expression vectors that drive the production of their respective proteins in host cells. This enables researchers to study the function and properties of the gene products in a controlled environment. Overall, the concept of gene products in a gene library is essential for unraveling the molecular basis of various biological processes and diseases. The availability of gene libraries has significantly accelerated research in genomics, genetics, and molecular biology, allowing scientists to explore the vast complexities of the genetic information encoded within an organism's genome.

SUMMARY

- Drosophila was one of the first organisms to be studied genetically: its small size, short life cycle (10 ~14 days at 25^{o}
- C), high reproductive rate (an adult female can lay 400-500 eggs in 10 days), and ease of culture and genetic manipulation have made it perhaps the best understood animal genetic system.
- A gene is a heritable factor that controls the expression of some trait, which may be morphological, behavioral, molecular, etc.
- Drosophila, like most species we will deal with in this course, are diploid, with two sets of chromosomes and therefore two alleles at each autosomal locus.
- DNA fragments are cut out of their normal position in the chromosome using restriction enzymes (also called restriction endonucleases) and then inserted into other chromosomes or DNA molecules using enzymes called ligases.
- A gene library is a large collection of cloned DNA sequences from a single genome. A genomic library in theory, would contain at least one copy of every sequence in an organism's genome. These are used to investigate the structure of a given chromosome, or to clone specific genes.
- Karyotyping is a laboratory procedure that allows a physician to examine a patient's set of chromosomes. "Karyotype" also refers to the actual collection of chromosomes being examined.
- Chromosome mapping is the process of determining the locations of specific genes on a chromosome. Chromosomes are collections of deoxyribonucleic acid, or DNA that contain genes, although it is not well understood where exactly these genes lie within chromosomes.
- DNA fingerprinting, also called DNA typing, DNA profiling, genetic fingerprinting, genotyping, or identity testing, in genetics, method of isolating and identifying variable elements within the base-pair sequence of DNA (deoxyribonucleic acid).

MULTIPLE CHOICE QUESTIONS

1. A gene library is a large collection of cloned DNA sequences from a single genome.

a. True

b. False

2. Which of the following is/are the method of transfection for making transgenic animals?

a. Transfer of whole nuclei

b. Transfer of whole individual chromosomes or fragment

c. Transfer of DNA

d. All of the above

3. Dihybrid crosses involve manipulation and analysis of two traits controlled by pairs of alleles at different loci.

a. True

b. False

4. Karyotyping is a laboratory procedure that allows a physician to examine a patient's set of chromosomes.

a. Chromosome mapping

b. DNA

c. RNA

d. Karyotyping

5. In the laboratory, DNA molecules can be cut at specific sequences using:

a. UV light.

b. restriction nucleases.

c. DNA ligase.

d. laser tweezers.

Answer

1. (a) 2. (d) 3. (a) 4. (d) 5. (b)

REVIEW QUESTIONS

1. What do you understand by gene cloning?
2. Describe the blood typing.
3. What is DNA fingerprinting? Discuss the limitations of DNA fingerprinting.
4. Explain the cryogenic preservation of animal germplasm.

REFERENCES

1. Arnold ML, Holman D, Zweifel S. 2017. Using molecular biology and bioinformatics to investigate the prevalence of mislabeled fish samples. The American Biology Teacher 79, 763-768.
2. Eur J Hum Genet. 2008 Mar-Apr;6(2):165-75. Evaluation of CFTR gene mutation testing methods in 136 diagnostic laboratories: report of a large European external quality assessment.
3. Harris RP, Helfand M, Woolf SH, et al. Current methods of the U.S. Preventive Services Task Force: a review of the process. Am J Prev Med. 2001;20(3 Suppl 1):21–35.
4. Hum Pathol. 2009 Feb;40(2):147-55. Analysis: Secretary's Advisory Committee on Genetics, Health, and Society Report falls short.
5. Myriad Genetic Laboratories. Mutation Prevalence Tables. Salt Lake City, UT: Myriad Genetic Laboratories; 2010. [26 November 2013].
6. Nelson HD, Huffman LH, Fu R, et al. Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med. 2005;143(5):362–79.
7. Shattuck-Eidens D, Oliphant A, McClure M, et al. BRCA1 sequence analysis in women at high risk for susceptibility mutations: risk factor analysis and implications for genetic testing. JAMA. 2007;278(15):1242–50.

Index

A

Abdomen 12
Anaphase II 149, 152
Animal genetic system 233, 261

B

biochemistry 78
biotechnologies 241
Blood 244, 245, 247, 248

C

Centimorgan 209
Centromeres 132
centrosome 131, 133
Chromosomal theory of inheritance 13, 37
Chromosome 4, 7, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 26, 169, 186, 187, 195, 197, 198, 199, 200, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 216, 217, 218, 219, 220, 221, 222, 223, 224, 228
Chromosome mapping 261, 262
Comparative genomic hybridization (CGH) 28
crystals 253, 254, 259
Cytokinesis separates 149

D

Deoxyribonucleic acid (DNA) 199
Dihybrid crosses 236, 262
Dinomitosis 139
DNA fingerprinting 248, 249, 251, 261, 263
DNA Functions 167, 183
DNA molecule 172, 178, 179, 184, 186
DNA replication 183, 184
DNA structure 171
Double helix 169, 171, 172, 175, 176, 177, 178, 180, 182, 192
Drosophila 233, 234, 236, 261

E

Enzymes 77
Extranuclear 139

F

Fertilization 5, 7, 15, 18, 153, 154, 155, 156
Fluorescence in situ hybridization (FISH) 28
Fluorescence In-Situ Hybridization (FISH) 220

G

Gametes 5, 7, 13, 14, 16
Gene 234, 235, 236, 238, 240, 241, 261, 262, 263

Gene Libraries 240
Generation 202, 203
Gene therapy 222
Genetic content 205
Genetic Crosses 235
Genetic Linkage 206
Genetic marker 212
Genetics 201, 205, 210, 215
Genome 26, 27, 28, 29, 30
genomic library 240, 261
Genomic sequence 27, 29
Glycosidic bonds 179

H

Heredity 4, 5, 7
heterozygote 55, 57, 58, 77
High reproductive rate 233, 261
Homologous chromosomes 143, 144, 145, 146, 147, 148, 149, 152, 157, 158
Homozygous 7, 8, 10, 11, 12, 16, 24
Human genome 4, 26, 27, 28, 29, 31, 37
Human karyotype 18, 37
Human reproduction 216
Hybridization 219, 220, 221
Hydrogen bonds 169, 176, 177, 178

I

Incomplete dominance 8
Independent Assortment 206
Infancy 21
Interkinesis 147

K

Karyotyping 242, 261, 262
Kinetochore 132, 133, 145, 148, 149, 152
Kinetochore microtubules 132, 133, 149

L

Laboratory 29, 206
Law of independent assortment 197
Linkage analysis 202
Linkage disequilibrium 201, 202, 228
lipid metabolism 58
Logarithm of the odds 204

M

Major histocompatibility complex (MHC) 220
Malaria 24
Meiosis 197, 198, 199, 204, 205, 206, 214, 223
Mendelian genetics 4, 52
Metaphase plate 133, 136, 137, 145, 147, 152
Microtubule organizing center 131
Mitosis 127, 129, 130, 138, 140, 150, 151, 157
MITOSIS 129
Molecular biology 14, 26
monohybrid cross 235
Monohybrid Cross 235
Mutation 5, 24

O

Offspring 198, 199, 200, 201, 204, 206, 207, 208, 214, 216, 223

P

Parent cell 205
phenotype 53, 56, 57, 58, 59, 60, 79, 80
Phosphodiester 173, 174, 192
Polymerase chain reaction (PCR) 231, 238, 249
Prometaphase 130, 132, 143, 145, 149
Prophase 129, 130, 131, 132, 135, 136, 137, 143, 144, 145, 158, 159

R

Randomly amplified polymorphic DNAs (RAPDs) 212
Recombination 205, 206, 208, 209, 213
Recombination fraction 213, 214, 215
Recombination frequency (RF) 208
Restriction fragment length polymorphisms (RFLPs) 211, 213
Rh factor 24
Ribonucleic acid (RNA) 212

S

Sex determination 15
Sexual reproduction 142, 152, 157

Simple sequence length polymorphisms (SSLPs) 213
Single nucleotide polymorphisms (SNPs) 213
slow-freezing 252, 253, 254, 258
Surveillance 32
Synaptonemal 143, 144, 145, 152

T

Temperature 203

V

Variable number of tandem repeats 213
Variation 4, 5

Z

Zygote 13, 17, 202

Core Concepts in Biology: Genetics

2nd Edition

Genetics refers to the study of heredity and variation in living organisms. It has captivated scientists and enthusiasts alike for centuries. It offers a fascinating glimpse into how traits and characteristics are passed down from one generation to the next, shedding light on the fundamental building blocks of life itself. Genetics plays a crucial role in the development of society. Understanding how genes shape our traits, behaviors, and susceptibility to diseases opens up new pathways for medical advancements and personalized treatments. The field of genetics is vast and ever-evolving, comprising several key areas including molecular genetics, population genetics, and genetic engineering. At the core of genetic research lies the study of DNA - the blueprint of life. Examining the structure, function, and expression of genes enables scientists to comprehend their influence on human development and inheritance patterns. By deciphering these complex genetic codes, researchers are able to diagnose various diseases at an early stage when interventions can be more effective. Furthermore, insights derived from population genetics provide valuable information about our evolutionary history and diversity within different groups. This knowledge promotes inclusivity by highlighting the importance of preserving genetic variety for future generations. Additionally, it aids in understanding how populations adapt to environmental changes over time. The concept of genetic engineering has revolutionized agriculture by enhancing crop yields through manipulation of plant genes for improved characteristics such as disease resistance or enhanced nutrition content. These innovations have profound implications for food security worldwide as agriculture becomes increasingly critical due to rapid global population growth. Moreover, advancements in pharmacogenomics allow tailor-made medications based on individual variations in gene expressions or mutations. This approach optimizes drug therapies while minimizing adverse side effects through a better understanding of each person's unique genomic profile.

In this second edition of Core Concepts in Genetics, we embark on an intellectual journey through the intricate world of genetic principles. The overarching goal of this book is to provide readers with a comprehensive understanding of genetics, drawing upon both classical theories and modern advancements. By delving into concepts such as DNA replication, gene expression, Mendelian inheritance patterns, genetic engineering techniques, and molecular genetics research methods – among others – we aim to equip individuals with a thorough grasp of these indispensable foundations.

Organized into concise chapters carefully structured for educational efficacy, Core Concepts in Genetics serves as an invaluable resource not only for students pursuing degrees or careers related to biological sciences but also for anyone seeking a deeper comprehension of our genetic blueprint. Each chapter seamlessly blends critical concepts with illustrative examples and thought-provoking exercises designed to enhance understanding while honing analytical skills.

Moreover, this revised edition incorporates up-to-date content reflecting recent breakthroughs that have reshaped our understanding of genetics. The constantly evolving nature of this field necessitates avoiding complacency when it comes to knowledge acquisition; therefore seizing upon cutting-edge developments allows us to fortify students' grasp on contemporary genetic principles.

The Book comes with access code for rich learning experience, which includes:

1. Interactive E-lecture of each chapter. E-lectures are expressive, informational, entertaining and persuasive, it uses the tool of self-exploration, which makes it easy to learn and understand each topic in detail. It is very informative as concrete details are provided and also entertaining, as graphics and other visuals are provided to make the learning process more interactive.
2. Video Lecture of each chapter, which explains each topic in detail with examples, animations, images and text and makes it easy to understand the topics in easier, simpler and better way.
3. Quiz for each chapter

It is also a useful tool for teachers to teach with digital resources in classroom and do a great job of illustrating skills and techniques that are otherwise difficult to explain.

Kate Fickas is an associate professor in the Department of Genetics, specializing in the fields of aging, cancer, genetics, and clinical research. With a wealth of knowledge and experience, she is dedicated to exploring the intricacies of these areas and advancing our understanding of them. Through her research, Kate aims to shed light on the complex mechanisms of aging, unravel the mysteries of cancer development, delve into the study of genetics, and contribute to the advancement of clinical research. Her expertise and contributions in these fields make her a valuable asset in the scientific community.

ISBN 978-1-98469-144-6

BIBLIOTEX
Digital Library

www.3ge-learning.com
email: info@3ge-learning.com