School of Social Sciences
Indira Gandhi National Open University

BANC – 111

Human Population Genetics

BANC-111

HUMAN POPULATION GENETICS

School of Social Sciences

Indira Gandhi National Open University

EXPERT COMMITTEE

Prof. A. K. Kapoor (retd), Department of Anthropology, University of Delhi, Delhi

Prof. Gautam Kumar Kshatriya (retd), Department of Anthropology, University of Delhi, Delhi

Prof. Rashmi Sinha, Discipline of Anthropology, IGNOU

Dr. Rukshana Zaman, Discipline of Anthropology, IGNOU

Dr. Mitoo Das, Discipline of Anthropology, IGNOU

Dr. Palla Venkatramana, Discipline of Anthropology, IGNOU

Dr. K. Anil Kumar, Discipline of Anthropology, IGNOU

COURSE PREPARATION TEAM

BLOCK		UNIT WRITER
Block 1	**Introduction to Human Population Genetics**	
Unit 1	Essentials of Population Genetics	Dr. Kumud Sarin, Indian Biosciences and Research Institute, Noida Dr. Gangaina Kameih, PhD (Anthropology), Thangmeiband Yumnam Leikai, Imphal West, Manipur
Unit 2	Haemoglobin Polymorphism and Thalassaemia	Dr. Palla Venkatramana, Discipline of Anthropology, SOSS, IGNOU
Unit 3	Genetics of Complex Diseases	Dr. Sudipta Ghosh, Department of Anthropology, North-Eastern Hill University, Shillong, Meghalaya
Block 2	**Genetic Structure of Human Populations**	
Unit 4	Hardy-Weinberg Principle	Dr. Gangaina Kameih, PhD (Anthropology), Thangmeiband Yumnam Leikai, Imphal West, Manipur
Unit 5	Mechanisms of Evolution	Dr. SAA Latheef, Department of Genetics and Biotechnology, Osmania University, Hyderabad
Unit 6	Genetic Polymorphisms	Dr. Ketaki Chandiok, PhD (Anthropology), Delhi
Block 3	**Human Population Structure and Disease Pattern**	
Unit 7	Mating Patterns	Khangembam Taibanganba Meitei and Dr. Huidrom Suraj Singh, Department of Anthropology, Manipur University, Imphal
Unit 8	Biological Consequences of Mating Systems	Thangjam Chitralekha Devi and Dr. Huidrom Suraj Singh, Department of Anthropology, Manipur University, Imphal
Unit 9	Population and Disease Association Studies	Dr. Kumud Sarin, Indian Biosciences and Research Institute, Noida Dr. SAA Latheef, Department of Genetics and Biotechnology, Osmania University, Hyderabad
Unit 10	Comparative Biology	Dr S. Yaiphaba Meitei and Henry Konjengbam, Department of Anthropology, Manipur University, Imphal

Practical Manual		
Unit 1	Blood Group Typing-ABO and Rh (D) Blood Groups	Dr. SAA Latheef, Department of Genetics and Biotechnology, Osmania University, Hyderabad
Unit 2	Colour Blindness	Dr. Palla Venkatramana, Discipline of Anthropology, SOSS, IGNOU
Unit 3	Glucose-6-Phosphate Dehydrogenase Deficiency (G6PD)	Dr. SAA Latheef, Department of Genetics and Biotechnology, Osmania University, Hyderabad
Unit 4	PTC Tasting Ability	Dr. Palla Venkatramana, Discipline of Anthropology, SOSS, IGNOU

Course Coordinator

Dr. Palla Venkatramana, Discipline of Anthropology, School of Social Sciences, IGNOU, New Delhi

General Editors

Prof. Gautam Kumar Kshatriya (retd), Department of Anthropology, University of Delhi, Delhi
Dr. Palla Venkatramana, Discipline of Anthropology, SOSS, IGNOU
Academic Consultancy : Dr. Smarika Awasthi Sharma

PRINT PRODUCTIONS

Mr. Rajiv Girdhar
Assistant Registrar,
MPDD, IGNOU, New Delhi

Mr. Hemant Kumar Parida
Section Officer
MPDD, IGNOU, New Delhi

January, 2022

ISBN:

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi by the Registrar, MPDD, IGNOU, New Delhi.

Composed by : Hi-Tech Graphics, D-4/3, Okhla Industrial Area, Phase-II, New Delhi-110020.

Print by:

ignou
THE PEOPLE'S
UNIVERSITY

COURSE CONTENTS

ignou
THE PEOPLE'S
UNIVERSITY

COURSE INTRODUCTION

Anthropology appreciates the significance of learning the similarities and differences within and between population groups in context of their cultural and genetic backgrounds. Understanding the biological attributes of human populations; and their interaction with culture is the principal component of biological anthropology. Population genetics is an integral part of biological anthropology in understanding human evolution. In view of the depletion in single nucleotide polymorphisms and increase in multigenetic disorders, population geneticists are challenged to understand the pattern of allele distribution in the presence of evolutionary forces.

Course Presentation

The course, Human Population Genetics consists of three blocks of theory and a practical manual.

Block-1 (Unit 1 through Unit 3) deals with the concept and scope of population genetics including complex genetic disorders. Unit 1 describes the essentials of population genetics, its concept and scope, and also focuses on the principles of Mendelian genetics. Unit 2 discusses in detail the haemoglobin polymorphism and thalassaemia giving an account of their distribution, consequences and mutation. Unit 3 focuses on the genetics of complex diseases, their prevalence and issues concerning public health.

Block-2 (Unit 4 through Unit 6) highlights Genetic equilibrium of the populations through Hardy-Weinberg principle and the predisposition of various diseases. Unit 4 explains the Hardy-Weinberg principle highlighting the frequencies, assumptions, applications and exceptions associated with it. Unit 5 covers the mechanisms of evolution including mutation, selection, genetic drift and migration, with adequate examples from India as well as world populations. Unit 6 gives an insight into types of Genetic polymorphisms and the models explaining the maintenance of genetic polymorphism.

Block-3 (Unit 7 through Unit 10) discusses mating patterns, the association of blood groups with diseases and the comparative account of various characters between man and apes. Unit 7 describes the various types of mating patterns. Unit 8 focuses on biological consequences of mating systems, covering areas like parental consanguinity, inbreeding, genetic load, etc. Unit 9 explores the association between populations and diseases prevalent among them, primarily the blood groups associated with diseases. Unit 10 gives an account of the comparative study of human and apes with reference to their biological and morphological characters.

The Practical Manual covers the blood group typing- the ABO and RH (D) blood group systems. It explains the classification of blood types according to the different types of antigens in the red blood cells and antibodies in the plasma. It also discusses Colour blindness, when one is not able to distinguish between certain colors. Then it explores Glucose-6-phosphate dehydrogenase deficiency (G6PD), a genetic disorder that mainly affects red blood cells, which carry oxygen from the lungs to tissues throughout the body. Further, the practical manual, describes the PTC-tasting ability, which is a simple genetic trait governed by a pair of alleles.

ignou
THE PEOPLE'S
UNIVERSITY

BLOCK 1
Introduction to Human Population Genetics

BLOCK 1 : INTRODUCTION TO HUMAN POPULATION GENETICS

UNIT 1 ESSENTIALS OF POPULATION GENETICS

Contents

Learning Objectives

After reading this unit, you would be able to:

- Discuss the concept and scope of Population Genetics;
- Explain various evolutionary forces; and
- Elucidate the principles of Mendelian Genetics.

1.0 INTRODUCTION

Modern humans evolved in Africa about 200,000 years ago and they migrated across different regions of the world (Lachance and Tishkoff, 2013). While travelling through diverse conditions, they underwent modifications both

Contributed by **Dr. Kumud Sarin**, Indian Biosciences and Research Institute, Noida and **Dr. Gangaina Kameih**, PhD (Anthropology), Thangmeiband Yumnam Leikai, Imphal West, Manipur

culturally as well as biologically to adapt to the environmental pressure for successful survival. These modifications left traces of variation in the human genetic structure. Thus, studying the genetic structure of human population can help us understand these modifications as well as the reasons behind them. Genetic evidence is given by patterns in variation in complete or a specific region of human genome. That means if you want to see the genetic evidence of any evolutionary force, you need to screen either the region of a DNA or you can scan a complete genome. It is useful as genetic evidence were imprints of the earliest evolutionary events. We also call these imprints as genetic signature of the evolution. However, screening genetic variation without the knowledge of other factors such as geography, environmental factors, and ethno-historic information can lead to incomplete understanding of this diversity. It means along with genetic information we also need information about other fields such as geography, environmental factors, and ethno-history of a population. An anthropological approach is therefore very important in gaining insight on genetic diversity by considering a range of factors to ascertain the quantity of variation observed. Dobzhansky rightly said that nothing in biology makes sense except in the light of evolution. So, in case if you are observing any biological phenomena, we need to understand them in terms of evolution. Evolutionary forces subject human population to change in biological traits. These traits are inherited, that means they are passed from one generation to the next generation. Studying these biological differences among human population in an evolutionary framework comes under the purview of biological anthropology in general and population genetics in particular.

Population genetics deals with the assessment of genetic differences underlying biological traits. It estimates allele and genotype frequencies and predicts the way they would change or remain constant over a period of time. Evolution basically means genetic change. Let's apply it in terms of allele or genotype frequency. If during a course of time, there is a change in allele or genotype frequency, we can say that evolution has occurred. The alterations in frequency are used in mathematical model to comprehend the interplay between different micro-evolutionary forces. It is not always that only a single micro-evolutionary force is acting upon the population. There can be many micro-evolutionary forces so it is always the joined effect of all these micro-evolutionary forces which bring alteration in the gene and genotype frequencies. The pattern of genetic variation is used to test hypothesis on demographic events that have occurred during the history of modern humans.

Population genetics can be defined as the study of the frequency of genes and genotypes in a Mendelian population and predicting the way these frequencies would change over a period of time under the influence of various evolutionary forces.

Check Your Progress

1) What is Population Genetics?

.....

.....

.....

.....

Population genetics is concerned with changes in genetic variations overtime. There are two ways of looking at genetic variation: variation within populations and variation between populations. The former refers to differences and similarities of individuals within a population; the latter refers to average differences between two or more populations. Let us use a simple analogy and consider you are in a large classroom filled with students, and imagine that we measured everyone's height. We would use these measurements to compute how much variation existed within the classroom. If for example, everyone in the class was of exactly the same height, there would be no variation. If, however, there were differences in height, with everyone being between 5 feet 8 inches tall and 5 feet 10 inches tall then the variation would exist because not everyone would be the same. If everyone were between 5 feet and 6 feet tall, there would be even more variations.

On the other hand, suppose that we want to compare the height in your classroom with the height in the next classroom. An example would be if the average heights in your classroom were 5 feet 9 inches and the average height in the other classroom were 5 feet 8 inches. The differences in average height would be 1 inch. This difference would be an example of variation between groups.

By studying genetic change over time and its effects on genetic variation within and between populations, we are able to apply a theory of population genetics to address a wide variety of questions about human biological variation and evolution. Some sample of such questions includes

- Where did the first humans come from?
- Why do some human populations have high frequencies of the harmful sickle cell allele?
- Genes are not resistant but are not susceptible to AIDS because of certain allelic combinations of genes.
- Why do some small populations differ genetically from their neighbours in terms of their genetic structure?
- What is the genetic structure of Zeliangrong Naga tribes of Manipur? Are these sub-tribes genetically similar?

Population genetics has also expanded its quest to understand the basis for genetic variation in susceptibility or resistance to disease and drugs responses. In recent years, new technologies and methodologies have been developed specifically to

explore the role of genetic variation in health and diseases. Proper understanding of human genetic variation is essential for the correct interpretation of the link between genetic variation and disease thereby its identification can assist in the development of diagnostic tests as well as effective treatments (ePG Pathshala, Anthropology, Paper no.08, Module 1).

1.1 EVOLUTIONARY FORCES SHAPING GENETIC VARIATION

The term evolution refers to all changes, large or small, visible or invisible, adaptive or non-adaptive (Kimura, 1983). In this context, existence of genetic variation becomes immensely important as they form the substrate for evolution. The key evolutionary forces that shape the pattern of genetic variation such as mutation, natural selection, gene flow, genetic drift, non-random mating and recombination are briefly mentioned in the sections below:

1.1.1 Mutation

Mutation is the change in the genetic materials- the change that is passed on from cell to cell (Das, 2008). Mutations are random changes in genetic code. Mutations are one of the fundamental forces of evolution because they fuel the variability in populations and thus enable evolutionary change. However, evolutionary consequences only follow from those changes that occur in germ-line (one that will give rise to gametes, i.e., egg or sperm cells) and not those occurring in somatic tissues (found elsewhere in an organism's body).

Mutations are caused by physical changes to the hereditary material and, because DNA is a long sequence of base pairs organized into physically unlinked chromosomes, there are many possible ways it can change (Loewe and Hill, 2010). There are (i) point mutations that change only a single letter and lead to so-called 'single nucleotide polymorphisms' in populations, (ii) insertions and deletions of various sizes (also called 'indels' if it is difficult to decide which of the two actually happened; these can also lead to 'copy number variants'), (iii) transpositions that move a sequence from one position to another, and can thereby cause mutations at the boundaries, (iv) inversions of various sizes that change the orientation of a stretch of DNA, (v) chromosome mutations that affect long enough pieces of DNA to become visible under the microscope and might even lead to the loss or duplication of a whole chromosome (also known as non-disjunctions), and (vi) changes in the ploidy level, where a whole copy of the genome is either gained or lost. Mutation essentially brings about new varieties.

1.1.2 Natural Selection

According to Charles Darwin natural selection is the most important mechanism by which evolution occurs. Natural selection is the process through which populations of living organisms adapt and change. Individuals in a population are naturally variable, meaning that they are all different in some ways. This variation means that some individuals have traits better suited to the environment

than others. Natural selection is entirely dependent on environment. Individuals with adaptive traits (traits that give them some advantage) are more likely to survive and reproduce. These individuals then pass the adaptive traits on to their offspring. Over time, these advantageous traits become more common in the population. Through this process of natural selection, favourable traits are transmitted through generations (National Geographic, Resource Library). Natural selection essentially involves differential survival and reproduction over generations.

Let us consider the classic example of peppered moth in England (Relethford, 2012). This species of moth comes in two forms, a dark coloured form and a light coloured form. Centuries ago, most of the moths were light coloured, and only about 1% was dark coloured. Dark-coloured moths were rare because more clearly visible against the light colour of the tree trunks, making it easier for birds to see them and eat them. Over time, the environment changed, and the frequency of dark coloured moths increased as the frequency of the light coloured moths decreased. Because the colour of the moths reflects genetic differences, this observed change is an example of the genetic composition of a species changing over time. In this case, the initial origin of a different form is due to mutation, and the change in moth colour over time reflects natural selection, because the environment had shifted following the Industrial Revolution, leading to darker tree trunks, thus creating a situation where dark coloured moths were less likely to be eaten by birds because of their survival advantage.

1.1.3 Gene Flow

Gene flow is the transfer of alleles from one population to another population through migration. Migration can cause some changes in the relative allele frequencies of the population either through immigration (when new organisms join a population, bring new alleles or changing frequencies) or by emigration (when members of a population leave it, taking with them their genes). As a result of migration, admixture is introduced. We can think of America to where coloured Africans as well as white Europeans migrated. There admixture took place between these populations giving rise to a mixed population. If you moved to another population and have a child, your genes have moved as well. Therefore, migration is one of the factors to bring about changes in allele frequencies in a population.

1.1.4 Genetic Drift

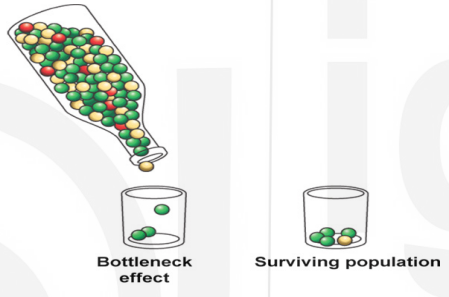
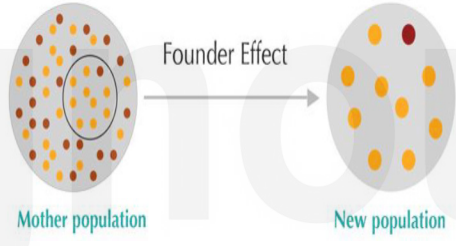
Genetic drift is random changes in allele frequencies in a population due to chance events. To be more exact, genetic drift is change due to "sampling error" in selecting the alleles for the next generation from the gene pool of the current generation. Its effects tend to be stronger in small populations.

Examples of genetic drift

- Bottleneck Effect

- Founder Effect

Table 1.1: Differences Between Bottleneck Effect and Founder Effect

Bottleneck Effect	Founder Effect
Reduction in population size	A new population is found by a small subset of the ancestral population
Subpopulation caused when a natural disaster reduces the size of an original larger population.	Founder establishes a subpopulation as a new colony drawn from a larger population.
Result of habit fragmentation and/or over exploitation of the species	Result of migration
Genetic drift that occurs after a bottleneck event with a little genetic variation	Genetic drift that occurs after the start of new population.
Cannot define	Can trace it back to one of the original communities
	

1.1.5 Inbreeding

The mating patterns of organisms in a population do not frequently exhibit random mating. In positive assortative mating, also known as consanguineous mating or inbreeding, mating between similar individuals occurs more frequently than chance would suggest. As a result, frequency of homozygotes increases compared to random mating. Negative assortative mating occurs when individuals with unlike genotype or phenotype mate. Both of these types of non-random mating affect genotype frequencies. Inbreeding tends to increase homozygosity and detrimental effects of inbreeding are well known.

Types of Inbreeding/Consanguineous Marriages

There are different types of inbreeding or consanguineous marriages that are practiced among human communities or societies. Depending on the type of consanguineous relationship between husband and wife, they are categorized as first degree and second-degree consanguineous marriages. The first cousin marriages are categorized as first-degree consanguinity. This involves individuals who belong to same generation but come from different families whose parents share a common ancestor (grandparental level). An individual if he or she marries his/her paternal or maternal daughter or son (uncle's daughter or son respectively) the cousin marriages are the first-cousin marriage. Similarly, in the case of second cousin marriages, the second cousins share a common ancestor at the great grandparent level. In between, there could be marriages

between individuals that belong to two generations e.g., one-and-half cousin marriages etc.

The cousin marriages could be of different types: for example, parallel, cross and double cousin marriages etc. The marriage between the cousins who were descendants of two brothers or sisters (or whose either parents are brothers or sisters) is referred as parallel cousin marriages. An individual marries his or her paternal uncle's daughter or son is referred as parallel cousin marriage. The word 'parallel' refers to same sex (either brothers or sisters) at the parental level. In case of cross-cousin marriages, it is the marriage between the offspring (cousins) who's (either) parents are brother and sisters. In Indian populations, Uncle-niece marriages are more common to Aunt-nephew marriages. Apart from the above two there could be one more type of cousin marriages, it is called as double cross cousin marriages. In double cross cousin marriage type both a boy and a girl marry their uncle's or aunts' girl or son (cousins) respectively. In Indian populations, Uncle-niece marriages are more common to Aunt-nephew marriages.

1.1.6 Recombination

Recombination refers to the exchange of genetic information between homologous chromosomes during meiosis. The process of recombination generates new combinations of alleles and in a way increases genetic diversity. Conversely, recombination can also disrupt the beneficial allele combinations, resulting in a drop in fitness known as "outbreeding depression". Alleles at closely placed loci generally do not segregate randomly during meiosis as recombination between them occurs infrequently. By virtue of this, evolutionary force operating on one locus affects the genetic diversity at the other locus. An allele that rises to high frequency through positive selection at a linked locus is said to be "Hitchhiking" while the reduction in the genetic diversity at loci linked to a recently fixed allele is called "Selective Sweep". Moreover, recombination rates are not uniform along a segment of chromosome. At population level, recombination can be measured indirectly by investigating whether specific alleles at different loci are associated with one another more or less often than would be expected by chance. This non-random association of alleles at different loci is termed as "Linkage Disequilibrium" (LD).

Check Your Progress

2) What are the evolutionary factors affecting the genetic variation in population genetics?

.....

.....

.....

.....

1.2 HARDY WEINBERG PRINCIPLE: A MATHEMATICAL MODEL OF POPULATION GENETICS

G.H. Hardy (a British Mathematician) and W. Weinberg (a German physician) independently proposed Hardy Weinberg Law which states that "the relative frequencies of various kinds of genes and alleles in a large and randomly mating (panmictic) population tend to remain constant (or in equilibrium) from one generation to another generation in the absence of evolutionary forces such as mutation, selection and gene flow, etc.". The principle can be represented in the form of equation and this is represented by:

$$p^2 + 2pq + q^2 = 1$$

Where, p and q are the allele frequencies for a genetic locus with two alleles. This equation can be used to measure whether the observed genotype frequencies in a population differ from the frequencies predicted by the equation. For more details about Hardy-Weinberg principles, assumptions and applications, you can refer to Unit 4.

1.3 PRINCIPLES OF MENDELIAN GENETICS

Gregor Johann Mendel, known as 'Father of Genetics', was an Austrian monk who worked on Pea plants (Pisum sativum) to understand the concept of heredity. His work laid the foundation of modern genetics. He was the first one to say that the characters are transmitted from one generation to the next through particles "factors" that are now known as genes. He proposed three basic laws of inheritance:

1) Law of Dominance

2) Law of Segregation

3) Law of Independent Assortment.

Check Your Progress

3) What are the three principles of Mendelian Genetics?

.....

.....

.....

.....

Reason for Choosing Garden Peas

- Short life span of three to four months with three generations in a year.

- Small herbaceous plant producing many seeds.
- Naturally self-pollinating but have seven contrasting characters (fig.1.1) such as tall, dwarf plant etc. which results in variation.
- Easy to self or cross pollinate the plant.

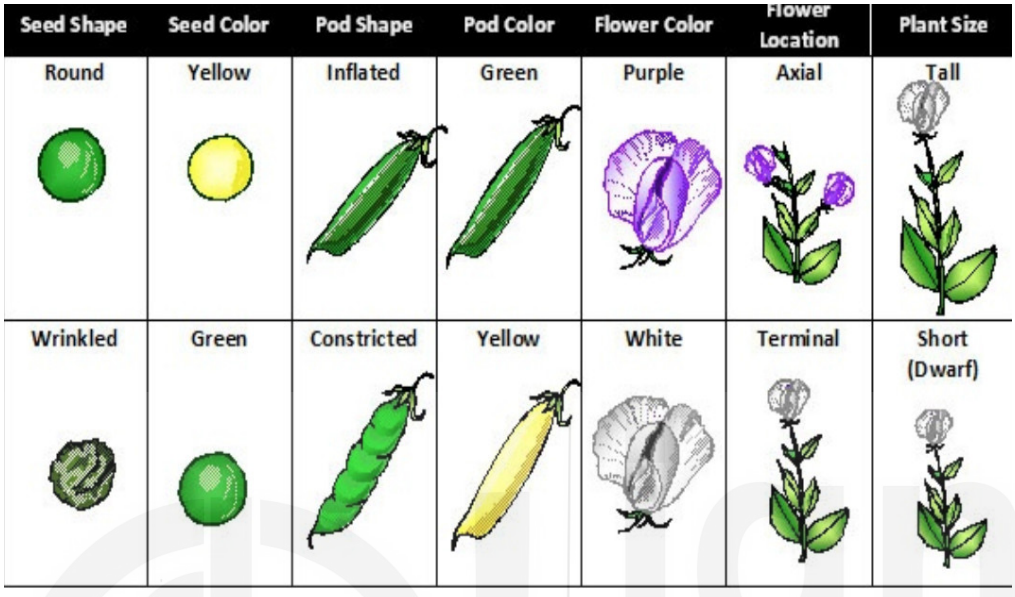


Fig. 1.1: Seven contrasting traits of Pea plants
(***Source:*** https://images.app.goo.gl/Vf4oMptP4Zh8dyWU8)

1.3.1 Mendel's Experiment

Mendel conducted 2 main experiments to determine the laws of inheritance. These experiments were:

1. Monohybrid Cross and
2. Dihybrid Cross

1.3.1.1 Monohybrid Cross

When only one character is considered while crossing two organisms, then such a cross is known as monohybrid cross. This monohybrid is used to observe how homozygous offspring express heterozygous genotypes inherited from their parents.

Monohybrid cross is responsible for the inheritance of one gene. In this experiment, Mendel took two pea plants of opposite traits (one short and one tall) and crossed them. He found the first generation offspring (F1 progeny) were all tall. Then he crossed F1 progeny and obtained a phenotypic ratio of 3:1 and Genetic ratio of 1:2:1 in F2 generation (Fig.1.2). He also conducted this experiment with other contrasting traits. In all the cases, he found a similar result. From this he formulated the law of dominance and the law of segregation.

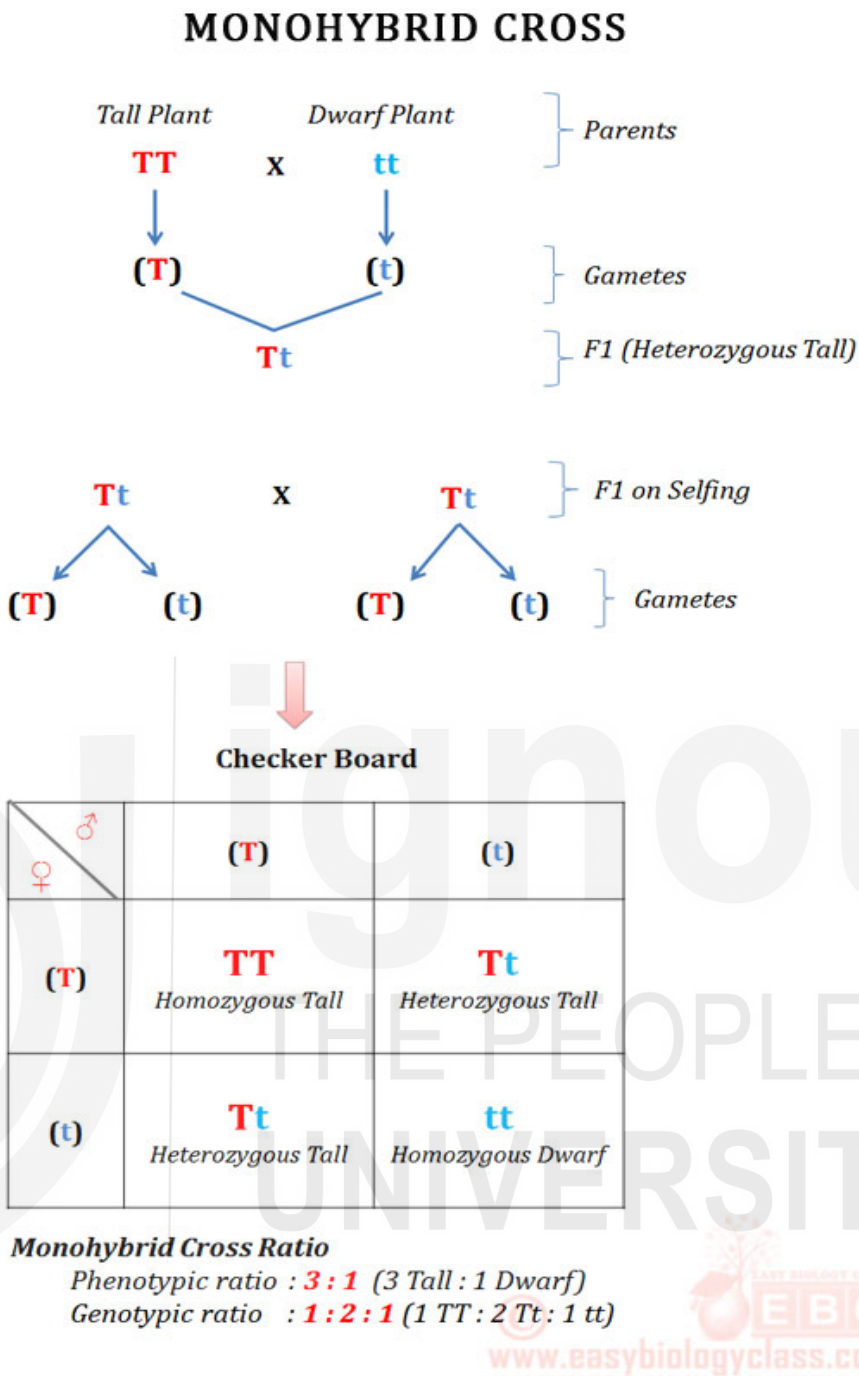


Fig. 1.2: Monohybrid Cross (Source: https://images.app.goo.gl/LBXKeBj6WNV1ENPz5)

Examples: Tongue rolling is inherited as a simple mendelian trait. 'R' is the allele for roller and 'r' is the allele for non-roller.

1.3.1.2 Dihybrid Cross

When two characters are considered while crossing two organisms, then such a cross is known as dihybrid cross.

In this experiment, Mendel considered two traits (colour and shape of seed), One parent carries homozygous dominant allele i.e., round-yellow seeds (YYRR), while the other one carries homozygous recessive allele i.e., wrinkle-green seeds (yyrr). He crossed them and observed that all offspring produced after the first generation progeny (F1 progeny) were all round-yellow (heterozygous

for specific traits). This meant that dominant traits were the round shape and yellow colour.

He then self-pollinated the F1 progeny and obtained 4 different traits round-yellow, round-green, wrinkle-yellow, and wrinkle-green seeds in phenotypic ratio of 9:3:3:1 in F2 generation (Fig. 1.3). From dihybrid cross, Mendel formulated law of independent assortment.

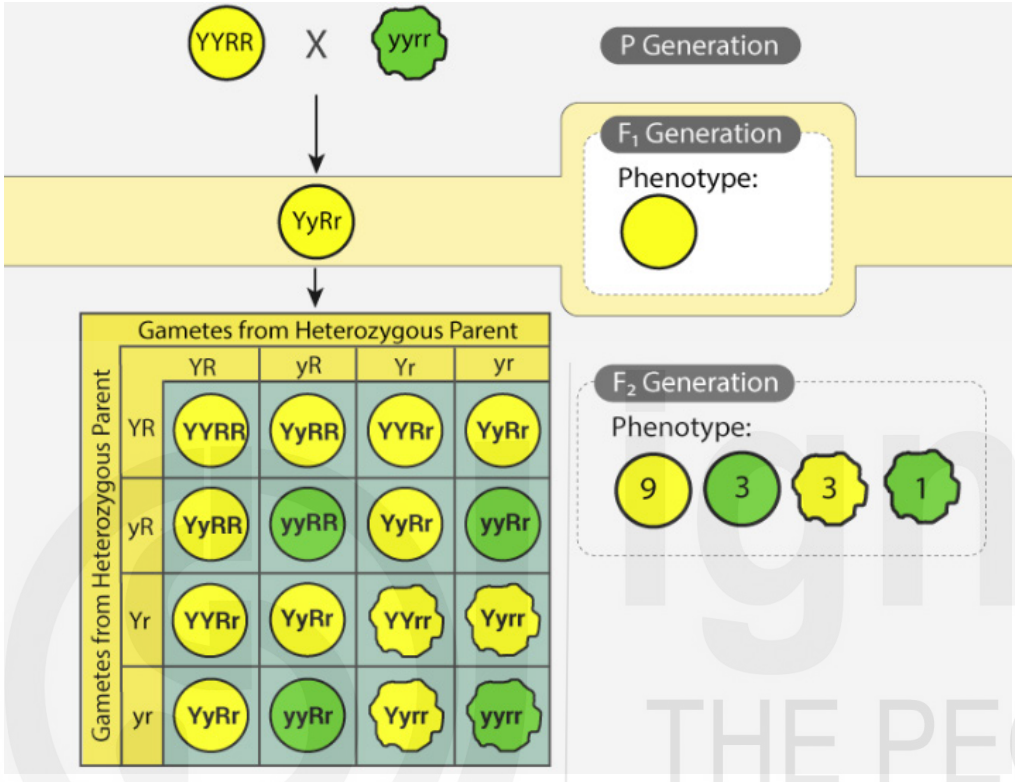


Fig. 1.3: Dihybrid Cross. (Source: https://images.app.goo.gl/BaZzYka3r9rQBvn1A)

Example: ABO and Rh blood group systems

1.3.2 Mendel's laws of Inheritance

1.3.2.1 Law of Dominance

According to the law of dominance, hybrid offspring will only inherit the dominant trait in the phenotype. In other words, out of pair of contrasting alleles, one dominates the other, the alleles which express/dominates are called as dominant trait and the alleles which are suppressed are called as recessive trait. This law is significant and true but not universally applicable.

1.3.2.2 Law of Segregation

Mendel's Law describes that member of allelic pair in a hybrid remain together without mixing with each other and separate or segregate during gamete formation and gametes receive only one of the two factors from each parent and are pure for given trait (law of purity of gametes). This law is universally applicable.

Much of our DNA exists on long strands in the nuclei of our cells, called chromosomes. Chromosomes come in pairs. Different species have different

numbers of chromosomes; humans have 23 pairs, whereas chimpanzees (our closest living relative) have 24 pairs. During the replication of body cells through mitosis, a single cell containing 23 pairs of chromosomes will duplicate, giving rise to two identical cells, each with 23 pairs of chromosomes. However, this is not what happens during reproduction. Instead of passing along 23 pairs of chromosomes to your offspring in a sex cell (sperm in males, egg in females), you pass on one of each pair through the process of meiosis. The process of chromosome pairs separating through meiosis is known as Mendel's law of segregation. You contribute 23 chromosomes (not 23 pairs) and your mate contributes 23 chromosomes, resulting in your child having 23+23=23 chromosome pairs. Likewise, your genetic inheritance also resulted from this process, as one of each chromosome pair came from your father and the other one came from your mother.

1.3.2.3 Law of Independent Assortment

Law of independent assortment states that the inheritance of one trait has no effect on the inheritance of another. It is caused by gene shuffling or random assortment of chromosomes during meiosis. In other words, as the individual heredity factors assort independently, different traits get equal opportunity to occur together. This law is not universally applicable.

As a bisexual organism (a species that has two distinct sexes, male and female), half of your genetic inheritance comes from your mother and half from your father. The same applies to any biological siblings. Apart from identical twins, why are you not identical to a sibling? If your brother and you received 50% of your DNA from your mother and 50% from your father, why are you not genetically the same? The answer relates to basic probability; you do not inherit the same 50%. For any given chromosome pair, there is 50:50 chance of one being passed on to an offspring, either the maternal chromosome (from your mother) or the paternal chromosome (from your father). The same probability applies to each chromosome pair, as they are all independent such that whatever chromosome you pass on from the first chromosome pair has no effect on the second pair, the third and so on.

Let us illustrate this principle with a simple analogy using coins. Imagine an organism with only three chromosome pairs, each represented by a penny with two sides- heads and tails. If we flip the first coin, we have 50:50 chances of getting heads (H) and tails (T). We will use this as a model for a chromosome pair consisting of one chromosome labelled H and one labelled T. If you flip heads for the first coin (chromosome pair), what is the probability of flipping heads on the second coin? It is still 50:50 because the coin flips are independent; the outcome of one coin flip does not influence any other coin flips. In terms of the genetic analogy, this hypothetical organism can produce eight different combinations of coin flips. One of these eight combinations would be getting heads for the first coin, heads for the second coin, and heads for the third coin. Another possibility would be heads for the first coin, heads for the second coin, and tails for the third coin. If we follow this pattern, we wind up with eight different combinations, each equally likely:

1) Heads-heads-heads

2) Heads-heads-tails

3) Heads-tails-heads

4) Heads-tails-tails

5) Tails-heads-heads

6) Tails-heads-tails

7) Tails-tails-heads

8) Tails-tails-tails

Because of chance, this organism could produce eight different combinations of chromosomes. This independent inheritance is known ad Mendel's law of independent assortment.

In principle, we could simulate the same process for human beings by using 23 different coins, but it will take too long to enumerate all possible combinations of coin flips. Instead, we can figure out the number of possibilities using the simple formula $2n$, where n is the number of coins/chromosomes pairs. For humans, n= 23 chromosome pairs, giving 223 = 8,388,608 combinations! Keep in mind that this is for one individual. The same rule applies to the production of sex cells in the individual's mate; they too, can produce up to 8,388,608 combinations. A child could therefore have any of the first parent's combinations paired with any of the second parent's combinations, giving a total of 8,388,608 × 8,388,608 = 70,368,744,177,664 possible genetic combinations in any given child. Given the number of possibilities, it is easy why it would be virtually impossible for you to be genetically identical to your non twin brother for your entire genome.

1.3.3 Test Cross

A genetic cross between a homozygous recessive individual and a corresponding suspected heterozygote to determine the genotype of the later is called test cross. Test cross is a simple method devised by Mendel to verify the genotype of F1 / F2 /F3 hybrids.

When hybrid expresses recessive trait e.g., dwarf plant, its genotype is definitely tt but if hybrid tall, it is not possible to know its genotype. Hence to determine whether the tall plant is homozygous or heterozygous, test cross is performed. Figure 1.4 shows how to determine the unknown genotype using the test cross. In test cross, F1 hybrid is crossed with the homozygous recessive parent. Since the offspring is crossed back with one of the parents, it is also called back cross.

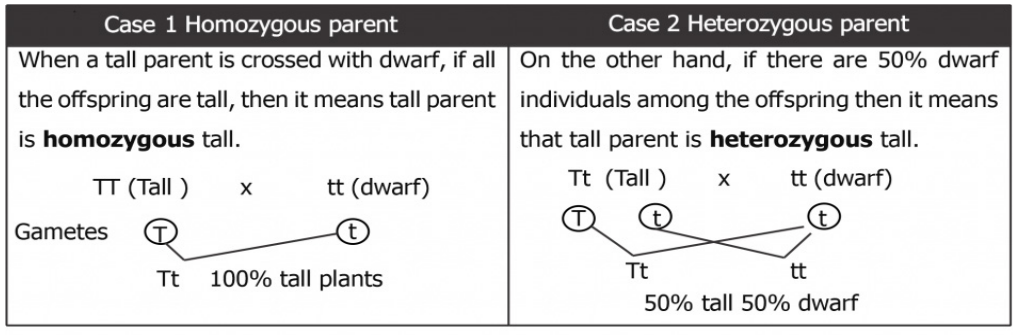


Fig. 1.4: Determination of unknown genotype using the Test Cross

1.4 SUMMARY

- Population genetics can be defined as the study of the frequency of genes and genotypes in a Mendelian population and predicting the way these frequencies would change under the combined influence of various factors, overtime in the population.
- The key evolutionary forces that shape the pattern of genetic variation are mutation, natural selection, gene flow, genetic drift, non-random mating and recombination.
- The Hardy-Weinberg equation can be used to measure whether the observed genotype frequencies in a population differ from the frequencies predicted by the equation.
- Gregor Johann Mendel was an Austrian monk who worked on pea plants (Pisum sativum) to understand the concept of heredity. His work laid the foundation of modern genetics. He proposed three basic laws of inheritance-law of dominance, law of segregation and law of independent assortment.
- Mendel formulated the law of dominance and law of segregation from monohybrid cross and he formulated the law of independent assortment from dihybrid cross.
- Test cross is used to determine the genotype of suspected heterozygote.

1.5 IMPORTANT TERMINOLOGY

Allele: An alternative form of a gene that is found at the same place on a chromosome.

Dominant alleles: The alleles that express themselves in an organism in every possible combination and can be seen.

Gene Pool: The sum total of genes of all the individuals of a Mendelian population is known as gene pool.

Genotype: The genetic makeup of an individual at a given locus defined by what has been inherited from both parents.

Heterozygous: An individual possessing dissimilar alleles for a particular trait is called heterozygous or hybrid for that trait.

Homozygous: An individual possessing similar alleles for a particular trait is called homozygous or pure for that trait.

Locus: The physical location of a gene or of a DNA polymorphism on a chromosome.

Phenotype: The external appearance of an individual for any trait is called phenotype for the trait.

Recessive alleles: The allele which is not expressed in the presence of a dominant allele.

Trait: Any characteristic that can be passed from parent to offspring.

1.6 REFERENCES AND WEB LINKS

Das, B. M. (2008). *Outlines of Physical Anthropology*. Delhi: Kitab Mahal.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. New York: Cambridge University Press.

Loewe, L. & Hill, W. G. (2010). The Population Genetics of Mutations: Good, Bad and Indifferent. *Philosophical Transactions of The Royal Society B Biological Sciences*, 365:1153-67.

Relethford, J.H. (2012). *Human Population Genetics.* (Vol. 7). New Jersey: John Wiley & Sons.

Lachance, J., & Tishkoff, S. A. (2013). Population Genomics of Human Adaptation. *Annual Review of Ecology, Evolution, and Systematics*, *44*:123-143.

National Geographic: encyclopaedia https://www.nationalgeographic.org/encyclopedia/natural-selection/ Retrieved on 15/08/2021

Mendel's laws of Inheritance https://byjus.com/biology/mendel-laws-of-inheritance/ Retrieved on 15/08/2021

Monohybrid cross https://images.app.goo.gl/LBXKeBj6WNV1ENPz5 Retrieved on 15/08/2021

ePG Pathshala, Anthropology, Paper no.08, Module 1 https://epgp.inflibnet.ac.in/epgpdata/uploads/epgp_content/anthropology/08._human_population_genetics_/01.definition_and_scope_of_population_genetics/et/6186_et_m1.pdf Retrieved on 13/08/2021

1.7 ANSWERS TO CHECK YOUR PROGRESS

1) Population genetics can be defined as the study of the frequency of genes and genotypes in a Mendelian population and predicting the way these frequencies change or remain constant under the combined influence of various factors, overtime, within and between populations.

2) The key evolutionary forces that shape the pattern of genetic variation are mutation, natural selection, gene flow, genetic drift, non-random mating and recombination.

3) The three principles of Mendelian genetics are:
   i) Law of dominance
   ii) Law of segregation
   iii) Law of independent assortment.

UNIT 2 HAEMOGLOBIN POLYMORPHISM AND THALASSAEMIA

Contents

Learning Objectives

After reading this unit, you will be able to:

- Elucidate about the blood and haemoglobin;
- Appreciate the types of hemoglobinopathies, their sub types, consequences; and
- Understand the clinically significant haemoglobin variants.

2.0 INTRODUCTION

Human blood contains 91%-92% of light yellowish or straw coloured liquid known as plasma and 8%-9% of solids consisting of red blood cells, white blood cells and platelets, fibrinogen, proteins (albumin, globulin, clotting factors, hormones, antibodies and enzymes), immunoglobulins and electrolytes (calcium, sodium, chloride, bicarbonate, potassium) (Matthew et al. 2021). The process by which different blood cell types are formed is called haematopoiesis. This process is initiated by haematopoietic stem cells residing in the bone

marrow. These haematopoietic stem cells divide into common myeloid or lymphoid progenitors (cells capable of differentiating into any cell type). Myeloid progenitor cells produce monocytes, erythrocytes or red blood cells, basophils, eosinophils, neutrophils and megakaryocytes. From megakaryocytes, platelets are formed. Lymphoid progenitor cells produce T cells, B cells, and natural killer cells (Chapman and Zhang, 2021).

Erythrocytes or red blood cells have biconcave, discoid shape. These cells are a nucleated means they don't have a nucleus. The life span of red blood cells (RBC) is 120 days. RBC transport oxygen from the lungs to the different tissues and collect carbon dioxide from tissues and transport to the lungs. RBC has two phospholipid layers in the membrane and this membrane is maintained by a network of proteins. The percentage of RBC present in a given volume of blood is called haematocrit. RBC play an importance role in health and disease. Low number of RBC is considered as anaemia and also observed in thalassemia, a genetic disorder. Different shapes of RBCs were reported in different diseases such as small RBCs in microcytic anaemia, sickle cell shape in sickle cell anaemia (genetic disorder) and non-discoid or atypical erythrocytes in systemic lupus erythematosus (autoimmune disorder) (Barbalato and Pillarisetty, 2021).

2.1 HAEMOGLOBIN

The transport of gases such as oxygen and carbon dioxide in RBC is carried out by haemoglobin (Hb) and it also gives red colour to the RBC. Each RBC contains 300 million molecules of Hb. Hb molecule is composed of two globin chains or clusters which are known as α (alpha) globin and β (beta) clusters. These two clusters are associated with heme moiety composed of organic protoporphyrin ring IX and iron ion in ferrous state. The α globin cluster contains ζ(zeta) and α globin chains and β cluster comprises of ε (epsilon), γ (gamma), β and δ (delta) chains. Six phenotypes of Hb are found in humans. They are: embryonic (Hb Gower I (ς2ε2), Gower II (α2ε2), and Portland (ς2γ2) (3-10 weeks of gestation); foetal haemoglobin (HbF; α2γ2) (pregnancy); HbA (α2β2) and HbA2 (α2δ2) (adult). The appearance of different species of Hb during different periods of human development is known as haemoglobin switching. The percent of A, A2 and foetal phenotypes of Hb in normal adult human beings are 97%, 2% and 1% respectively. The genes of α cluster (*HBZ, HBA2, HBA1, HBM and HBQ1*) are located on chromosome 16 and genes of β cluster (*HBE1, HBG2, HBG1, HBD* and *HBB*) are housed on chromosome 11. Nucleotide sequences flanking the genes of either cluster, perform the regulatory role. Globin chains are produced in the cytoplasm of RBC and heme in cytoplasm and mitochondria of RBC.

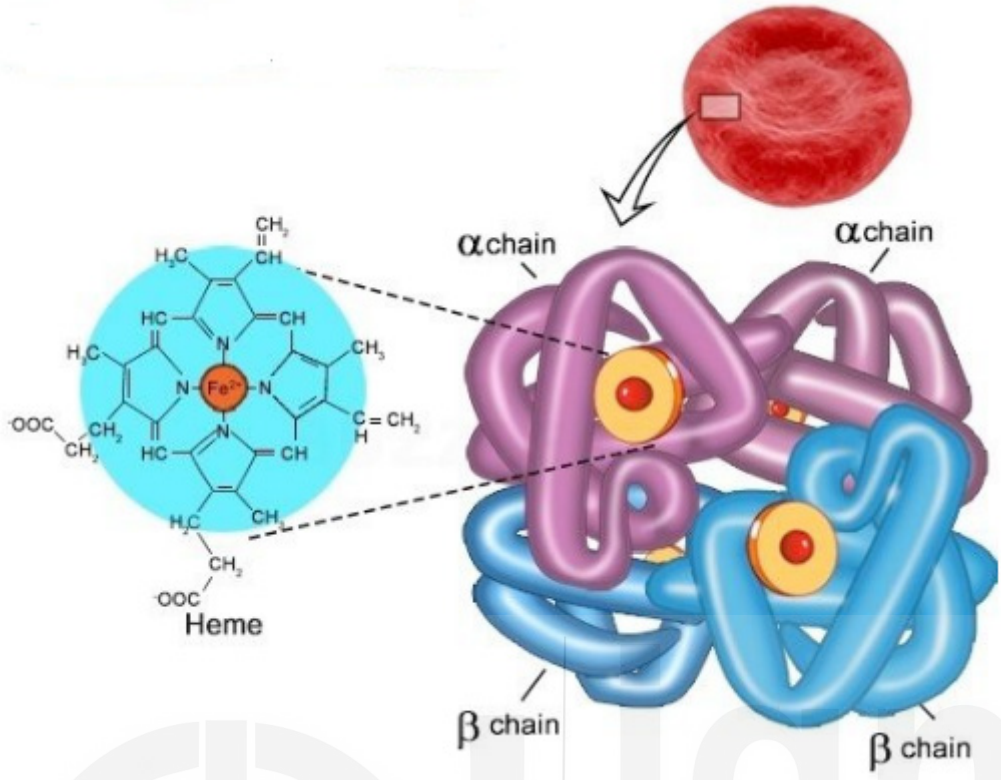


Fig: 2.1: Structure of Haemoglobin
(***Source:*** https://www.pinterest.com/pin/342836590378990703/)

2.2 HEMOGLOBINOPATHIES

These are groups of monogenic inherited disorders of red blood cells. They are classified into two groups namely Thalassemias and haemoglobin variants. Hemoglobinopathies are a result of the mutations affecting the alpha or beta globin production or structure and function of haemoglobin. The higher incidence of hemoglobinopathies were reported in the Middle East, the Mediterranean, southern China and south east Asia (Zhang et al. 2019). Hemoglobinopathies show autosomal recessive inheritance. Each year around 300,000-400,000 babies are born with a severe haemoglobin disorders and its frequency reaches to 90% in low and middle income countries (Williams and Weatherall, 2012; Christianson et al. 2006). The incidence of hemoglobinopathies in India is 1.2 per 1000 live births and it is estimated that each year 32,400 babies are born with a serious haemoglobin disorders (Christianson et al. 2006). The higher incidence of hemoglobinopathies was attributed to natural selection mediated resistance against *Plasmodium falciparum* malaria in the carriers, high frequency of consanguineous marriages, genetic drift and higher maternal age. Global distribution of the hemoglobinopathies is due to the migration and marriages between different ethnic groups (Thalassemia International Federation, 2021). The spread of hemoglobinopathies was found to be heterogeneous across countries and small geographical regions (Christianson et al. 2006).

Mutations of Hb genes were reported in 7% of global populations. Significant frequency of sickle cell gene (40%) and HbE (60%) in human populations has also

been observed (Heaton et al. 2021). At the time of writing this unit, information on1405 variants in haemoglobin gene and 536 variants in thalassemia was documented in HbVar database, a database of human haemoglobin variants and thalassemia (globin.bx.psu.edu/cgi-bin/hbvar/counter). Various methods such as isoelectric focussing, high performance liquid chromatography, mass spectrometry, cellulose acetate, citrate agar, alkaline globin chain and capillary zone, electrophoresis, restriction fragment polymorphism, allelic discrimination using real time PCR, reverse dot blot, the amplification-refractory mutation system and DNA sequencing are used to detect hemoglobinopathies.

Check Your Progress

1) What are the Hemoglobinopathies and its types?

.....

.....

.....

.....

2.2.1 Thalassemias

Thalassemia word is derived from Greek word 'thalassa' meaning sea. These are a group of Hb diseases in which reduced or no synthesis of one or more α or β globin chains or subunits is observed. Reduced production of RBC, low levels of Hb and anaemia is a characteristic feature of thalassemia. Imbalances of globin chains in thalassemias results in the hemolysis and impairment of erythropoiesis. Both sexes are equally affected. It is estimated that worldwide 56,000 conceptions would have a major thalassemia disorder (Cao and Kan, 2013) and each year in India, 10,000-12,000 children are born with thalassemia (Colah et al. 2017). Depending on the impairment of globin chain synthesis these are known as β, α, δ, γ, δβ or εγδβ thalassaemias (Thalassemia International Federation (TIF), 2021). Most of the transmission of thalassemia is autosomal recessive and rarely show autosomal dominant inheritance. Among these thalassemias, β and α thalassemias are important from the point of view of the severity of symptoms and requirement of RBC transfusion in homozygous condition. Heterozygotes in both α and β thalassemias show no symptoms and require no treatment. Homozygotes in both α and β thalassemias, compound heterozygotes and interaction of thalassemias and Hb variants produce various thalassemic syndromes. Depending on the severity of symptoms following thalassaemic syndrome people require lifelong (non-deletional HbH, β thalassemia major, severe HbE/ β thalassemia), intermittent (deletional HbH, moderate HbE/ β thalassemia), occasional (mild HbE/ β thalassemia, HbC/ β thalassemia) and seldom (α-thalassemia trait and β- thalassemia minor) RBC transfusions. Thalassemias are diagnosed by hematologic and molecular testing. Thalassemias are prevented by public education, screening of carriers and pregnant women and preventing of marriages between carriers.

2.2.1.1 The *α* Thalassemia

Deletion or deficient synthesis of alpha globin chains and excess accumulation of β chains observed in α Thalassemia. This is due to deletions, point mutations and insertions of one or more alpha globin cluster genes (CDC, 2015). The carrier rate of α thalassemia is 1%-80% in India (Panigrahi and Marwa, 2007). α Thalassemia were detected in South East Asia, China, Middle East, Gulf, Italy, Greece, North Europe (TIF, 2021), Africa, Asia, the Mediterranean basis (Dominguez et al. 2013). Carrier rate of α thalassemia in China (40%) (TIF, 2021) and Afro-descent population of Uruguay (15.4%) (Dominguez et al. 2013) were reported. The commonly found α thalassemia mutations are -α3.7, - α4.2, - α SA (South African deletion) and Hb Constant Spring (Panigrahi and Marwa, 2007).

2.2.1.1.1 Subtypes of α Thalassemia

There are 7 subtypes of α Thalassemia which have been mentioned below:

Alpha Thalassemia or Silent Carrier (α+ -thalassemia) (one gene deletion in alpha globin cluster).

This subtype of thalassemia is found in sub-Saharan Africa, Mediterranean regions, Middle East, Indian subcontinent, East and South East Asia. Frequency ranges from 10-25% in the aforementioned geographical areas and in North India and Papua New Guinea it is prevalent in 80% of people (Williams and Weatherall, 2012).

Alpha Thalassemia Minor or Trait (α0 -thalassemia) (two genes deleted in alpha globin cluster on same chromosome).

This is of two genotypes i.e. *cis* (if occur on the same chromosome, common in Asian populations) or *trans* (if occur on opposite chromosome, common in African/Black populations). Mild anaemia is observed in the carriers of this type of thalassemia.

Haemoglobin H (HbH) or Thalassemia Intermedia (three genes deleted or decreased expression). In this type most of the patients show mild to moderate anaemia. Severity depends on the type of mutation involved (deletional or non-deletional either alone or in combination). In severe cases, growth retardation occurs. Other symptoms include enlargement of liver and spleen, feeding problems, leg ulcers, gallstones and folic acid deficiency.

Hydrops Fetalis with Bart's or α-Thalassemia Major (four genes deleted in alpha globin cluster or decreased expression of four genes in alpha globin cluster).

In this type of thalassemia, fluid accumulation is observed in the tissues, organs and skull (hydrocephaly) of foetus. Anaemia, enlarged liver and spleen, impaired brain development, signs of heart failure, abnormalities of skeleton and urinary tract also found. It can be fatal condition before (prenatal as still birth) and after birth (neonatal period).

Hb Constant Spring or Non-Deletion Thalassemia

It is caused by Synthesis of elongated α-globin subunit due to chain termination mutation. The carriers of this phenotype accumulate very low levels of α-globin subunit in their RBC. This phenotype is common in the residents of South East Asia. This Phenotype leads to severe form of HbH disease when coinherited with α0 cis genotype on the different chromosome.

The α-thalassemia with Mental Retardation Syndrome (ATR)

This Syndrome is of two types. They are the ATR-16 syndrome (involve very large deletions in α-globin and adjacent genes on chromosome 16 that thought be arise denovo in embryo stage of development) and ATR-X syndrome (contain normal α-globin gene and are caused by germline mutation in ATRX transcription factor gene encoding the transcription factor involved in the regulation of α-globin gene expression.

Acquired α -thalassemia (HbH disease) Associated with Myelodysplastic Syndromes (ATMDS)

MDS is a bone marrow disorder with characteristic features of disturbed haematopoiesis with cytopenia, abnormal myeloid differentiation and RBC abnormalities. The α-globin gene is structurally normal in this type of thalassemia but the impairment in its gene expression is observed due to the somatic mutation in the ATR-X gene and comparatively severe impairment of α-globin gene was observed than ATR-X syndrome (Forget and Bunn, 2013).

2.2.1.1.2 Consequences of α-Thalassemia

Alpha thalassemia minima or silent carrier has no haematological consequences. Mild microcytic (small RBC <83 μm) in or hypochromic anaemia (RBC pallor or less colour) without haemolysis occurs in both (*cis* and *trans*) genotypes of α+-thalassemia. In Haemoglobin H, compensated anaemia is seen requiring no RBC transfusions. β globin self associate and form β globin chain tetramers (HbH). These are unstable species damage the aged RBC and reduce their life span by forming inclusion bodies. Hb Bart's is composed of four γ globin subunits. Hb Bart's has high affinity to oxygen and don't release the oxygen to the body tissues as a result of which infant carriers (their RBC lack adult or foetal Hb) of Hb Bart's phenotype suffer from hypoxia (lack of oxygen) condition known as hydrops fetalis (Forget and Bunn, 2013). HbH causes microcytic anaemia, haemolysis and splenomegaly.

Check Your Progress?

2) What are the characteristic features of thalassemias and its causes?

.....

.....

.....

.....

2.2.1.2 The β-Thalassemias

These thalassemias are characterised by absent or reduced synthesis of beta globin chains leading to the excess accumulation of α globin chains. Reduced RBC synthesis, decreased Hb in RBC and anemia is observed in patients with β-Thalassemias. β-Thalassemias are caused by 300 mutations and deletion of two beta globin genes. In India, 80 mutations in β-Thalassemias were observed, the most common mutation is IVS 1-5 (G>C). Along with this mutation, other mutations such as IVS1-1 G→T, Codon 41/42(-TCTT), Codon 8/9 and the 619 bp deletion are responsible for 90% of mutations in β-Thalassemias in India (Panigrahi and Marwa, 2007). The mutations are point mutations involving a one or more nucleotides in the β globin gene. Others include deletions caused by unequal crossover between the linked and partially homologous δ - and β -globin genes in δβ globin or Lepore gene and large deletions involving partial or total β globin clusters in δβ or εγδβ thalassemias or hereditary persistence of foetal haemoglobin (HPFH) syndromes. It is estimated that in India 30,000 people would have β-Thalassemias (Colah et al. 2017). Incidence of β-Thalassemias range from 4%-17% in different ethnic groups and incidence of β-Thalassemia carrier status was found in 3-4% or 35-45 million in India (colah et al., 2017). The carrier rate of β thalassemia is 14% in Cyprus and 10.3% in Sardina. It is estimated that 1.5% of global population or 80-90 million are β-thalassemia carriers; 1 in 100,000 in world population is diagnosed with beta thalassemia and 200,000 patients with β-thalassemia alive globally and the most common compound β-thalassemia syndrome is HbE/ β-thalassemia found in 50% of residents of South East Asia (Galanello and Origa, 2010).

2.2.1.2.1 Subtypes β-Thalassemias

There are 4 subtypes of β-Thalassemia which have been mentioned below:

β-thalassemia Minor or Carrier or Trait or Heterozygous beta Thalassemia: In this type of thalassemia, a mutation in one Hb beta globin gene is observed.

β-thalassemia Major or Cooley Anaemia or Mediterranean Anaemia: This is caused by a mutation in two Hb beta globin genes resulting in almost complete lack of beta chains.

β-thalassemia Intermedia: This is due to a mutation in two Hb beta globin genes resulting in reduced levels of beta chains.

Dominant β-thalassemia: This is caused by a mutation in Hb beta gene and produce unstable Hb**.**

2.2.1.2.2 Consequences of β Thalassemias

In all types of β thalassemias commonly observed features are deficiency of HbA tetramers and accumulation of free α-sub units. These α-sub units (due to the deficiency of β subunits) fail to form Hb tetramers. Mild to moderate hypochromic, microcytic anaemia without haemolysis in heterozygotes (β thalassemia trait or β -thalassemia minor) and transfusion dependent haemolytic anaemia associated with weakness, fatigue, dizziness, shortness of breath, headaches and jaundice; recurrent fever, irritability, diarrhoea,

feeding problems, enlargement of liver and spleen and bone deformities, (β –thalassemia major) in homozygotes or compound heterozygotes major are cardinal clinical features. Symptoms in β –thalassemia major appear after 6 months of age. Haemolytic anaemia, poor growth and skeletal abnormalities are observed in patients with β –thalassemia major. Patients with β –thalassemia major die after 30 years of age due to cardiac complications of iron overload. In β –thalassemia intermedia, mild to moderate and partially compensated haemolytic anaemia requiring occasional transfusions, pallor, moderate to severe skeletal abnormalities, jaundice, gall stones, pallor, enlargement of liver and spleen are the characteristic features. In this type of thalassemia less severe imbalance of α and non- α globin is observed leading to the less accumulation of free α subunits. This is due to the inheritance of milder β+-thalassemias with less severity or co-inheritance of α like thalassemia or coinheritance of genetic trait associated with increased production of the γ subunit of foetal haemoglobin or heterozygosity for a single mutant β-globin gene associated with production of highly unstable β-globin subunit causing RBC damage (dominant β-thalassemia). In δβ-thalassemia, β-globin sub unit is not produced due to the persistent high level of expression of γ subunit of foetal haemoglobin that reduces the excess of α subunit. Haemolytic anaemia in neonatal stage was observed in δβ-thalassemia. Persistence high levels of γ globin production takes place in HPFH. Mild to moderate anaemia, jaundice and enlarged spleen is observed in patients with dominant β-thalassemia.

β-Thalassemia major patients are clinically managed by regular and safe blood transfusions and adequate iron chelation and removal of spleen and cured by allogenic stem cell transplantation. β-Thalassemias are prevented by public education, screening of carriers and pregnant women for prenatal diagnosis. β-Thalassemia intermedia is clinically managed by folic acid supplementation, spleen removal, prevention of thromboembolic events and treatment of leg ulcers and extramedullary erythropoietic masses.

2.3 HAEMOGLOBIN (Hb) VARIANTS

These variants are caused by missense mutations, deletions, antitermination mutations, multiple amino acid substitutions and post translational modifications. Some of the Hb variants produce clinically significant symptoms. There are more than 1000 naturally occurring Hb variants. These are named after the geographic origin of the affected individuals. Hb variants are identified by physical examination, laboratory testing (physical methods of Hb separation, Hb-O2 binding curves, visible wave length spectroscopy, Hb stability testing), electrophoresis, chromatography, mass spectrometry, DNA sequencing and crystallography techniques. In this unit Hb variants HbS, HbE, HbC and HbD are discussed.

2.3.1 HbS (Sickle Cell Haemoglobin)

This is also called sickle cell disease. It is a monogenic disease and shows autosomal recessive inheritance. It is caused by single base A>T mutation in

the triplet encoding the sixth residue of β globin subunit causing replacement of valine for glutamic acid and HbS. Sickle shaped RBC was first reported by Robert Herrick from Grenada in 1910. If a person inherits HbS from a single parent he/she is called heterozygote (HbAS) or a carrier of sickle cell trait. The incidence of sickle trait in African Americans is 80% and from 0.3%- 35% in different communities and 1%-40% in tribes of India. If a person inherits HbS from two parents, he/she is termed homozygote (HbSS) or carrier of sickle cell disease. There are four haplotypes in HbS mutations and they are Benin, Bantu, Senegal (Central Africa) and Arab-Indian/Asian. Studies on Indian sickle cell disease (HbSS) showed higher levels of foetal Hb, enlargement of spleen, lower reticulocyte counts, higher total Hb, >3 transfusions/year in under five years children, bacterial infections caused by staphylococcus aureus and gram negative bacteria. Vaso-occlusion in HbSS disease damages kidney, retina, spleen and liver. Symptoms are precipitated by hypoxia, hypertonic solutions, fever, infection, temperature changes and excess exercise. In 11% of SCD patients stroke occurs due to stenotic cranial arterial lesions.

Sickle cell disease was first identified by Lehman and Cutbush in Nilgiri Hills of Tamil Nadu in 1952. RBC of heterozygote person is composed of 40% of sickle (S) haemoglobin and 56-58% of adult haemoglobin (A) that confer protection to them against *Plasmodium falciparum* malaria and this condition is cited as good example of balanced polymorphism in population genetics. Heterozygotes show no symptoms. Sickling of Hb is due to the formation of gel and altering the shape of RBC from biconcave disc to elongated crescent as a result of polymerisation of deoxy HbS. Sickling causes abnormalities in permeability and cellular dehydration of RBC resulting in the premature destruction and haemolytic anaemia.

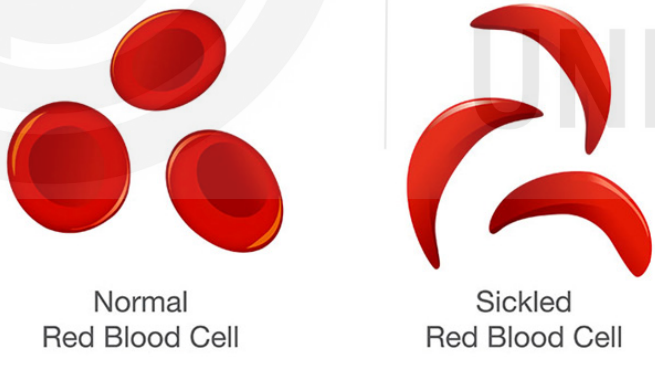


Fig. 2.2: Normal and Sickled Red blood cell

(***Source:*** https://www.cdc.gov/ncbddd/sicklecell/features/images/normal-and-sickled-red-blood-cell.jpg?_=76443)

Sickling RBC increase blood viscosity and obstruct capillary flow leading to the hypoxia, cell death, tissue necrosis and these group of events are known as vaso-occlusive pain crisis. HbSS or HbAS status is diagnosed by solubility test, electrophoresis, high performance liquid chromatography (HPLC), isoelectric focussing and capillary electrophoresis techniques. In India, newborn screening for sickle cell disorders was initiated after 2010 in few states collecting heel prick/cord blood on Guthrie cards followed by HPLC. Prenatal diagnosis is the

only choice for couple at risk of having sickle cell disease which is done by collecting chorionic villus sample at 10 to 12 weeks of gestation followed by extraction of DNA and analysis by reverse dot blot technique or allele specific priming method (Forget and Bunn,2013; Serjeant et al. 2016; Colah et al., 2015). Clinically HbSS is managed by avoiding pain episodes, treatment of infections, spleen removal and reliving symptoms and planning cure by stem cell transplantation.

2.3.2 HbE (Haemoglobin E)

This is a common Hb variant in South East Asia and its incidence is about 50% in some groups of Thailand. This is formed due to the single point mutation at codon 26 of beta globin gene and replaces amino acid glutamic acid with lysine. As a result of which abnormal RNA processing occurs and produces irregular RBC. Homozygotes (HbEE) are characterised by microcytic hypochromic anaemia while heterozygotes (HbAE) are identified by mild anaemia. Carriers of HbE variant develop jaundice, enlargement of liver and spleen and growth retardation. Severity of symptoms depends on its coinheritance with β-thalassemia which is prevalent in South East Asia. HbE is suggested to have multiple origins in Asia. Consanguineous marriages for variations in the frequency of HbE and meiotic recombination for appearance of more than one haplotype in a given population were postulated. A positive association between HbE and reduced malaria in epidemiological studies suggested the role of HbE conferring protection against malaria. Significant incidence of HbE was observed in the populations of Assam (HbAE: 4.6%-55.5%; HbEE: 1%- 45.4%). Arunachal Pradesh (HbAE: 9.1%-35.2%; HbEE: 0%- 5.3%), Meghalaya (HbAE: 4.5%-38.8%; HbEE: 0%- 4.29%), Manipur (HbAE: 5.6%-18.3%; HbEE: 0%-6.7%) Nagaland (HbAE: 0%-6.8%; HbEE: 0%- 0%) and Tripura (HbAE: 0%-53.9%; HbEE: 5.4%- 31.2%) of India (Ha et al. 2019; Sikdar, 2016).

2.3.3 HbC (Haemoglobin C)

This Hb variant resulted due to the substitution of amino acid lysine in place of glutamic acid at position 6 of the β globin. Heterozygotes (HbAC) are asymptomatic and homozygotes (HbCC) show mild haemolytic anaemia. When coinherited with HbS, they cause chronic haemolytic anaemia and show moderate haemolytic anaemia with splenomegaly when coinherited with β thalassemia. It has been shown that in homozygous state (HbCC) this Hb variant provide protection against *Plasmodium falciparum* and with the same malaria it provides intermediate protection in heterozygous state (HbAC). HbC is detected in West Africa (15%), and in South East Asia, Brazil, United States and Europe suggesting the presence of immigrants from West Africa (Piel et al. 2013)

2.3.4 HbD (Haemoglobin D)

This Hb variant is also known as haemoglobin D-Punjab or haemoglobin D-Los Angeles. It formed due to the point mutation in the beta globin gene in the first base of 121 codon replace glutamine for glutamic acid. It is found in Punjab

and North Western India (2%) and Gujarat. Hb D-Punjab is also detected in Italy, Austria, Belgium (0.4%), Turkey (0.2%) and China (55.6%). HbD exhibit normal solubility and show similar mobility as HbS in alkaline pH and mobility as HbA in acidic pH. Analysis of beta gene polymorphic restriction sites suggested HbD-Punjab originated in multiple geographical regions. This Hb variant was described by Itano in 1951. Hb with similar pattern was reported by Baglioni in 1962 known as Hb D-Chicago, Hb D-North Carolina, Hb D-Punjab, HbD-Portugal and Hb D-Oak showing similar chemical composition as that of Hb D-Los Angeles. This Hb variant is frequently referred as either Hb D-Punjab or Hb D-Los Angeles. Seven other types of HbD caused by different point mutations in Hb beta globin gene were reported and they are HbDAgri, HbD-Bushman, HbD-Ouled Rabah, HbD-Iran, HbD-Granada, HbD-Ibadan and HbD-Neath.

Homozygous (HbDD) state shows mild to moderate haemolytic anaemia and heterozygous (HbAD) state shows no clinical significant symptoms. It is also coinherited with thalassemias or HbS. Co-inheritance of HbD with β-thalassemia show mild microcytic and hypochromic anaemia. In compound form with HbS, HbD show symptoms similar to homozygous HbS (higher polymerisation rate with HbS, vulnerability of RBC for lysis, vaso-occlusion related pain, stroke and acute chest syndromes). HbD-Punjab variant is identified by cellulose acetate/ agarose/capillary electrophoresis, isoelectric focussing, HPLC and Restriction Fragment Length Polymorphism methods (Torres et al. 2015).

2.4 SUMMARY

Red blood cells are one type of blood cells. They contain a protein known as haemoglobin which carry oxygen from lungs to the tissues of the body and transport carbon dioxide from tissues to the lungs**.** Hb molecule is composed of two globin chains or clusters which are known as α (alpha) globin and β (beta) clusters. These two clusters are associated with heme moiety composed of organic protoporphyrin ring IX and iron ion in ferrous state. Hemoglobinopathies are a group of monogenic inherited disorders of red blood cells. They are classified into two groups namely thalassemias and haemoglobin variants. Hemoglobinopathies result from the mutations affecting the alpha or beta globin production or structure and function of haemoglobin. The higher incidence of hemoglobinopathies was attributed to natural selection mediated resistance against *Plasmodium falciparum* malaria in the carriers, high frequency of consanguineous marriages, genetic drift and higher maternal age. Global distribution of the hemoglobinopathies is due to the migration and marriages between different ethnic groups (Thalassemia International Federation, 2021). The spread of hemoglobinopathies was found to be heterogeneous across countries and small geographical regions.

Thalassemias word is derived from Greek word ‘thalassa’ meaning sea. These are a group of Hb diseases in which reduced or no synthesis of one or more α or β globin chains or subunits are observed. They are of two types namely α and β thalassemias. Within α-thalassemia seven subtypes are observed and in β

thalassemia, 4 subtypes. There are more than 1000 haemoglobin variants reported which are formed by missense mutations, deletions, antitermination mutations, multiple amino acid substitutions and post translational modifications. In this module Hb variants such as HbS, HbE, HbC and HbD are discussed.

2.5 REFERENCES

Alpha thalassemia. National organization of rare disorders. Retrieved from https://rarediseases.org/rare-diseases/alpha-thalassemia/

Barbalato, L & Pillarisetty LS. (2021) Histology, Red Blood Cell. In: *Stat Pearls*. Treasure Island (FL): StatPearls Publishing; Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539702/

Beta thalassemia. National organization of rare disorders. https://rarediseases.org/rare-diseases/thalassemia-major/

Cao A & Kan YW. (2013). The prevention of thalassemia. Cold Spring Harb Perspect Med 3. a011775

Colah R, Italia & Gorakshakar A. (2017). Burden of thalassemia in India: the road map for control. *Pediatric Hematology Oncology Journal*; 2:79-84.

Colah RB, Mukherjee MB, Martin S, & Ghosh K. (2015). Sickle cell disease in tribal population in India. *Indian J Med Res*;141: 509-515.

Chapman J & Zhang Y. (2015). Histology, Hematopoiesis. [Updated 2021 May 10]. In: Stat Pearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK534246/

Christianson A, Howson CP, Modell B. (2006). March of Dimes global report on birth defects. March of Dimes Birth Defects Foundation, New York.

Dominiguez Y, Zurita C, Calvopina D, Villacis J & Mora M. (2013). Prevalence of common haemoglobin variants in an afro-descent Ecuadorian population. *BMC Research Note*;6:132.

Forget BG & Bunn HF. (2013). Classification of the disorders of haemoglobin. In: Cold Spring Harbor Perspectives in Medicine. Eds: Weatherall D, Schechter AN, Nathan DG. Cold Spring Harbor Laboratory Press.

Galanello R & Origa R. (2010). Beta-thalassemia. *Orphanet journal of Rare Diseases*, 5:11.

Ha J, Martinson R, Iwamoto SK, & Nishi A. (2019). Hemoglobin E, malaria and natural selection.Evolution, *Medicine and Public Health*, 232-241.

Hemoglobinopathies (2015). current practices for screening, confirmation and follow-up. Centre for disease control and prevention.

HbVar: A Database of Human Hemoglobin Variants and Thalassemias. Available at https://globin.bx.psu.edu/cgi-bin/hbvar/counter.

Heaton C, Witt M, Cole L, Eyre J, Tazzyman S, McColm R & Francese S. (2021). Detection and mapping of haemoglobin variants in blood finger marks by MALDI MS for suspect profiling. *Analyst* , 146:4290.

Mathew J, Sankar P, Varacallo M. Physiology, Blood Plasma. [Updated 2021 Apr 28]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK531504/

Panigrahi I & Marwaha RK. (2007). Mutational spectrum of thalassemias in India. *Indian Journal of Human Genetics,* 13(1):36-37.

Piel FB, Howves RE, Patil AP, Nyangiri OA, Gething PW, Bhatt S, Williams TN, Weatherall DJ & Hay SI. (2013). The distribution of haemoglobin C and its prevalence in newborns in Africa. *Scientific Reports* 3:1671.

Serjeant GR, Ghosh K & Patel. (2016). Sickle cell disease in India: a perspective. *Indian J Med Res*, 143:21-24.

Sikdar M. (2016). Hemoglobin E in Northeast India: A review of its origin, distribution, migration and health. *Anthropological Review*; 79: 241-263.

Torres LDS, Okumura JV, da Silva DGH & Bonini-Domingos CR. (2015). Hemoglobin D-Punjab: origin, distribution and laboratory diagnosis. *Brazilian Journal of Hematology and Hemotherapy;* 37(2):120-126.

Cappellini, MD, Farmakis D, Porter J, Taher A. Thalassemia International Federation, (2021). Guidelines for the management of transfusion dependent Thalassemia.4th edition.

Williams TN & Weatherall DJ. (2012). World distribution, population genetics and health burden of the hemoglobinopathies. Cold Spring Harb Prospects Med 2. a011692.

Zhang J, Li P, Yang Y, Yan Y, Zeng X, Li D, Hong Chen1, Jie su1 & Baosheng Zhu. (2019). Molecular epidemiology, pathogenicity, and structural analysis of haemoglobin variants in the Yunnan province population of southwestern China. *Scientific Repo Rts*, 9:8264 https://doi.org/10.1038/s41598-019-44793-0.

2.6 ANSWERS TO CHECK YOUR PROGRESS

1) Hemoglobinopathies are a group of monogenic inherited disorders of red blood cells. They are classified into two groups namely thalassemias and haemoglobin variants.

2) The characteristic features of thalassemias are deletion or deficient synthesis of alpha globin chains and excess accumulation of β chains observed in α thalassemias is due to deletions, point mutations and insertions of one or more alpha globin cluster genes.

UNIT 3 GENETICS OF COMPLEX DISEASES

Contents

Learning Objectives

After reading this unit, you will be able to:

- Discuss basic principles of genetics and inheritance in humans;
- Describe the heritability of complex diseases;
- Elucidate the prevalence of complex diseases worldwide and in India; and
- Explain the details about important and most prevalent complex diseases in humans.

Contributed by **Dr. Sudipta Ghosh**, Department of Anthropology, North-Eastern Hill University, Shillong, Meghalaya

3.0 INTRODUCTION

In the last few decades, prevalence of complex diseases across the world and in almost every human population has gained tremendous attention among the biologists, anthropologists and health scientists. Initially, complex diseases were thought to be a manifestation of wealthy and sedentary lifestyle. Hence it was considered as the public health concern of the developed countries only. However, within no time these diseases started reaching every corner of the earth including developing countries like India. Scientists then started becoming aware about the complexities and multifactorial components of these diseases. They realised complex diseases are not simply life style diseases. They are in fact, the result of a complicated interaction between genetic and environmental components, where each of these components again has several other components controlling their expression. Here, we are going to learn details about various complex diseases emphasising on the genetic and environmental factors responsible for their prevalence.

3.1 WHAT IS GENETICS?

Genetics is a term that was derived from the Greek word '*Gen*', which means to become or to grow into something. Thus, it signifies that genetics deals not only with the transmission of hereditary factors or materials but also with the ways in which hereditary characters express themselves during the growth and development of the individual. In some families, one child may carry features of both parents, whereas another one resembles only the father or the mother. Even there can be a child who doesn't show any physical resemblance to any family member. All living organisms must carry specific genetic/hereditary information which can be transmitted to the next generation.

Hence, the science of 'Genetics' includes a study of the nature of the hereditary material, how it is transmitted from one generation to next and how it interacts with the environment to bring about an effect on a cell in particular and on an organism in general. Knowledge of genetics helps us to explain why an individual may resemble both the parents, one parent or neither of them. In your own family consider how each member differs from the others and yet how remarkably they resemble each other in some of their features. As a consequence, such similarities and dissimilarities create diversity in a family. Diversity or variation is the important factor for the survival of an organism at the micro level and ultimately the species at the macro level.

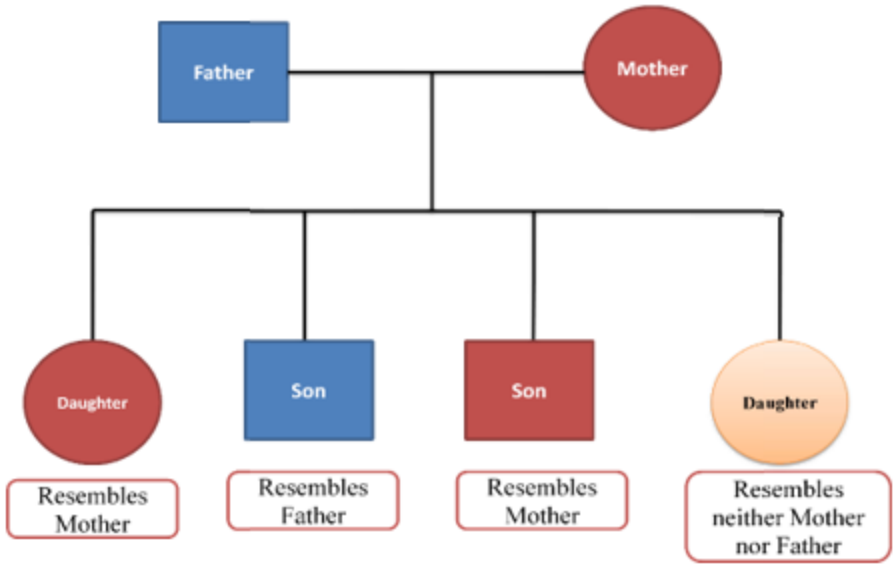


Fig. 3.1: Pedigree showing inheritance of physical features

3.2 PRINCIPLES IN HUMAN GENETICS

Most of the information on patterns of inheritance and the action of the hereditary material have been gathered from organisms that are not human. It is true that for several decades, certain simple patterns of inheritance in humans were well known. For example, traits like 'the presence of extra fingers (Polydactyly)' and 'color blindness' occur in a characteristic fashion in human pedigrees, but the basis for the transmission was not well understood for several years. As a solution, pioneer investigators used both plants and lower animals in their genetic studies. In fact, using bacteria and viruses as genetic tools have provided an insight into the very nature of the hereditary material and the way in which it acts in the cell. The basic principles established in these lower groups of organisms have been found to successfully apply to humans.

Now, the question is if it is the humans that we are interested in knowing about, why is it that biologists have not concentrated solely on humans to obtain fundamental genetic information? Well, there are several reasons behind this, but most important one is because humans are not the ideal creature for such research. In genetic research it is very important to follow not just one or two but several generations in order to determine the way in which a particular trait occurs among members of a family line. The span of a human generation is long, around 25 years. As a result, not many generations of one family can be followed in the lifetime of an investigator.

3.2.1 Heritability and Inheritance

When we study growth and development of humans, we are interested in the variation of metric characters in a population. One major and crucial component of this variation is attributable to genetic causes. When we try to evaluate the normality of a child or when we try to make a prediction of his/her height, we must take these genetic factors into account. This genetic component of any physical character of humans is called 'heritability'. The

mechanism of the transmission of hereditary material from parent generation to the offspring generation is known as 'inheritance'. In other words, inheritance is the transmission of particular characteristics from generation to generation by means of the genetic code, which is transferred to offspring in the gametes.

The physical characters that we study in growth and development are the combination of genetic and environmental components. Using mathematical formulation, scientists can actually calculate the degree of the genetic and the environmental components of any physical character that can be found in humans. Thus, such computation can answer the question like what is the percentage of a character being genetic and what is the percentage of a character is environmental. We can however find many situations where the genetic factors are not independent of the environmental factors. For example, an intelligent child will be more stimulated by intellectual influences of families and/or of schools.

Generally, the effects of genetic and environmental factors are additive and are randomly associated. Meaning, the genetic component values do not influence the environmental values. The environmental factors include all non-genetic factors like nutritional, climatic, cultural and so on. Heritability determines the degree of resemblance between relatives and is therefore used to estimate the heritability of any character among a particular population between parents and their offspring. However, even for traits with good additivity, i.e. genetic component, such as finger ridge count or even stature, environmental influences are not negligible. These traits do not develop outside the environment and we cannot conclude that physical differences are exclusively genetically determined. Heritability depends on gene frequencies and on the interactions between gene and environment (Susanne, 1982). Heritability is a quantitative index of the importance of genetics for individual differences in a physical character. Heritability (h^2) is a proportion and it ranges from 0 to 1.0. When h^2 is equal to 0, it implies only environmental influences, but when $h^2 > 0$, it implies that genetic influences are acting in some way.

Box 3.1 : Heritability

The heritability, as designated by h^2, of character expresses the proportion of the total phenotypic variation which can be attributed to the average effects of the genes controlling the character. It is the ratio of Genetic variance (VG) and total variance (VP) of a physical character.

That is, $h^2 = VG / VP$

It is also known as degree of genetic determination. It, therefore, indicates the relative importance of the genetic and environmental sources of variation in the character (Narain, 1988).

3.3 WHAT IS A COMPLEX DISEASE?

According to the Oxford dictionary, disease is a condition in which the normal function of some part of the body (cells, tissues or organs) is disturbed. A variety of microorganisms and environmental agents are capable of causing a disease. The functional disturbances are often accompanied by structural changes in tissue.

Complex or multifactorial diseases are defined as diseases that are ultimately determined by a number of genetic and environmental factors (Schork, 1997). In other words, 'Complex Disease' is a complex trait, which is polygenic in nature, meaning controlled by more than one gene or many genes at the cellular level. Such disease or trait is the product of an intricate interaction between 'nature and nurture', that is between inherited (nature) characters and the environmental (nurture) factors. For example, cancer, diabetes, coronary arterial disease etc. These diseases are also known as 'Noncommunicable Diseases'.

Check Your Progress

1) What is a complex disease?

.....

.....

.....

.....

Thus, complex diseases involve many genes in complex interactions, in addition to environmental influences. An individual may not be born with such a disease but may be at high risk of acquiring it. This is called as 'genetic predisposition' or 'susceptibility'. So, the genetic susceptibility to a particular disease due to the presence of one or more gene mutations, and/or a combination of alleles may not develop into a complex disease in an individual. However, if that individual lives a life which is unhealthy in terms of physical inactivity, fat rich food intake and so on then he/she will have much higher risk of developing a complex disease than a normal individual in a society.

The functioning of human body is an extraordinarily complex biological phenomenon. A number of biochemical networks, physiological systems, and organs must work together in intricately coordinated fashion for the human body to function properly. There are a great many features of the human body that require subtle control of a number of biochemical and physiological mechanisms. For example, respiratory control, blood pressure regulation, nutrient extraction and dispersion from food, and immune system surveillance, all involve and implicate a number of biochemical and physiological systems. These systems are typically under genetic control and often respond via feedback mechanisms to environmental stimuli, both to ensure stability (i.e., homeostasis) and to allow adaptability. What is important to consider about human physiologic complexity is that it suggests that (*1*) a number of things could go wrong that might ultimately influence a particular clinical endpoint

taken as a sign or symptom of a disease, and (*2*) different individuals could have different underlying pathologies that lead to similar phenotypic endpoints (Schork, 1997).

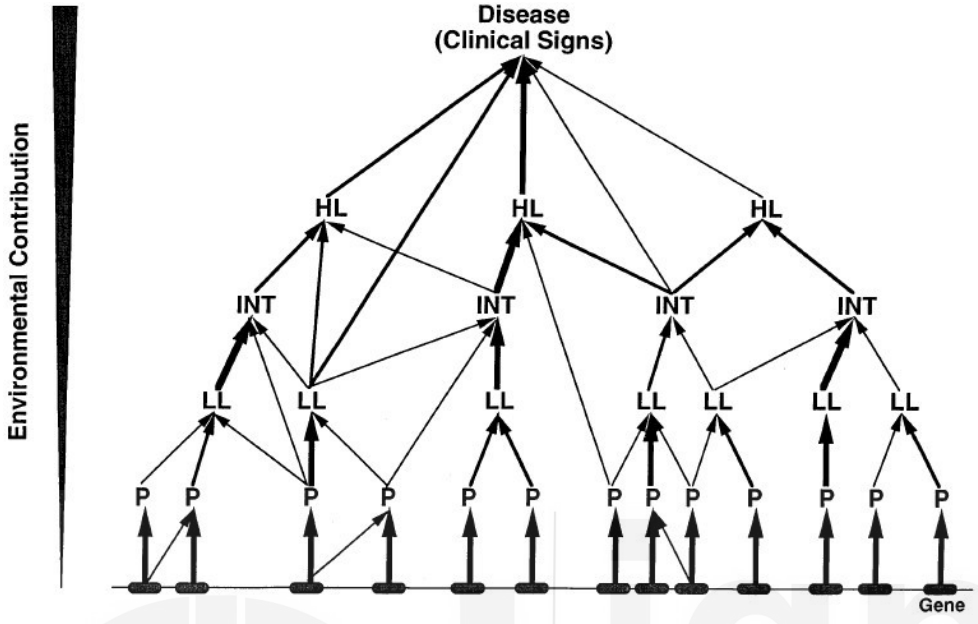


Fig. 3.2: Relationship of genes and their products to intermediate phenotypes as well as the more overt clinical manifestations of a complex disease. The thickness of the arrows denotes the strength of the contribution of a lower-level factor to a higher-level factor

(***Source:*** Schork, 1997).

3.4 HERITABILITY OF COMPLEX DISEASES

Heritability plays an important role in the development of complex diseases. For example, people whose parents or other close blood relatives have complex diseases are more likely to develop it themselves. In addition, ethnicity or genetic background is included in almost all the complex diseases risk factors due to the underlying genetic components.

3.4.1 Gene Expression

The study of inheritance would be simpler than what it is if a given genotype always expressed itself in exactly the same way in all individuals. Although there are many genes whose effect during the developmental stages of an embryo is constant under all circumstances, there are many others of which this is not true. Examples of unmistakable expression like albinism or woolly hair in human best demonstrate the regularities of gene transmission as well as expression. However, it is important to consider genes whose expressions or effects are variable.

Gene and its effect or expression are not related to each other directly, but through numerous interconnecting steps that involve complex network of enzymes, hormones, amino acids and other biochemical pathways. So, a change anywhere along this complex network of interconnections may lead

to a change in the expression of the gene. This will create variable expressions of the same gene in different individuals. However, you must also remember that the variable expression of a gene may also be caused by variability of the gene itself, that is through mutation rather than by variability in the network of reactions. This is not impossible but the probability that it is true is very small and hence negligible. All these probabilities may lead to genetic heterogeneity of the disease.

3.4.2 Gene-Environment Interaction

Just as heredity and environment are involved in the expression of any phenotypic feature, like eye colour, height etc., they are also involved in the expression of a disease. In a sense, disease is a kind of phenotype from human biology point of view. The genetic component may play the more significant role in the case of one particular disease, whereas the environment may exert a more pronounced effect in the case of some other one. For example, sex-linked muscular dystrophy and Huntington's diseases are examples of disorders in which the genetic component overshadows the environmental component. On the other hand, phenylketonuria is a disease with a very significant hereditary component, but the environmental factors can considerably alter its expression. Further, complex diseases like diabetes, cancer and hypertension are diseases in which the environmental and the genetic factors are so interrelated that their relative importance cannot be evaluated.

A disease may be known to be triggered by some infectious agent, such as a virus (smallpox, measles or even covid-19) or a bacterium (tuberculosis, leprosy). Certainly, the environmental component overshadows the genetic in these cases in the sense that the microorganism must be present for the disease to manifest itself. However, it is well known that people vary in their susceptibilities to different infectious diseases. As a consequence, when surrounded by a germ-laden environment, for example the present world-wide pandemic situation due to Covid-19, some persons will develop the disease, while others will not. Even those who will develop the disease will experience many different degrees of severity. Studies of disease incidence among identical twins, unrelated persons and comparisons of different populations and ethnic communities have confirmed that even infectious diseases have a genetic component which determines the relative susceptibility of the individual to the microorganism or agent which spread the disease.

Thus, the interrelationship of environmental and genetic factors in disease may be so complex, especially in complex diseases, that the very distinction between the words "environmental agent" and "hereditary factor" become blurred.

2) How genetics play a role in the prevalence of complex diseases?

.....

.....

.....

.....

3.4.3 Epidemiology of Complex Diseases

Dictionary meaning of 'Epidemiology' is the study of diseases that affect large number of people. Originally, epidemiologists were concerned only with infectious diseases like typhoid and influenza, which arise and spread rapidly among the human populations as epidemic. However, in recent years epidemiology also deals with non-infectious or non-communicable or complex diseases like diabetes, heart disease and cancer. Typically, the distribution of a disease is charted in order to discover patterns that might give information about its mode of transmission or the susceptibility of certain ethnic groups. This in turn can reveal insights about the causes of the complex disease and possible preventive measures.

Thus, the epidemiologist records a variety of dietary, lifestyle, behavioral, and social factors, providing an interface between the environment and the genome. Epidemiological variation, whether genetically or environmentally determined, contributes to inter-individual variation in gene expression and thus to variation in common complex disease risk.

3.5 ENVIRONMENTAL FACTORS AFFECTING COMPLEX DISEASES

In non-communicable or complex diseases, the whole complex of environmental factors and biological responses (inborn and acquired) are generally considered when dealing with the regional variation. For example, the fact that Africans are more susceptible to frostbite than are Eskimos or North American Indians may well be attributed to both lack of acclimatization and some genetic susceptibility. The malformation of the central nervous system like spina bifida and anencephaly have been shown from different family and ethnic studies, the role of genetic factors, social class, birth order and maternal age effects as well as secular seasonal variation, which indicate that environmental factors are also important in their causation. Similarly, the geographical aspect of the complex disease like cancer has received academic attention in the hope that studies of local environmental conditions associated with high or low incidences of cancer may give valuable information about the causes behind the disease. There are many environmental factors that aggravate complex diseases. Few important ones are discussed below:

3.5.1 Nutrition

Nutrition in terms of the food that we eat affects almost all the complex diseases directly. Our dietary habits have considerably changed from our prehistoric ancestors' time. For example, we get most of our calories from grains, domesticated livestock, dairy products and refined sugars, whereas our ancestors ate naturally occurring plant foods and wild animals. They almost never consumed cereal grains, dairy foods, separated oils or commercially processed food. In fact, our ancestors consumed twice the animal protein that current western populations do. The proportion of total fat consumed by them was also roughly equal to the contemporary Americans, yet they never developed obesity or heart diseases or any other complex diseases. The reason behind this is that they never had cholesterol rich diet and they received carbohydrate almost entirely from fruits and vegetables. These made their diet healthy with less sodium, but more potassium, fiber and micronutrients. Drastic shift in the nutrition intake from fruits and vegetables based diet towards dairy and cholesterol rich diet is one of the most crucial factors for complex diseases.

3.5.2 Physical Inactivity

In the past people had to work physically in order to eat. With the rapid industrialization and modernisation, energy expenditure through physical exertion has decreased drastically for all human populations. Physical activity has important effects on muscular strength, skeletal robusticity, all of which were present in our ancestors. Lacking such physically active life style results into development of huge adipose tissues in our body, thereby increasing the risk of diabetes, atherosclerosis and other complex diseases.

3.5.3 Psycho-social Factor

Genes affecting human behaviour are ancient and probably evolved much in the same manner as other physical characters. Like current sedentary life style and diet, the social circumstances of contemporary societies are also new and very different from our ancestor's time. For example, physical contact between infant and adults is much reduced. In most affluent societies, babies do not sleep with their mothers. All these differences along with work related stress, frequent contact with strangers, conflicting social roles, wage labour, working in bureaucracies, reduced support from relatives and siblings may contribute to syndrome such as attention deficit or hyperactivity, depression, anxiety disorders and substance abuse (Eaton, 2010). These psychosocial factors ultimately aggravate the risk of developing various complex diseases.

3.6 REPRESENTATIVE COMPLEX DISEASES

The entire list of complex diseases is so long. But only five complex diseases that have received worldwide attention due to their prevalence in several human populations will be discussed here.

3.6.1 Diabetes

Diabetes' is the condition in which the body does not properly process food for use as energy. Most of the food that we eat, especially carbohydrate rich food, is turned into glucose, or sugar, for our bodies to use for energy. The pancreas, an organ that lies near the stomach, makes a hormone called 'Insulin' to help glucose get into the cells of our bodies after digestion. When you have diabetes, your body doesn't make enough insulin to complete this metabolic pathway. This causes sugars to build up in your blood causing serious health complications, including heart disease, blindness, kidney failure, and lower-extremity amputations.

There are three major classes of diabetes: Type I, Type II and Gestational Diabetes. Type I diabetes, previously called insulin-dependent diabetes mellitus (IDDM) or juvenile-onset diabetes, generally accounts for 5%-10% of all diagnosed cases of diabetes. Risk factors are less well defined for Type I diabetes than for Type II diabetes. Type II diabetes was previously called non-insulin-dependent diabetes mellitus (NIDDM) or adult-onset diabetes. Type II diabetes generally accounts for about 90%- 95% of all diagnosed cases of diabetes. Risk factors for Type II diabetes include older age, obesity, family history of diabetes (hereditary), prior history of gestational diabetes, impaired glucose tolerance, physical inactivity, and race/ethnicity. Gestational diabetes is a type of diabetes, or high blood sugar, that only pregnant women get. It is one of the top health concerns related to pregnancy. If not treated, gestational diabetes can cause problems for mothers and babies (WHO, 2015).

Box 3.2 : Symptoms of Diabetes

☐ Frequent urination	☐ Tingling or numbness in hands or feet
☐ Excessive thirst	☐ Feeling very tired much of the time
☐ Unexplained weight loss	☐ Very dry skin
☐ Extreme hunger	☐ Sores that are slow to heal
☐ Sudden vision changes	☐ More infections than usual

☐ Nausea, vomiting, or stomach pains generally accompany some of these symptoms in the abrupt onset of insulin-dependent diabetes, now called Type I diabetes.

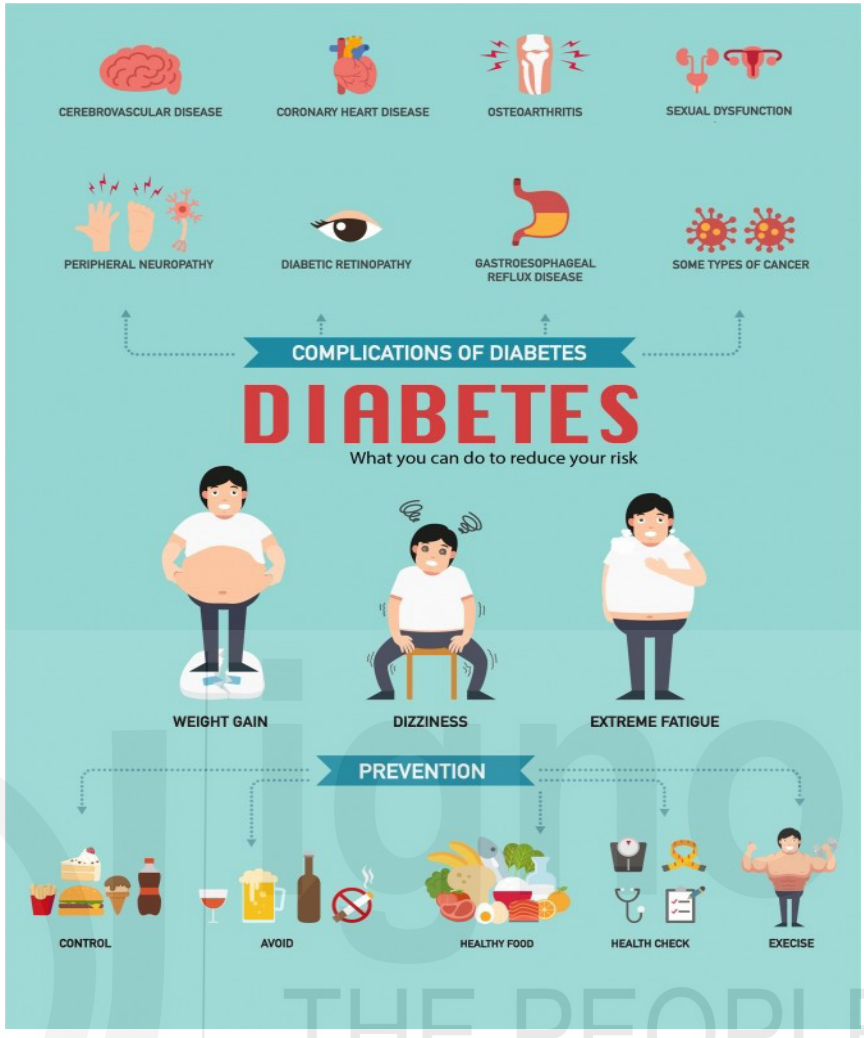


Fig. 3.3: Symptoms and Prevention of Diabetes
(***Source:*** https://www.healthysolutionsmd.com.)

3.6.2 Hypertension

Blood Pressure (BP) is the force exerted by the blood against a vessel wall. BP depends on the volume of blood contained within the vessel and the compliance of the vessel walls, i.e. how easily they can be stretched. The maximum pressure exerted in the arteries when blood enters into them during systole is called the Systolic Blood Pressure (SBP). The normal value for SBP is 120 mm Hg. The minimum pressure within the arteries when blood is draining off into the remainder of the vessels during diastole is called the Diastolic Blood Pressure (DBP). The normal value for DBP is 80 mm Hg. So, the systolic blood pressure is the peak pressure exerted in the arteries when blood is pumped into them during ventricular systole. The diastolic blood pressure is the lowest pressure exerted in the arteries when blood is draining off into the vessels downstream during ventricular diastole.

Elevated systolic or diastolic blood pressure is known as "Hypertension". More specifically, when an individual shows resting systolic blood pressure as greater than 140 mm Hg and/or diastolic pressure exceeding 90 mm Hg, it is considered as the presence of hypertension. One of the main factors that affects blood pressure is peripheral resistance (Figure 3.4). If resistance increases, then more

pressure is needed to keep blood moving, this causes hypertension. Three main sources of peripheral resistance: blood vessel diameter, blood viscosity and total vessel length. If blood vessel diameter is smaller more blood will be near wall and hence resistance will increase and so the BP. Blood viscosity is related to the thickness of blood. The greater the viscosity, the less easily molecules slide past one another and the more difficult it is to get the blood moving, which increases resistance and in turn BP. Because of this greater resistance to flow, a greater pressure is required to pump the same volume of viscous or thick blood. Total vessel length affects peripheral resistance and in effect BP. The longer the total vessel length, the greater the resistance encountered, and the greater the blood pressure.

Besides peripheral resistance, blood vessel elasticity also affects blood pressure. A healthy elastic artery expands during blood flow, the elastic recoil of the vessel then maintains the continued flow of blood. Blood volume affects blood pressure. When there is a greater volume of blood, more blood presses against the walls of the arteries resulting in a greater BP. An increase in cardiac output, which is the amount of blood that moves out of heart per minute, results in increased blood pressure.

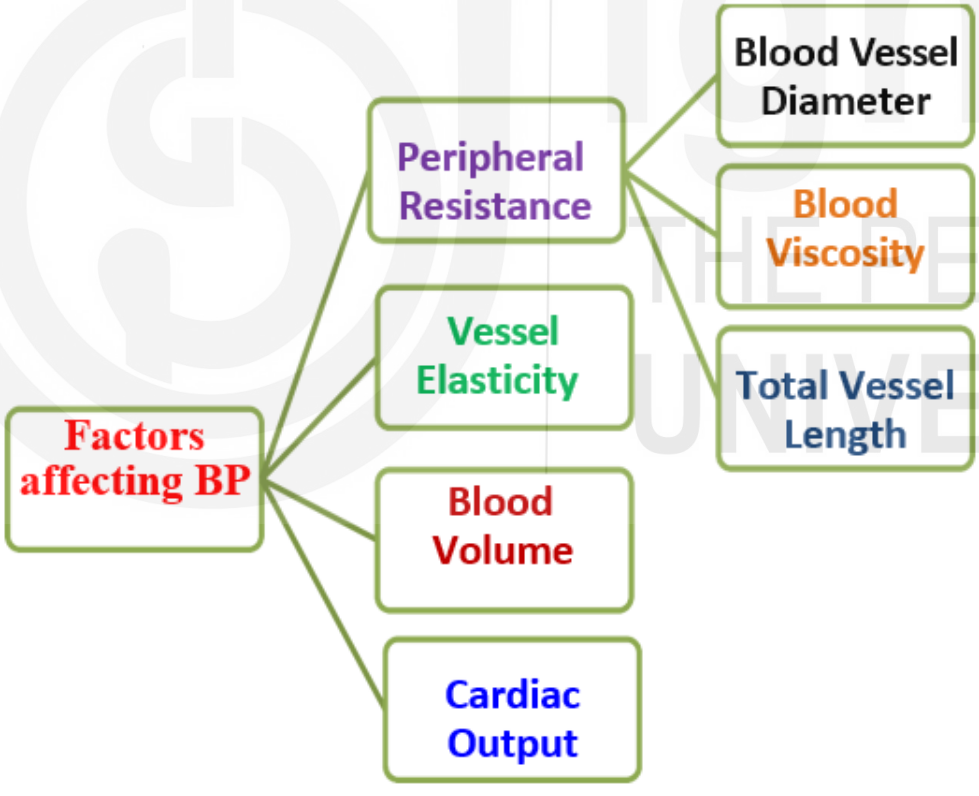


Fig. 3.4: Factors affecting Blood Pressure and causing Hypertension

Hypertension or high blood pressure is a complex disease that produces no symptoms. It has become known as the "Silent Killer", because its complications affecting the eyes, kidneys, brain and heart can be deadly. Persons with hypertension double their risk of heart attack and have four times more risk of stroke than do individuals with normal BP.

3.6.3 Cardiovascular Disease

Major cardiovascular diseases (CVD) are a group of disorders of the heart and blood vessels that include coronary heart disease (CHD), cerebro-vascular disease, heart failure, rheumatic heart disease and congenital heart disease. The major risk factors associated with cardiovascular diseases are cigarette smoking, unhealthy diet, physical inactivity, hypertension, diabetes and high blood cholesterol. CVD may also result from a variety of genetic causes, including single-gene mutations, the interaction of multiple genes and environmental factors. Economic transition, urbanisation, industrialization and globalisation bring about lifestyle changes that promote heart disease.

More specifically, CVD involves degenerative changes in the heart muscle and the inner lining of the larger arteries that are mainly responsible for smooth and efficient blood flow. In other words, deposition of fatty cells in the inner wall of arteries reduce the diameter of blood vessel (Figure 3.5) that results into high BP. Damage to arterial walls begins as a multifactorial response to injury, mostly from hypertension, cigarette smoking, and high cholesterol level in blood.

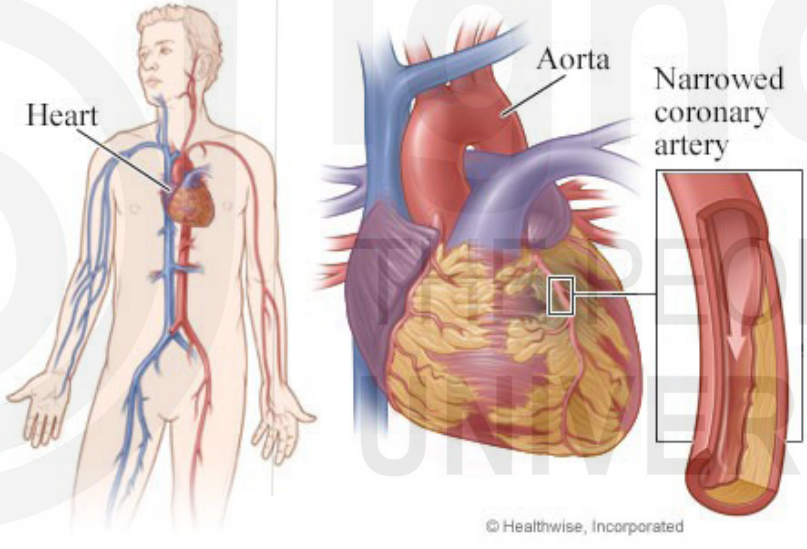


Fig. 3.5: Narrowing of larger artery like Aorta due to fat deposition
(***Source:*** https://myhealth.alberta.ca/Health/pages/conditions.Govt.of Alberta)

3.6.4 Obesity

A person with a Body Mass Index (BMI: Height and Weight ratio) above 30 for world population and above 25 for Asian population is considered as "Obese". However, you should remember that a person may weigh considerably more than the average weight-for-height standard yet still be rated as 'under-fat' for body composition. Extra weight for this person exists as excessive muscle mass. Therefore, the term obesity or overweight refers only to a body mass in excess of some standard, usually the average for a given height.

Excessive fat in youth is even more of an adult health risk than obesity begun in adult hood. Overweight children and adolescents, exhibit a significantly higher

risk of a broad range of diseases as adult than adolescents of normal weight. Obesity leads to a constellation of complex diseases like diabetes, hypertension, coronary heart disease, cancer and many more. It is excess body fat, not excess body weight that explains the relationship between above average body weight and disease risk. Thus, obesity defined as an excess accumulation of body fat, is a heterogeneous disorder with a final common pathway in which energy intake chronically exceeds energy expenditure. Disruption in energy balance often begins in childhood, and if this occurs, the chance for adult obesity is significantly higher than in children with normal body fat levels. If parental obesity also exists, the child's risk of obesity in adulthood is two to three times that of normal weight children without obese parents. Obesity results from a complex interaction of factors, including genetic, environmental, metabolic, physiologic, behavioural, social and even racial influences. The significant interaction between genetics and environment makes it difficult to quantify the role of each in the development of obesity. One's genetic makeup does not necessarily cause obesity, but can make one more vulnerable to develop obesity in conducive environment.

3.6.5 Cancer

Cancer is a complex disease in which a person's body cells divide uncontrollably and destroy body tissues. Cancer is the name given to a collection of related diseases. There can be various types of cancer in humans. For example, breast cancer, prostate cancer, lung cancer, skin cancer, colon cancer, blood cancer (Leukemia) and so on. In all types of cancer, some of the body's cells begin to divide without stopping and spread into surrounding tissues. Cancer can start almost anywhere in the human body, which is made up of trillions of cells. Normally, human cells grow and divide to form new cells as the body needs them. When cells grow old or become damaged, they die, and new cells take their place. When cancer develops, however, this orderly process breaks down. As cells become more and more abnormal, old or damaged cells survive when they should die, and new cells form when they are not needed. These extra cells can divide without stopping and may form growths called tumors (National Institute of Health).

According to the World Health Organisation (WHO, 2015), cancer occurs because of mutations in the genes responsible for cell multiplication and repair. The change which a cell undergoes in the process of abnormal transformation is a reflection of the genetic alterations.

WHO's department of Noncommunicable Diseases and Mental Health (NMH) has done extensive work on major complex or noncommunicable diseases, like cancer, diabetes, cardiovascular disease, asthma, and some mental illnesses. In some cases, such as cancer, individuals are born with genes that are altered by lifestyle habits or exposure to chemicals. Cancer, for example, may involve tumour-suppressor genes, genes which suppress tumour formation, which lose their function, thus giving rise to carcinomas.

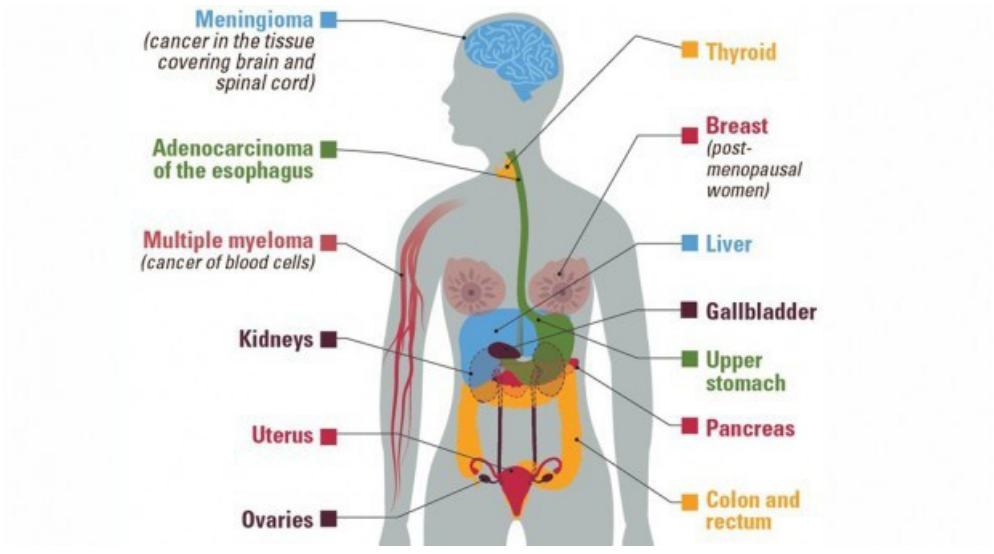


Fig. 3.6: Thirteen types of Cancer that are linked to Obesity
(***Source:*** https://www.cancerhealth.com)

3.7 PREVALENCE OF COMPLEX DISEASES

The comparison of populations for complex diseases in contrasted environments showed that the prevalence or incidence of these diseases varies according to the ethnicity, food habits, climatic conditions and life style of the population. So, the frequency of complex diseases is not equally distributed across the world. Here, we will learn about the occurrence of complex diseases in different populations across the world and specifically in India.

3.7.1 Worldwide

Diabetes is the seventh leading cause of death in the United States, where one third of the entire population has diabetes. African Americans, Hispanic/Latino Americans, American Indians, and some Asian Americans and Pacific Islanders are at particularly high risk for type 2 diabetes. Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year. Four out of 5CVD deaths are due to heart attacks and strokes, and one third of these deaths occur prematurely in people under 70 years of age. Cardiovascular disease (CVD) tends to manifest itself in specific ways unique to various communities. For example, African communities tend to have strokes as a result of cardiovascular disease, while South Asians tend to have heart attacks. Asian populations, especially Asian Pacific Islanders (Hawaii, Japan, etc) consistently have the lowest demonstrated risk of developing hypertension, with an average lifetime risk of about 9.5% in men and 8.5% in women, as compared to the high prevalence of hypertension in African- American (36%), Caucasian (20%), Native American, and Hispanic populations.

Various Asian populations have shown that the decreased overall incidence of hypertension tends to moderate when these people are placed into different cultural circumstances. For example, when native Vietnamese populations migrate to the United States, their risk of developing hypertension and

cardiovascular diseases tend to approach that of Caucasians within a short time. These results raise the question that it might be other, societal forces that contribute to differing rates of complex diseases among various ethnic groups in the developed world.

Obesity is now the second leading cause of preventable deaths in the United States; 300,000 deaths early, after CVD. Obesity among youth has more than doubled in the last 15 years. Obesity now claims 15-20% of American children and 12% of adolescents. By 2025, 75% of the American population will classify as overweight, with about one-third of these individuals rated obese (WHO).

3.7.2 In India

CVD is the leading cause of mortality in India, where the country contributes almost one-fifth (18.6%) of the global CVD burden. Overall, the mean 10-year risk of a CVD event in the total Indian population is 12.7% among females and 21.4% among males. Geographic variation of CVD risk varied from 10.2% among females in Assam to 24.2% among males in Nagaland and Himachal Pradesh (Figure 3.7). Similarly, prevalence of a high CVD risk varied from 5.0% among females in Assam to 30.4% among males in Kerala. Among both males and females, CVD risk tended to be highest in South India (including Goa), three northern states like Himachal Pradesh, Punjab, and Uttarakhand), the North-Eastern states (except Assam), and West Bengal (Geldsetzer et al., 2018). It has been predicted that CVD will be the most important cause of mortality in India by the year 2025.

India is experiencing rapidly escalating epidemics of diabetes and obesity. The prevalence of type 2 diabetes in urban Indian adults has increased from <3% in the 1970s to >12% in 2000.

India has the highest number of diabetic patients of any country. Over the past 25years, the prevalence of coronary heart disease (CHD) in Indian adults has increased from <2% to >10%. It is predicted that by 2025 India will have more than 60 million diabetic patients. In other words, one in five diabetic patients in the world will be Indian (Yajnik, 2004).

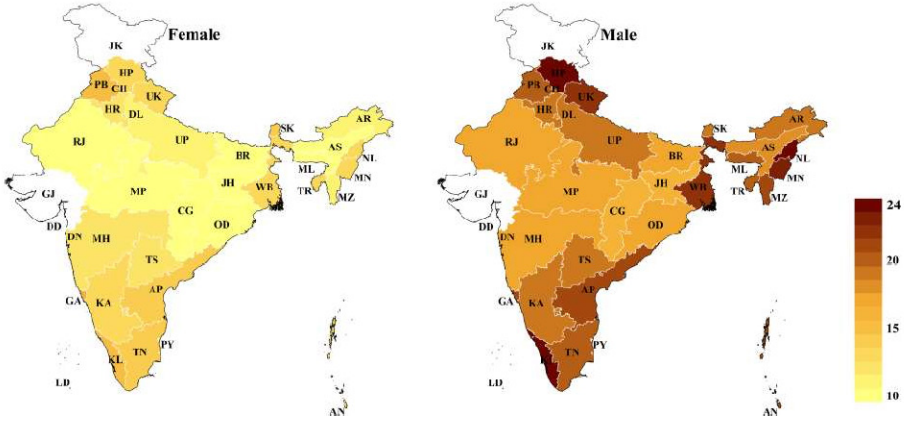


Fig. 3.7: State wise percentage distribution of CVD risk events in India
(***Source:*** Geldsetzer et al., 2018)

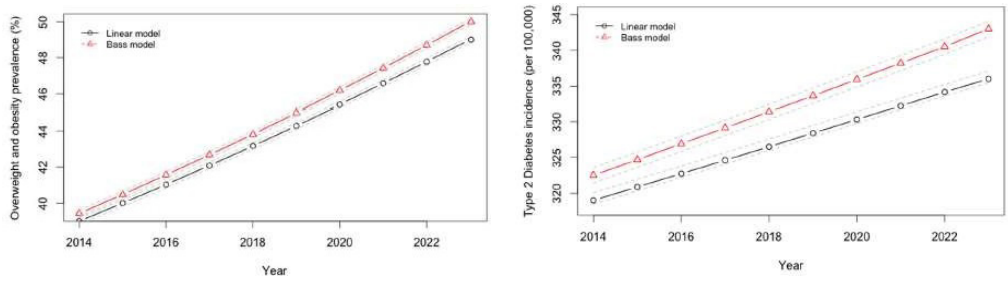


Fig. 3.8: Prediction of Obesity and Type 2 Diabetes incidence in India, 2013-2024
(***Source:*** Basu et al., 2014)

3.8 SUMMARY

Complex diseases are a manifestation of a complicated network between gene and environment. Again, gene or the genetic component itself is a complex interface between the genotype, enzymes, hormones, amino acid and other factors responsible for effective biochemical pathways leading to gene expression. Similarly, environmental component in turn is a multifaceted interaction between factors like ethnicity, life style, climate, food, physical activity, psychological development and so on.

So, the question of possible genetic influence of different ethnic groups on developing complex diseases is not simply an academic curiosity. Rather, these differences, if present, hold the promise for individualized treatment targeted to specific factors that may differ between ethnic groups. For instances, there is a clear body of scientific evidence supporting the idea that members of certain ethnicities do tend to develop hypertension (a complex disease) more often than people from other ethnic backgrounds. In nearly all of the studies on ethnicity as a risk factor for hypertension, two groups emerge as having risk that is much different than that for the population on average. African-Americans, for example, consistently lead incidence profiles in hypertension studies, with about 36% of the population developing hypertension at some point. This is compared to about 20% in the Caucasian, Native American, and Hispanic populations.

3.9 REFERENCES

Basu S, Vellakkal S, Agrawal S, Stuckler D & Popkin B. (2014). Averting Obesity and Type 2 Diabetes in India through Sugar-Sweetened Beverage Taxation: An Economic- Epidemiologic Modeling Study". *PLoS Med* . 11: e1001582.

Cavalli-Sforza, L.L. & Bodmer, W.F (1971). *The Genetics of Human Population*, San Francisco: Freeman

Cancer Research: https://www.cancer.gov/about-cancer/

Centre for Disease Control and Prevention. www.cdc.gov/media/presskits/aahd/diabetes.pdf

Cummings M.R. (2011). *Human Heredity: Principles and Issues*. Ninth Edition. Brooks/Cole, Cengage Learning

Eaton, SB. (2010). Complex chronic diseases in evolutionary perspective, pp. 491-501. In: "Human evolutionary biology". Muehlenbein, MP (Ed.). Cambridge University Press, UK.

Geldsetzer P, Manne-Goehler J, Theilmann M, Davies JI, Awasthi A & Danaei G. (2018). Geographic and sociodemographic variation of cardiovascular disease risk in India: A cross-sectional study of 797,540 adults". *PLoS Med* 15(6): e1002581.

Montagu, A. (1960). Human Heredity. USA: Signet Press.

Narain, P. (1988). *Quantitative Inheritance*, pp. 311-332. In: "Statistical methods in human population genetics". Malhotra, K.C. (Ed.). Indian Institute of Bio-social Research and Development. Indian Statistical Institute. Indian Society of Human genetics.

Relton, CL & Smith, DG. (2010). "Epigenetic Epidemiology of Common Complex Disease: Prospects for Prediction, Prevention, and Treatment". *PLoS Med* Vol. 7(10): e1000356. https://doi.org/10.1371/journal.pmed.1000356

Relethford J.H. (2012). *Human Population Genetics*. USA: Wiley-Blackwell.

Schork, NJ. (1997). "Genetics of complex disease: approaches, problems, and solutions". *Am J Respir Crit Care Med*, Vol.156:S103–S109.

Susanne, C. (1982). Genetic and environmental factors during the growth period. Belgium: Free university of Brussels, Brussels.

WHO(2015).https://www.who.int/genomics/public/geneticdiseases/en/index3.html

Yajnik CS. (2004). Early life origins of insulin resistance and type 2 diabetes in India and other Asian countries. *J Nutr*. 134:205-210. doi:10.1093/jn/134.1.205.

3.10 ANSWERS TO CHECK YOUR PROGRESS

1) When a disease is caused due to the interaction between multiple genes and various environmental factors, including food, life style, climate and psychosocial elements, such disease is called 'Complex Disease'. Examples of complex diseases are cancer, diabetes, hypertension and other cardiovascular diseases.

2) Complex diseases involve many genes in complex interactions, in addition to environmental influences. An individual may not be born with such disease but may be at high risk of acquiring it as compared to the general population. Meaning, an individual is genetically susceptible to develop complex diseases, if they carry candidate genes or genes responsible for these diseases in their cells. So, the genetic susceptibility to a particular complex disease due to the presence of one or more genes may not develop into a complex disease in an individual. However, if those individual lives a life which is unhealthy in terms of physical inactivity, fat rich food intake and so on then he/she will have much higher risk of developing a complex disease than a normal individual in a society. On the other hand, if that genetically predisposed individual lives a very healthy life then he/she may not develop any complex disease in their life time, despite carrying the gene.

ignou
THE PEOPLE'S
UNIVERSITY