# MANE-001
# Human Genetics

Indira Gandhi
National Open University
School of Social Sciences

Block

# 4

## HUMAN BIOCHEMICAL GENETICS

## Expert Committee

Prof. P. Dash Sharma (Retd.)
Dept. of Anthropology
Ranchi University, Ranchi

Prof. P. Veerraju (Retd.)
Dept. of Human Genetics
Andhra University
Visakhapatnam

Prof. S.M.S. Chahal
Dept. of Human Biology
Punjabi University, Patiala

Prof. A. Papa Rao
Dept. of Anthropology
Sri Venkateswara University
Tirupati

Dr. Roli Mathur
Scientist 'C'
Division of Basic Medical Sciences
Indian Council of Medical Research
New Delhi

Dr. Seema Kalra
Assistant Professor
Dept. of Biochemistry
IGNOU, New Delhi

**Faculty of Anthropology**
**SOSS, IGNOU**

Dr. Rashmi Sinha, Reader
Discipline of Anthropology
IGNOU, New Delhi

Dr. P. Venkatramana
Assistant Professor
Discipline of Anthropology
IGNOU, New Delhi

Dr. Rukshana Zaman
Assistant Professor
Discipline of Anthropology
IGNOU, New Delhi

Dr. Mitoo Das
Assistant Professor
Discipline of Anthropology
IGNOU, New Delhi

Dr. K. Anil Kumar
Assistant Professor
Discipline of Anthropology
IGNOU, New Delhi

Academic Assistance provided by
Dr. N.K. Mungreiphy, Research Associate
(DBT) for the Expert Committee meeting.

## Programme Coordinator: Dr. Rashmi Sinha, SOSS, IGNOU, New Delhi

## Course Coordinator: Dr. P. Venkatramana, SOSS, IGNOU, New Delhi

## Content Editor

Prof. P. Veerraju (Retd.)
Department of Human Genetics
Andhra University, Visakhapatnam

## Block Preparation Team

**Unit Writers**

Prof. C.R. Srikumari Srisailapathy (Unit 1)
Department of Genetics, University of Madras
Chennai

Prof. S.M.S.Chahal (Unit 2)
Department of Human Biology, Punjabi University, Patiala

Prof. P.K.Das (Unit 3)
Department of Anthropology, Utkal University, Bhubaneswar

### Cover Design

Dr. Mitoo Das
Asstt. Professor
Anthropology
SOSS, IGNOU
New Delhi

*Authors are responsible for the academic content of this course as far as the copyright issues are concerned.*

## Print Production

Mr. Manjit Singh
Section Officer (Pub.), SOSS, IGNOU, New Delhi

July, 2012



ISBN-978-81-266-6136-7



*Further information on Indira Gandhi National Open University courses may be obtained from the University's office at Maidan Garhi. New Delhi-110 068.*

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi by the Director, School of Social Sciences.

Laser Typeset by : Tessa Media & Computers, C-206, A.F.E.-II, Okhla, New Delhi

Printed at :

# BLOCK 4 HUMAN BIOCHEMICAL GENETICS

Human genetic variation can be normal and abnormal. The nature of normal variation is in general more frequent while that of abnormal variation is usually rare. Further for most of the genetic characters/diseases described in literature so far, biochemical basis is illustrated only in some conditions. The discussion of these conditions (normal genetic characters as well as abnormal genetic diseases/ disorders) whose biochemical basis is identified constitute the subject matter of human biochemical genetics. Therefore the discussion on metabolic diseases falls under the purview of abnormal variation while that of genetic characters such as blood groups, red cell enzymes and serum proteins falls under the purview of normal variation and again the sum total of the discussion of these two categories of variation fall within the framework of human biochemical genetics.

Concerning the normal genetic characters such as blood groups, red cell enzymes and serum proteins, the techniques employed to detect these types of the characters are different. While the agglutination techniques are employed for detection of blood group antigens present on the surface of the red blood cell, the electrophoretic techniques are employed for detection of enzymes found within the red cell (that is, in the cytoplasm of the cell). For identification of the serum proteins also electrophoretic techniques are employed. That is to say while the corpuscular or cellular part of the blood, that is, the red cell is used for the elucidation of blood groups and red cell enzymes, the liquid portion of the blood i.e., the plasma or serum is used for illustration of plasma/serum proteins.

To enlighten yourself with the different aspects of biochemical genetics we have included few examples of blood groups systems and metabolic diseases (otherwise called inborn errors metabolism).

Here in this Block, you will be able to enlighten yourself with human genetic variation having biochemical basis to enable us to detect these variations in a way to focus further on treatment aspect of the genetic diseases, having a clinical significance.

ignou
THE PEOPLE'S
UNIVERSITY

# UNIT 1 MEANING, SCOPE AND GENETIC VARIATION

**Contents**



**Learning Objectives**

Once you have studied this unit, you will be able to:

- know about introduction to human bio-chemical genetics;
- describe one gene one polypeptide concept;
- depict Inborn errors of metabolism with typical examples; and
- elucidate human ABO blood group system and its fundamentals.

## 1.1 BIOCHEMICAL GENETICS

In the light of the understanding of basic aspects of genetic variation, let us now move on to one of the very important specialty in human genetics – known as Biochemical Genetics.

### 1.1.1 History

Gregor Mendel (1822 – 1884) (Fig. 1.1) is well known as the founder of modern genetics. He conducted breeding experiments from 1856 to 1863 and presented his results in 1865. Using the different varieties of pea plants interms of the flower colour, the seed coat type and many such simple characteristics carefully chosen as the varieties, he analysed the pattern of transmission of these traits in subsequent generations. This led to the discovery of principles of heredity which is now known as Mendelian Inheritance in his honour.

The impact of this research, appeared to be initially minimal and no one realized that Mendel has discovered the basic principles of inheritance. Unfortunately, the work remained dormant for nearly 44 years.

In 1900, three botanists Hugo de Vries, Erich von Tschermark, and Karl Correns did similar experiments with plants and arrived at conclusions similar to those of Mendel.

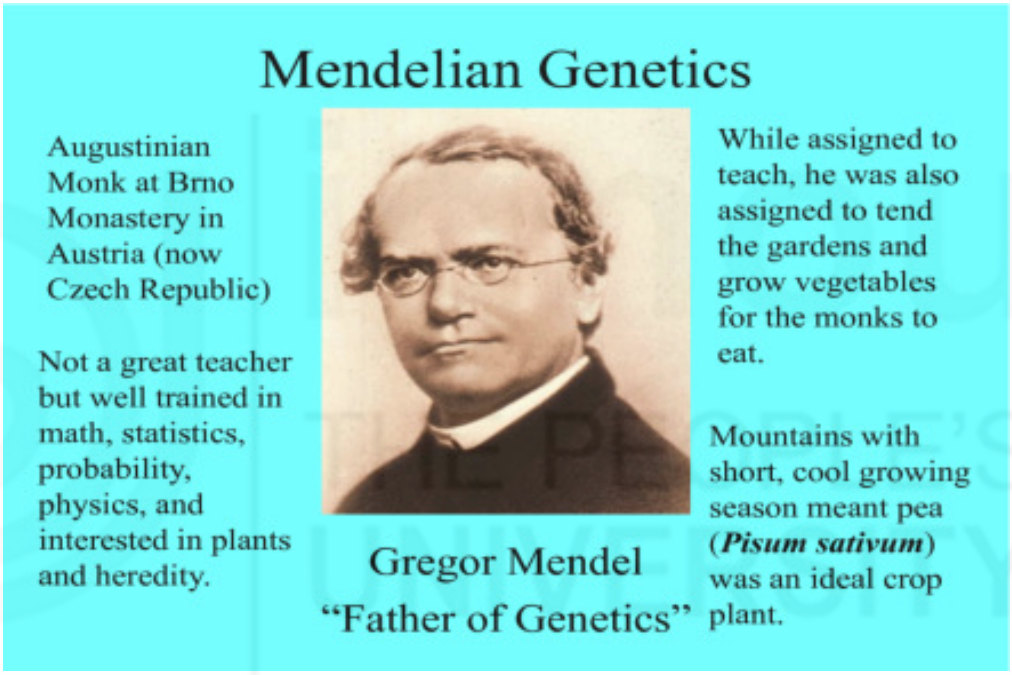



**Fig. 1.1: Gregor Mendel (Source:rocketmanagement.co.uk)**

### 1.1.2 Garrod's Inborn Errors of Metabolism (IEM)

At the time that Mendel's work was being discovered, English Physician and biochemist Archibald Garrod was studying a rare condition (incidence of 1 in 250,000) known as alkaptonuria (uria means "in the urine") otherwise healthy new born with the disorder, usually had one striking feature. Their urine upon standing in contact with air, turned black. In their later years, affected individuals show darkening of portions of ears, whites of eye, cartilage and tendons. One to the deposits in the cartilage they suffer with arthritis. Males develop black stones in the prostrate at later age). The dark substance in the urine was the oxidation product of alkapton now called homogentisic acid, which contains a six- carbon ring structure. Other researchers found that the excretion of homogentisic acid in urine was increased, when people with alkaptonuria were fed excess protein or the amino acid tyrosine and phenylalanine, both of which contain benzene rings. Yet individuals who have no alkaptonuria never excrete any homogentisic acid in urine.

Garrod suggested that homogentisic acid is an ordinary product of metabolism that is broken down in normal people but is not degraded in people with Alkaptonuria. This metabolic block results in homogentisic acid being excreted intact in the urine. More specifically, Garrod thought that people with alkaptonuria do not properly metabolize ring- containing fractions of proteins because they lack a special enzyme that is present in unaffected individuals. He proposed that during the normal break down of proteins, phenylalanine is converted to tyrosine, which in turn is changed to homogentisic acid and then simpler to complex products.

Garrod's analysis did not stop with the concept of metabolic blocks, however. The brilliance of his proposal lay in the genetic connection that he and population geneticist William Bateson were able to make. Garrod noted that alkaptonuria tended to occur in several sibs of a family whose parents were unaffected and also that many affected children were the offspring of first-cousin marriages. This pattern of inheritance fit the requirements for a rare recessive trait, the first human application of Mendel's newly discovered laws.

Garrod classified Alkaptonuria and three other conditions that he called as Inborn Errors of Metabolism. The four canonical inborn errors described by Garrod are Albinism, Alkaptonuria, Cystinuria and Pentosuria. He suggested that these four genetic conditions must represent only a tiny fraction of the errors of metabolism that exist in human populations, pointing out that many of them may produce no obvious externally manifesting phenotypic effects.

Archibald Garrod (Fig. 1.2) served from 1915 to 1919 as a colonel in the British Army Medical Services in Malta. Then although Garrod's medical career was extremely successful and distinguished, his pioneering insights into the nature of metabolic disorders and the principle of biochemical individuality was not appreciated during his life time.

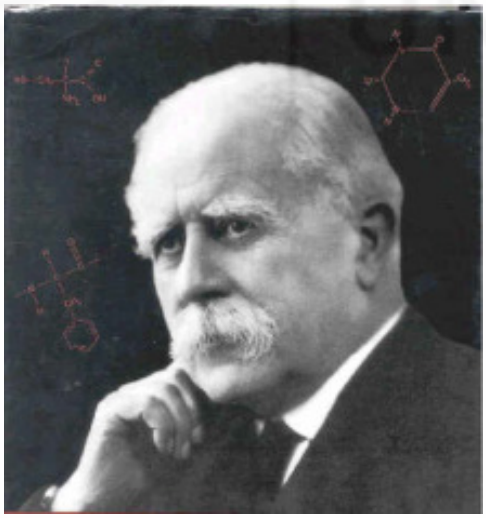


**Fig.1.2: Archibald Garrod (1857-1936) (Source: powerofthegene.com).**

Garrod an extraordinary physician scientist made discoveries that were ahead of his time and died long before, we would appreciate fully the information, the knowledge and the wisdom he had passed on to us. Garrod belonged to the period 1857 to 1936. During his lifetime the natural sciences changed our view of the world, and genetics began its journey to the double helix.

Thus was born the Garrod's bold hypothesis of "inborn errors of metabolism" with its far- reaching assumption that genes were there to produce chemical catalysts, one gene to each highly specialized catalyst.

Garrod the initial founder of human "Biochemical Genetics" belonged to a unique community ahead of his times as rightly appreciated by Scriver (2001) in his felicitation lecture, that Garrod belonged to the clinical-scientic community. He introduced a paradigm new for its day in medicine: he thought that biochemistry is dynamic and different from the static nature of organic chemistry. So this led him to think about metabolic path ways and heredity could explain an "inborn error of metabolism". During Garrod's time, he had no idea about the nature of gene. Genes are now well understood: Genomes are being described for Homo sapiens and it is understood that genomes "speak biochemistry ( not phenotype)" Scriver (2011).

### 1.1.3 Beadle and Tautum – Genetic Analysis Led to the One-Gene-One Enzyme Hypothesis

Garrod's ideas, like Mendel's were largely ignored for over thirty years although Garrod published a book and several papers on inborn errors of metabolism. The concepts have to be rediscovered by the American geneticists in their influential experiments connecting genes with enzymes. This was carried out in 1940s by George W. Beadle and Edward L. Tautum using a filamentous fungus called *Neurospora crassa* commonly called red bread mould. They chose this organism because both genetic and biochemical analysis could be done with ease.


**Fig.1.3: Beadle and Tautum (Source: Sandwalk.blogspot.com)**

In these experiments they identified new mutations that each caused a block in the metabolic pathway for the synthesis of some needed nutrients and showed that each of these blocks corresponded to a defective enzyme needed for one step in the pathway. The experimental approach now called genetic analysis was important because it solidified the link between the genetics and biochemistry.

*N. crassa* grows in the form of filaments on a great variety of substrates including laboratory medium containing only inorganic salts, a sugar and one vitamin. Such a medium is called as minimal medium because it contains only the nutrients that are essential for growth of the organism. The filaments consists of a mass of branched threads separated into interconnected, multinucleate compartments

allowing free interchange of nuclei and cytoplasm. Each nucleus contains a single set of seven chromosomes. Beadle and Tatum recognized that the ability of *Neurospora* to grow in minimal medium implied that the organism must be able to synthesize all of the other small molecules needed for growth such as amino acids. If the biosynthetic pathways needed for growth are controlled by genes, then a mutation in a gene responsible for synthesizing an essential nutrient would be expected to render a strain unable to grow unless the strain were provided with the nutrient.

The classic experiments of Beadle and Tautum thus showed that the relationship is usually remarkably simple: one gene code for one enzyme. The pioneering experiments united genetics and biochemistry, and for the "one gene one enzyme" concept Beadle and Tautum were awarded a Nobel prize in 1958.

Beadle and Tautum were fortunate to study metabolic pathways in relatively simple organism in which each gene specifies a single enzyme, a relation often called the one enzyme hypothesis.

### 1.1.4 Genetic Code

In 1953 James Watson and Francis Crick solved the structure of DNA and identified the base sequence as the carrier of genetic information. But the way in which the base sequence of DNA specifies the amino acid sequence of proteins (the genetic code) was not obvious until 10 years.

**Fig. 1.4: Watson (left) and Crick (right) (Source: A.Barrington Brown/Science Photo Library/Photo Researchers).**

Three nucleotides are necessary to specify a single amino acid. This is the basic unit of the genetic code. So the set of basis that encode a single amino acid is called as a codon. Using limitations in bacteriophage Francis Crick and his

colleagues confirmed in 1961 that the genetic code is a triplet code in which three nucleotides encode each amino acid in a protein. Proteins are very important to all living processes; they are in some cases enzymes, which are the biological catalyst and conduct the chemical reactions of the cell. In other cases proteins are structural components of a cell like muscle, nail, lens in the eye. Some proteins help in the function of transporting substances others help in the regulation of vital pathways, communication between cell to cell or in defending a cell from external threat.

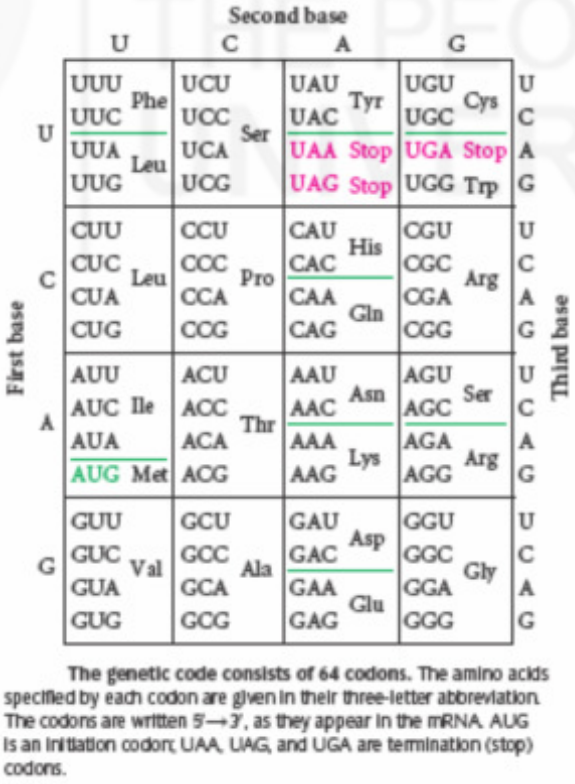
Every protein is composed of amino acids, aligned end to end. Twenty common amino acids are known to constitute different proteins. Each amino acid has a central carbon atom bonded to an amino group ($NH_3^+$), a hydrogen atom ,a carboxyl group ($COO^-$)and an R (Radical) group. The Radical group differs for each amino acid. The amino acids in proteins are joined together by peptide bonds.

Thus it is very important for you to remember some important concepts like

1) The products of many genes are proteins whose action produce the traits encoded by these genes. Proteins are polymers consisting of amino acids linked by peptide bonds. The amino acid sequence of a protein is its primary structure.

2) This structure folds to create the secondary and tertiary structures; two or more polypeptide chains may associate to create a quaternary structure.

3) The genetic code is a triplet code in which three nucleotides encode each amino acid in a protein.




The genetic code consists of 64 codons. The amino acids specified by each codon are given in their three-letter abbreviation. The codons are written 5′→3′, as they appear in the mRNA. AUG is an initiation codon; UAA, UAG, and UGA are termination (stop) codons.

**Fig. 1.5: Genetic Code (Source: Benjamin A. Pierce Genetics: A Conceptual Approach 4th Edition 2012)**

Because we now know that some enzymes contain polypeptide chains encoded by two (or occasionally more) different genes, a more accurate statement of the principle is "one gene, one polypeptide".

### 1.1.5 Revisiting the Question about what is a Gene?

After studying the developments of Genetics as it unfolded above you may be able to appreciate in a nut shell that:

i) The Mendelian concept of a gene views it as a discrete unit of inheritance that affects phenotype.

ii) Morgan and his colleagues assigned genes to specific loci on chromosomes.

iii) We can also view gene as a specific type nucleotide sequence along a region of a DNA molecule as put forth by Watson and Crick.

iv) We can define a gene functionally as a DNA sequence that codes for a specific polypeptide chain.

Even the one gene one polypeptide definition must be refined and applied selectively, because:

i) Most eukaryotic genes contain large introns that have no corresponding segments in polypeptides.

ii) Promoters and other regulatory regions of DNA are not transcribed either, but they must be present for transcription to occur.

iii) Our definition must also include the various types of RNA that are not translated into polypeptides.

Hence a gene is a region of DNA whose final product is either a polypeptide or an RNA molecule.

## 1.2 BIOCHEMICAL OR METABOLIC DISEASES

Now we will move on to consider single gene biochemical diseases. The spectrum of disorders known under this category is vast (200 inborn errors of metabolism are known) and a flavor of this fascinating branch of genetic medicine will be given. The metabolic disorders can be grouped based on either the metabolite, metabolic pathway, function of the enzymes or cellular organelle involved.

Only a few IEM are inherited in an autosomal dominant manner. This is because the defective protein in the most inborn errors is an enzyme which is diffusible, and there is usually sufficient residual activity in the heterozygous state for the enzyme to function normally in most situations. If, however, the reaction catalyzed by an enzyme, is rate limiting or the gene product is part of a multimeric complex, the disorder can manifest in the heterozygous state, i.e., be dominantly inherited.

### 1.2.1 Disorders of Amino Acid Metabolism

The best known under this category are phenylketonuria (PKU), alkaptonuria, occulocutaneous albinism, homocystinuria, maple syrup urine disease.

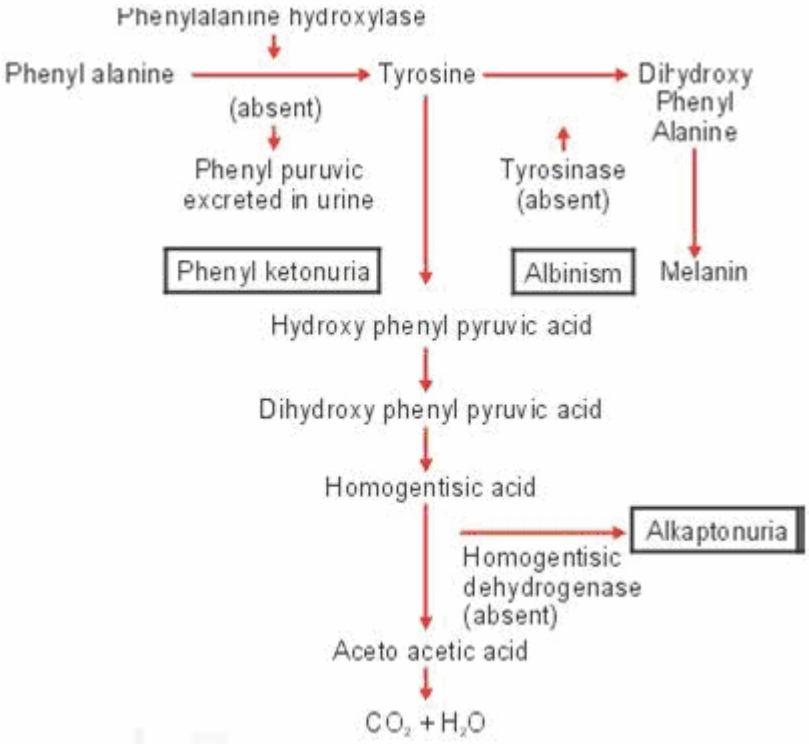



**Fig. 1.6: Sites of biochemical block in phenylketonuria, alkaptonuria and albinism (Source:transtutors.com)**

**Fig. 1.7: A boy with albinism (Source: human-albinism-eye-color.seebyseeing.net)**

**Fig. 1.8: A photograph of siblings born with Phenylketonuria (PKU) (Source:drustapbio.wikia.com).**

The untreated 11 year old boy (left) is severely retarded. His two and half year old sister (right) who was treated from early infacy with a low phenylalanine diet, has normal intelligence. Phenylketonuria can lead to severe brain damage and mental retardation.

### 1.2.2 Disorders of Branched Chain Amino Acid Metabolism

The essential branched chain amino acids leucine, isoleucine and valine have a part of their metabolic pathways in common like the previous set mentioned above. Maple syrup urine disease comes under this category. It is an autosomal recessive disorder that presents with vomiting in the first week of life followed by death within a few weeks. There is a specific odour of the urine similar to the maple syrup. This results in the excretion of increased branched chain amino acids leucine, isoleucine and valine in the urine. Treatment involves a diet that limits the three branched chain amino acids to the extent that it is necessary only to help the new born to grow. Affected individuals are susceptible to deterioration due to catabolic degradation of proteins.

### 1.2.3 Urea Cycle Disorders

Urea cycle helps to remove waste nitrogens from the amino groups of amino acids arising from the day to day production of proteins. The sight of action is mainly liver cells and the cycle has a five step metabolic pathway. Deficiencies of enzyme in the urea cycle will result in intolerance of protein due to the accumulation of ammonia in the body. This is also known as hyperammonemia. In normal individuals without enzyme deficiency it converts two molecules of ammonia and one bicarbonate into urea. Excess of ammonia levels are toxic to the central nervous system and it can also result in coma and death when untreated. These disorders are fortunately rare.

### 1.2.4 Disorders of Carbohydrate Metabolism

Examples are Galactosemia, hereditary fructose intolerance, glycogen storage diseases of liver and muscle.

### 1.2.5 Disorders of Steroid Metabolism

This is an inborn error in the biosynthetic pathways of cortisol. Due to this disorder new born female infants present with virilization of the external gentalia that is the commonest cause of ambiguous gentalia in female new borns. Affected infants in addition to requiring urgent correct assignment of gender are treated with replacement cortisol along with fludrocortisone if they have salt losing form of defect.

### 1.2.6 Disorders of Lipid Metabolism

Familial hypercholesterolemia where the person will have raised cholesterol level with high risk of developing coronary artery disease. The high cholesterol level is due to deficient or defective function of LDL receptors leading to increased levels of endogenous cholesterol synthesis.

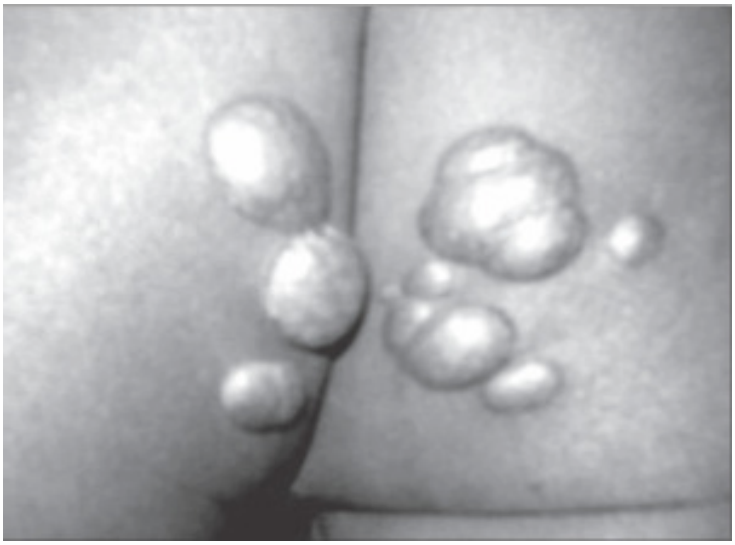


**Fig. 1.9: Familial hypercholesterolemia with multiple cutaneous xanthomas (Source: gfmer.ch)**

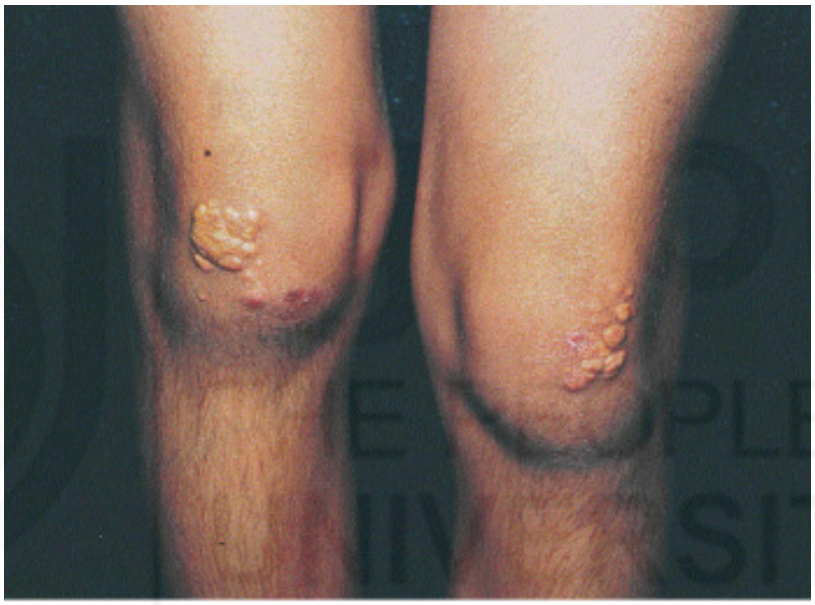


**Fig. 1.10: Cutaneous Xanthomas in Homozygous Familial Hypercholesterolemia (Source:nejm.org)**

### 1.2.7 Lysosomal Storage Disorders

While the previous six types of inborn errors of metabolism were due to enzyme defects leading to accumulation of intermediate metabolic precursors, there are other metabolic disorders due to accumulation of very large polysaccharides. Certain macro molecules which are complex need to be degraded consistently in the cell.

Children born with lysosomal disease are apparently normal at birth but tend to accumulate a variety of macro molecules due to a deficiency of lysosomal enzymes. Examples are Hurler Syndrome, Hunter Syndrome, Sanfilippo Syndrome, Morquio syndrome, all put together known as Mucopolysaccharidoses (MPS).

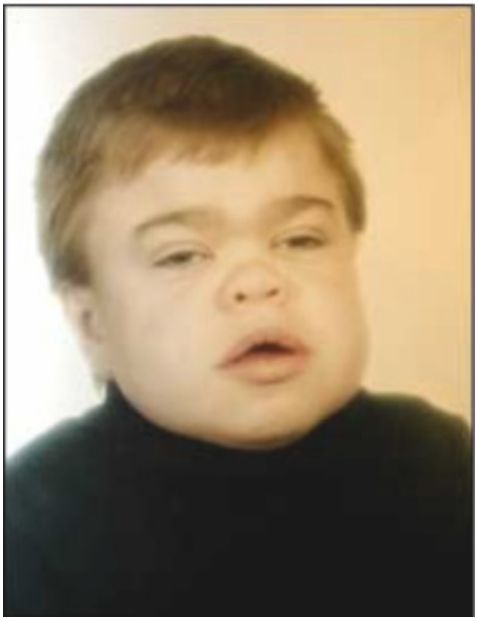


**Fig. 1.11: Hunter Syndrome (Source:thesun.co.uk)**

Face of a male with mucopolysaccharidosis. Note the characteristic coarsening of their facial features.

### 1.2.8 Lipid Storage Diseases

Tay-Sachs disease is the best known lipid storage diseases where in there is a progressive deposition of lipid or glycolipid primarily in the liver and spleen. Central nervous system involvement results in progressive mental deterioration, deafness, visual impairment, spasticity and progressive rigidity . 1 in 3600 persons of Ashkenazi Jewish ancestry are known to be affected with this disorder.

### 1.2.9 Disorders of Purine/Pyramidine Metabolism

Lesch-Nyham Syndrome is a good example. Accumulation of excessive amounts of uric acid and some of its metabolic precursors. The main effect on the central nervous system resulting in uncontrolled movements, spasticity, mental retardation and compulsive self- mutilation.

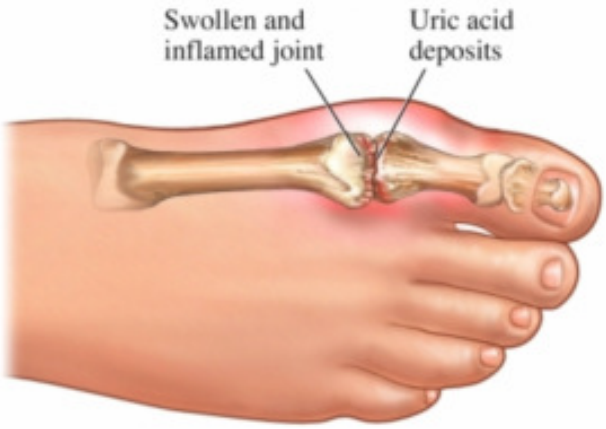



**Fig. 1.12: Lesch-Nyhan Syndrome (Source:EmpowHER.com)**

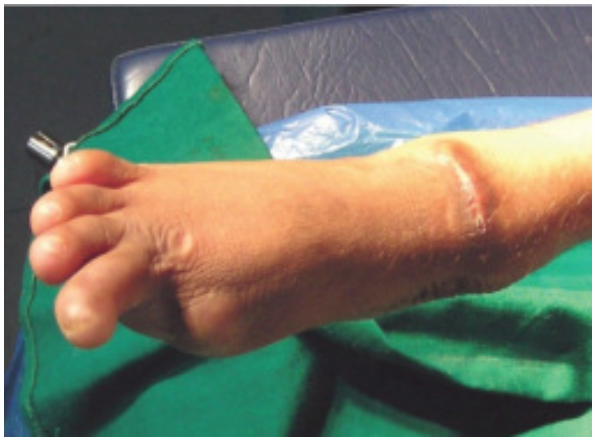


**Fig. 1.13: Lesch-Nyhan syndrome: A case report (Source:gfmer.ch)**

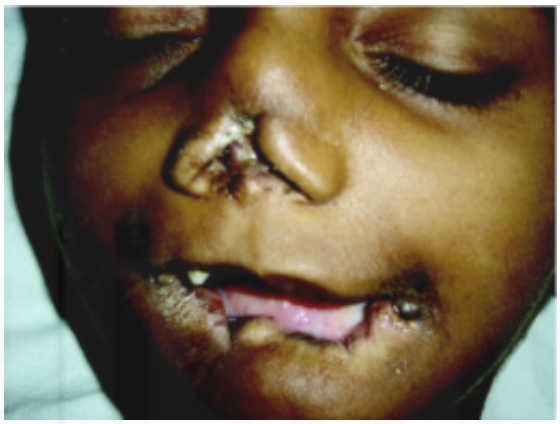


**Fig. 1.14: Self-mutilation in the Lesch–Nyhan syndrome (Source:Neurology.org)**

### 1.2.10 Disorders of Copper Metabolism

Good examples are Menkes Disease and Wilson Disease.

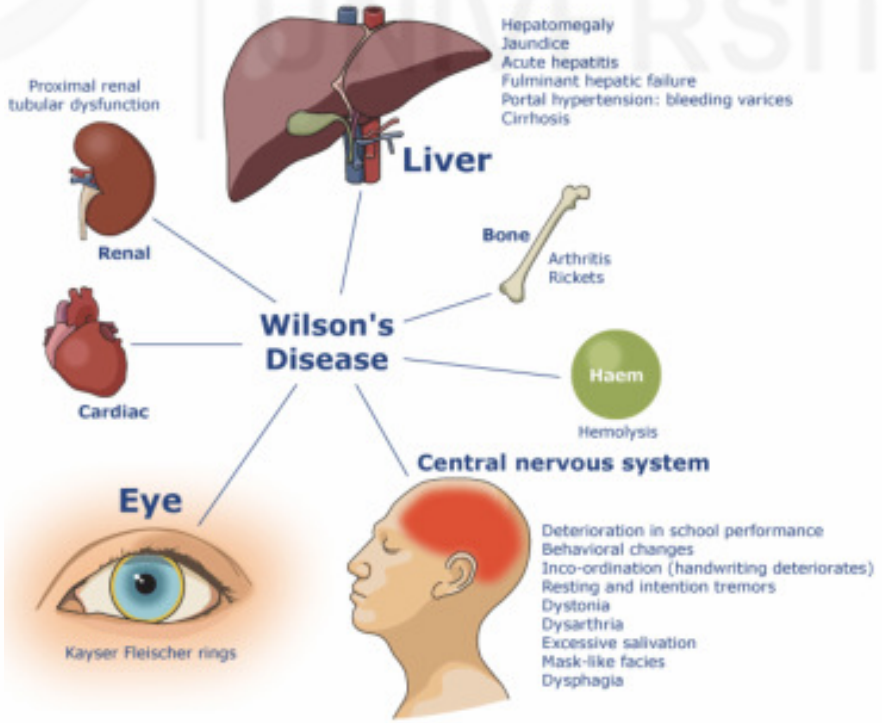



**Fig. 1.15: Wilson's disease (Source:eurowilson.org)**

# 1.3 HUMAN BLOOD GROUP SUBSTANCES

## 1.3.1 Introduction to Human Blood Groups

The red blood cells of human blood contain a great many antigens mostly proteins and according to the presence of antigens, human blood can be classified into different blood group systems, e.g ABO blood group, MN blood group, RH blood group. All of these blood groups in man are under genetic control, each blood group system being under the genetic control of genes at a single locus or of genes that are closely linked and behave in heredity as though they were in single locus.

## 1.3.2 The ABO Blood Group System- History

Blood has two main components: Serum and cells. Karl Landsteiner noted that the sera of some individuals caused the red cells of others to agglutinate. This observation led to the discovery of the ABO blood group systems. Based on the reactions between the red blood cells and the sera, he was able to divide individuals into three groups: A, B, and O. Two years later, two of his students discovered the fourth and rarest type, namely AB. Landsteiner described A, B, and O groups; Alfred von decastello and Adriano sturli discovered the fourth type, AB, in 1902.

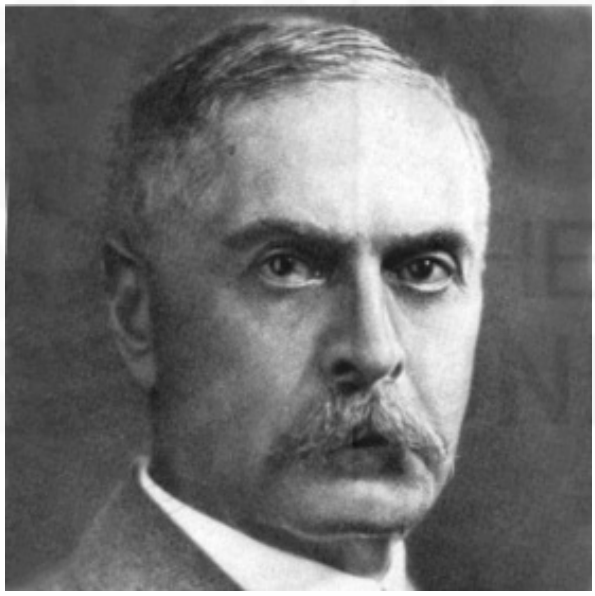


**Fig. 1.16: Karl Landsteiner (Source: http://www.biography.com/people/karl-landsteiner-9372736)**

Karl Landsteiner was born June 14, 1868 in Vienna, Austria. In 1897 he pursued his interest in the emerging field of immunology and in 1901 published his discovery of the human ABO blood group system. He won the 1930 Nobel Prize for Physiology or Medicine for his discovery of the major blood groups and the development of the ABO system of blood typing that has enabled blood transfusions.

Due to less developed communication system in those days, it was subsequently found that Czech serologist Jan Jansky had also independently pioneered the classification of human blood into four groups, although his name is not so much talked about, except in Russia and former states of USSR around that time in America, Moss published a similar work on blood group in 1910, from USA.

Ludwik Hirszfeld and von Dungern discovered the heritability of ABO blood groups in 1910–11. Felix Bernstein earns the credit for studying the blood group inheritance pattern to be due to, multiple alleles at one locus in 1924.

### 1.3.3 ABO Blood Groups Substances

Watkins and Morgan, in England, discovered that the ABO epitopes were conferred by sugars, to be specific:

i) N-acetylgalactosamine for the A-type

ii) Galactose for the B-type.

Published literature states that the ABH substances are all attached to glycosphingolipids – a long polylactosamine chain that contains the major portion of the ABH substances attached to it.

Later, Yamamoto's group showed the precise glycosyl transferase designates the A, B and O epitopes. Epitope is the antigenic determinant on an antigen to which the paratope (the specific site in the immunoglobulin) on an antibody binds.

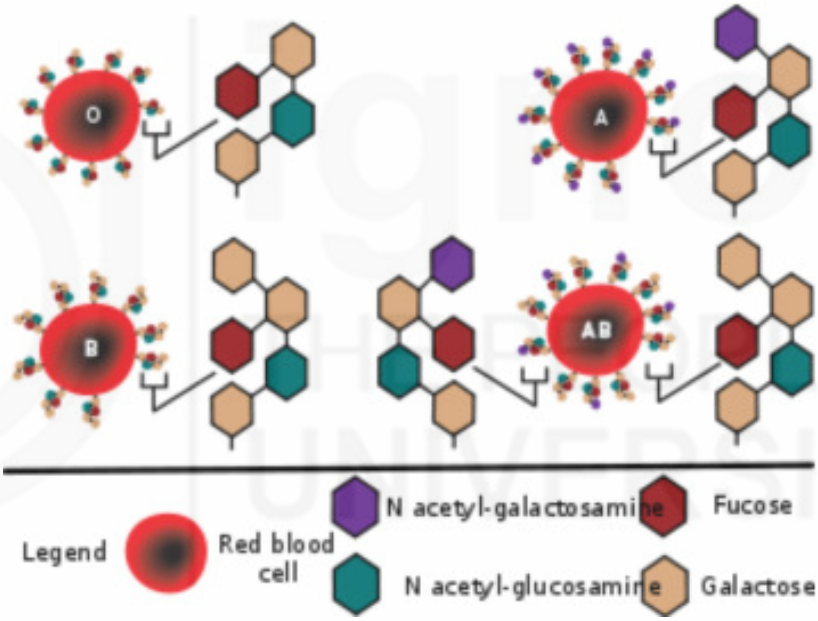



**Fig. 1.17: Diagram showing the carbohydrate chains that determine the ABO blood group (Source: http://en.wikipedia.org/wiki/ABO_blood_group_system)**

The genetic basis of the ABO blood group system is an example of multiple alleles. There are three alleles, A, B, and O, at the ABO locus on chromosome 9. The expression of the O allele is recessive to that of A and B, which are said to be co-dominant. Thus, the genotypes AO and AA express blood type A, BO and BB express blood type B, AB expresses blood type AB, and OO expresses blood type O. In the past, ABO blood group typing was used extensively both in forensic cases as well as for paternity testing.

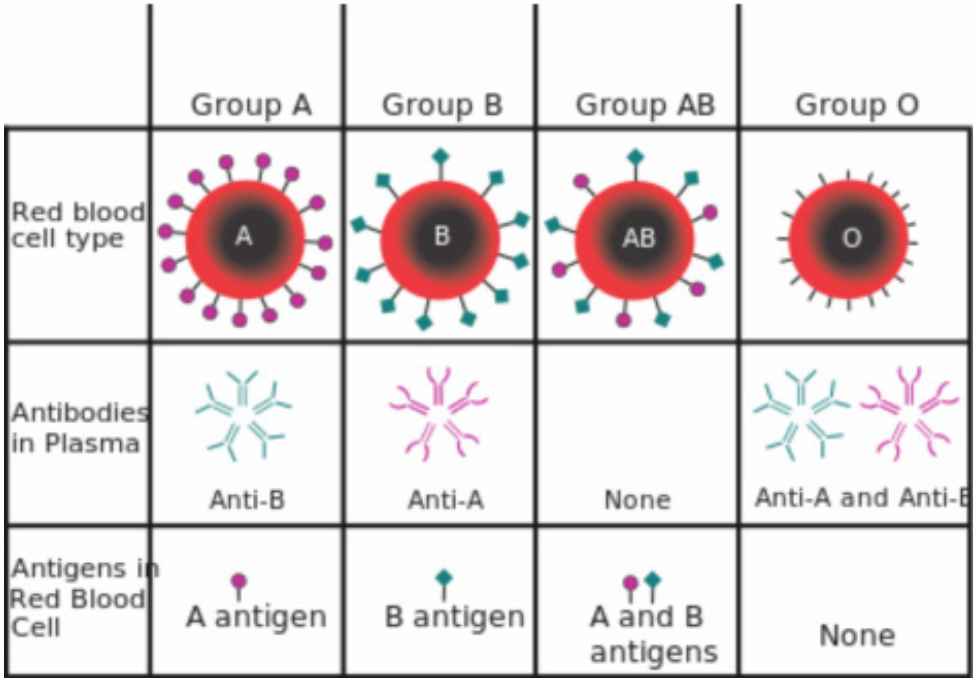



**Fig. 1.18: ABO blood group antigens present on red bllod cells and IgM antibodies present in the serum (Source: http://en.wikipedia.org/wiki/ABO_blood_group_system)**

The ABO blood group system is an important blood group system to be considered during human blood transfusion. The associated anti-A antibodies and anti-B antibodies are usually IgM antibodies, which are usually produced in the first years of life by sensitization to environmental substances such as food, bacteria, and viruses. ABO blood types are also present in apes such as chimpanzees, bonobos, and gorillas.

Thus ABO group substances are distinct, genetically determined group of human erythrocyte antigens represented by two blood factors (A and B) and four blood types (A, B, AB, and O).

### 1.3.4 Antigens and Antibodies

Blood grouping, is essentially based on the definition of antigen and antibody. An antigen is a substance, usually a protein or a glycoprotein**,** when injected into a human (or other organism) that does not have the antigen, will cause antibody to be produced. Antibodies are a specific type of immune-system proteins known as immunoglobulins, whose role is to fight infections by binding themselves to antigens. In the case of the ABO blood groups, the antigens are present on the surface of the red blood cell, while the antibodies are in the serum. These antibodies are unique to the ABO system and are termed "naturally occurring antibodies." The table shows the relationships between blood types and antibodies.

**Relationships Between Blood Types And Antibodies**

| Blood Type | Antigens on Red Blood Cell | Can Donate Blood To | Antibodies in Serum | Can Receive Blood From |
|---|---|---|---|---|
| A | A | A, AB | Anti-B | A, O |
| B | B | B, AB | Anti-A | B, O |
| AB | A and B | AB | None | AB, O |
| O | None | A, B, AB, O | Anti-A and anti-B | O |

ABO blood group-matching is very important in transfusion. Blood group O individuals are said to be universal donors, because their blood can be used for transfusion in individuals who have any one of the four blood types. On the other hand, individuals with blood type A can only donate to either type A or type AB, and individuals with blood type B can only donate to B or AB types. AB individuals can only donate to type AB. However, before any transfusions, donor blood is mixed with serum from the recipient (a process called cross matching) to ensure that no agglutination will occur after transfusion.

## 1.4 SUMMARY

The scope for Genetics is amazing and most of the sub disciplines emerged as the understanding of Genetics nurtured the development of the science. Human biochemical genetics was the first ideal model to pave way for this development. Improvements in genetic technologies have helped us now stand on a platform where we have a detailed knowledge of the human genome and its variations.

### Suggested Reading

Benjamin, A. Pierce. 2008. *Genetics: A Conceptual Approach*. 3rd Edition W.H. Freeman & Company.

Elaine Johansen Mange and Arthur P.Mange. 1999. *Basic Human Genetics*. 2nd Edition. Sunderland Publishers.

Mandal, S. 2002. *Fundamentals of Human Genetics*. 2nd Edition. New Central Book Agency.

Peter, D Turnpenny and SianEllard. 2007. *Emery's Elements of Medical Genetics*. 13th Edition. Elsevier Publications.

Robert, C. King and William D. StansField. 2002. *A Dictionary of Genetics*. 6th Edition. Oxford University Press.

Scriver, CR. 2001. *Work, the clinician-scientist and human biochemical genetics*. *Clin Invest Med,* Vol. 24, No. 4.

Thomas, D.Gelehrter and Francis S.Collins. 1999. *Principles of Medical Genetics*. 3rd Edition. Williams and Wilkins.

### Sample Questions

1) Describe with example some classic examples of inborn errors of metabolism.
2) Discuss the concept of gene in the light of Beadle and Tautum's work.
3) Define the basis of human ABO blood group systems.

# UNIT 2 ENZYME AND PROTEIN DIVERSITY IN HUMAN POPULATIONS

**Contents**



**Learning Objectives**

After reading this unit, you will be able to:

- define genes and isozymes and explain the concept of genetic polymorphism of red cell enzymes and proteins in man;
- discuss with reference to normal and abnormal haemoglobins and their structural variation; and
- explain the geographical distribution of some abnormal haemoglobins in Indian populations.

## 2.1 INTRODUCTION

Human beings are exceedingly diverse. They differ from one another in their normal physical, physiological and behavioural attributes. These variations are caused partially by differences in the environmental conditions in which they live but more importantly they also depend on inborn (genetic) differences. Human variation can be visible (e.g., differences in skin colour, hair colour and form and or head shape) or invisible (biochemical differences, e.g., blood groups, blood protein/red cell enzyme polymorphisms or DNA markers). At the beginning of the 20$^{th}$ century, Landsteiner discovered ABO blood groups and Hirschfeld and Hirschfeld (1919) suggested that these could be used to delineate biochemical races. Blood protein haemoglobin polymorphism, including the gene for sickle cell anaemia, was reported by the middle of the century, followed by serum protein haptoglobin (HP) polymorphism in mid-1950's and by mid-1970's most red cell enzyme and blood protein polymorphisms were discovered. Anthropologists had studied these genetic markers with the primary aim of documenting genetic differences among various populations inhabiting different parts of the world and also for human racial classifications. Thus the existence of Mendelian genetic traits, demonstrated from human blood in the 20$^{th}$ century,

provided important powerful tools for investigation of biological variation in humans along with traditional anthropometric/morphological and dermatoglyphic traits, among others.

## 2.2 GENES AND ISOZYMES

A gene can be defined as a unit of information and corresponds to a discrete segment of DNA (exon) that encodes the amino acid sequence of a polypeptide. Human cells contain about 25,000 genes which are dispersed on chromosomes and are separated by noncoding inter-genic DNA (introns). Isozymes result from point mutations or from insertion-deletion (indel) events that affect the DNA coding sequence of the gene.

The term isozyme was coined by Hunter and Markert (1957) who defined isozymes as different variants of the same enzyme having identical functions or present in the different individuals. In fact, isozymes refer to the existence of different molecular forms of an enzyme that catalyze the same biochemical reaction but which differ in their electrophoretic properties. The existence of isozymes permits the fine-tuning of metabolism to meet the particular needs of a given tissue or developmental stage.

## 2.3 GENETIC VARIATION OF RED CELL (ERYTHROCYTE) ENZYMES AND SERUM PROTEINS

A variety of different enzymes and proteins are synthesized in the human body, and the primary amino acid sequence of each of their distinctive polypeptide chains is coded in the DNA of a separate gene locus. Blood proteins, including red cell (erythrocyte) enzymes are composed of amino acids joined by covalent peptide bonds to form polypeptides. The sequences or "primary structures" are genetically determined. Each of the 20 amino acids has a unique side chain, characterized by its shape, size and charge. The side chains on lysine, arginine and histidine are positively charged ($NH_3^+$) and thus basic; those on aspartic acid and glutamic acid are negatively charged ($COO^-$) and thus acidic.

The charged side chains are responsible for the movement of the proteins or enzymes through a gel matrix during electrophoresis. In an electric field, the anions (i.e. negatively charged molecules) migrate towards the anode (i.e. positive electrode), and the cations (i.e. positively charged molecules) migrate towards the cathode (i.e. negative electrode). The speed of migration is primarily related to net charge on the protein or enzyme molecule and the electrical field strength applied through the electrodes. Thus proteins/enzymes, the main products of genes, move in an electric field with mobility that depends on their chemical structure. The amino acid sequences of proteins are changed by mutations in the encoding DNA locus. Such mutations may alter shape and net charge and electrophoresis aims to reveal as many of these changes as possible.

The electrophoretic enzyme pattern in homozygous individuals usually show one major zone of enzyme activity, which may be accompanied by minor zones of secondary isozymes (Shaw, 1969). Such pattern in heterozygous individuals may, however, show different degrees of complexity i.e. it shows one or more

components which are not present in either homozygote. These extra enzyme components found only in heterozygotes are referred to as "hybrid enzymes" and show an electrophoretic mobility intermediate to those of the parental homozygote enzymes. In a heterozygote, in the case of a dimer enzyme there will be a triple banded pattern, in the case of a trimer enzyme a four banded pattern and in the case of a tetramer enzyme a five banded pattern. In each case the outermost bands correspond to the two parental homozygotes, and the band(s) with intermediate mobilities represent the "hybrid enzyme(s)".

Electrophoretic investigations on a variety of different enzymes and proteins have led to the discovery of a large number of structurally variant forms which are genetically controlled. The inherited variants of enzymes/proteins which we find in human populations today must be attributed to specific gene mutations which occurred in single individuals among our ancestors in earlier generations. Thus it appears that at such loci a large number of different mutant alleles which may be generated by separate mutations, actually exist among living members of our species. The incidence of the majority of such mutant alleles is rather low in human populations. Occasionally, some occur with a polymorphic frequency (=1%) giving rise to the well-known phenomenon of genetic polymorphism in which the individual members of a population are sharply classified into two or more relatively common genetically determined phenotypes due to occurrence together of (usually) two or more alleles at a particular locus. The first example of electrophoretically detectable biochemical polymorphism in humans was that of the well-known blood protein haemoglobin (HB) (Pauling et al., 1949).

In man, both proteins (present in serum portion of the blood) and enzymes (present within the red blood cells/erythrocytes) have been screened for polymorphisms. Generally, the discovery of polymorphisms of serum protein preceded that of red cell enzymes in man. The first to be discovered was haptoglobin (HP) reported by Smithies (1955), followed by Transferrin (TF) (Poulik and Smithies, 1958), group specific component (GC) (Hirschfeld, 1959), complement component 3 (C3) (Wieme and Demeulenaere, 1967) and properdin factor B (BF) (Alper et al., 1971), among others.

Using the technique of starch gel electrophoresis pioneered by Smithies (1955) and through the use of very sensitive biochemical staining reactions (Hunter and Markert, 1957), the polymorphisms of red cell enzymes began to be reported from the early 1960's. The first example was that of acid phosphatase locus 1 (ACP1) discovered by Hopkinson et al. (1963), followed by phosphoglucomutase locus 1 (PGM1) (Spencer et al., 1964), adenylate kinase locus 1 (AK1) (Fildes and Harris, 1966), adenosine deaminase (ADA) (Spencer et al., 1968), phosphohexose isomerase (PHI) (Detter et al., 1968), esterase D (ESD) (Hopkinson et al., 1973) and glyoxalase locus 1 (GLO1) (Kompf et al., 1975), among others. Harris et al. (1977) observed that of the 104 different loci coding for enzymes that had been investigated, 33 were found to exhibit an electrophoretically detectable polymorphism. In other words, one out of every three human loci coding for enzymes is polymorphic. It is important to mention that although an electrophoretic difference indicates a difference in structure between the polymorphic forms of an enzyme/protein, it does not in general provide information about possible functional difference, if any.

Like different blood groups which are serological markers, various red cell enzyme and serum protein polymorphisms are biochemical markers, and alongwith with

the former help in characterization of human populations genetically. A large body of data on serum protein and red cell enzyme polymorphisms has been generated in world populations and these have been compiled and authoritative accounts written on their distribution (Mourant et al., 1976). For the people of India, such an exercise was performed by Bhasin et al. (1992, 1994a) and a brief account of the distribution of different biochemical genetic markers in Indian populations is given in the following.

In serum protein haptoglobin polymorphism, the average frequency of *HP*1* is 0.160 in Indian populations; in general, the frequency of this allele is higher in people of North India (0.208) than in South India (0.131). The frequency is also high in people of West India (0.215), followed by East India (0.192).

In transferrin polymorphism, the average frequency of *TF*C* is 0.991 in Indian populations with a range of 0.898 to 1. The frequencies of other alleles of this serum protein polymorphism such as *TF*D* and *TF*B* are rather meager. The frequency of *TF*D* is highest in peninsular (South) Indian populations and it starts decreasing in the Indus-Ganga-Brahmaputra plains of North India (0.005).

In serum protein group specific component polymorphism, the average frequency of *GC*2* in people of India is 0.253 with a range of 0.089 to 0.409. The frequency of the allele is low in West India (0.224) and high in North India (0.277). In the Himalayas, the frequency is almost similar in Eastern and Western regions (0.244 and 0.248, respectively) while in Central Himalayan region it is comparatively higher (0.312).

In acid phosphatase locus 1 system, like other world populations, in Indian populations the *ACP1*B* allele is preponderant (0.756), followed by *ACP1*A* (0.242) while the third allele of the system *ACP1*C* has a trace frequency (0.002). The average frequency of *ACP1*A* allele in people of India is 0.242, ranging from 0.035 to 0.467. The distribution of this red cell enzyme polymorphism in five geographical regions of the country i.e. North, West, Central, East and South India was studied by Chahal et al. (1985) who observed that the populations of North India are characterized by relatively high frequencies of *ACP1*A* and *ACP1*C,* which differentiate them from populations of all other regions. It was also observed that the incidence of both these alleles is generally higher in the non-tribal populations compared to the tribals of India. Indeed, as stated by Roberts et al. (1980) "the distribution of *ACP1*C* is essentially non-tribal in India."

In red cell enzyme phosphoglucomutase locus 1 polymorphism, the average frequency of *PGM1*2* is 0.300 in Indian populations, varying from 0.05 in Chaudhuri of West India to 0.558 in Kurumba of South India. In populations with Mongoloid affinities inhabiting Eastern Himalayas the frequency of the allele is somewhat higher (0.309) compared to those of South India (0.298). In general, the frequency of *PGM1*2* is low in North India and it gradually starts increasing in West and East India while it decreases in East, Central and South India.

In adenylate kinase locus 1 polymorphism, the average frequency of *AK1*2* in Indian populations is is 0.076 with a range of nil to 0.205. The frequency of the allele is low in the people of Himalayas having varying degrees of Mongoloid

admixture (0.058) but in populations of the peninsular (South) India it is comparatively high (0.082). Chahal et al. (1986a) studied the distribution pattern of this red cell enzyme polymorphism in the five geographical regions of India and found that there is a clear distinction between non-tribal (range 0.086-0.099) and tribal (range 0.042-0.064) populations inhabiting these regions. Thus tribal populations of India are characterized by having comparatively much lower *AK1*2* frequency.

In red cell enzyme adenosine deaminase system, the average frequency of *ADA*2* is 0.118 in various populations of India, varying from 0.015 in Jalari of Andhra Pradesh to 0.5 in Muslims and Kacharis of Assam. However, most of the Indian populations fall in a range of 0.015 – 0.214 for this allele. The frequency of the allele is quite low in the populations with Mongoloid ethnicity inhabiting Himalayas (0.087), followed by the peninsular populations of South India with Australoid (Pre-Dravidian) and Caucasoid (Dravidian) racial affinities compared to people of Indus-Ganga-Brahmaputra plains of North India with Caucasoid (Aryan) affinities (0.151).

In esterase D system, the average frequency of *ESD*2* in India is 0.271, ranging from as low as 0.022 in Gaddi of Kangra in Himachal Pradesh to 0.582 in people of Andhra Pradesh. The frequency is low in people of Himalayas (0.253) and Indus-Ganga-Brahmaputra plains (0.256) while it is comparatively higher in that of peninsular India (0.322). In fact, there exists a north-south cline of increasing *ESD*2* frequency in this red cell enzyme polymorphism in India (Chahal et al., 1986b).

The data on red cell enzyme glyoxalase I polymorphism in Indian populations are rather limited, especially among tribals. Apparently there are no great differences in *GLO1*1* frequency in them. The frequency of *GLO1*1* ranges from about 0.2 to 0.3 in most of them. Chahal et al. (1986) found somewhat elevated frequenciy of the allele in people of West India and attributed it mainly to the inclusion of two immigrant populations of the Parsi and Irani.

In red cell enzyme glucose phosphate isomerase (phosphohexose isomerase) system the frequency of the *GPI*1* allele is unity in most world populations, except Asiatic Indians and Japanese in which rare alleles, respectively, *GPI*3 and GPI*4* are present. Various Indian populations have been screened for variants of GPI (Papiha and Chahal, 1984) and the rare allele *GPI*3* attain polymorphic proportions in many of them. With a frequency as high as 0.110 of the allele, the Bhotia of Chamoli district in Garhwal region of Uttarakhand stand out from all populations of India (Chahal et al., 2008). In addition to *GPI*3,* other variant alleles reported from the country include *GPI*2, GPI*4, GPI*5, GPI*7, GPI*8 and GPI*9*, albeit some of them are limited to specific ethnic groups or geographical regions.

## 2.4 HAEMOGLOBINS: NORMAL AND ABNORMAL HAEMOGLOBINS

Haemoglobin (HB), the red respiratory protein found in mammalian erythrocytes, is one of the most informative molecules in primate blood. It is the best known blood protein that gives rise to the most thoroughly studied genetic polymorphism

in man – the polymorphism that includes the gene for sickle cell anaemia (Cavalli-Sforza and Bodmer, 1971). The study of haemoglobin has made several significant contributions towards the development of molecular biology.

Normal adult human haemoglobin is composed of three different types which can be separated by electrophoresis. They are all tetramers having the general formula $X_2Y_2$, where $X_2$ refers to a pair of α polypeptide chains and $Y_2$ to a pair of β, γ, δ or ε chains (Giblett, 1969). For example, HBA, the major haemoglobin in the blood of adults, has the molecular formula $\alpha^{A}{}_{2}\beta^{A}{}_{2}$. The sequence of the 141 amino acid residues in the α chain and of the 146 residues in the β and other non-α chains has been determined. While most of the haemoglobin in normal human subjects past infancy is HBA (95-98%), HBA2 is also present, but in a concentration of only 1.5-3% of the total haemoglobin content. Its molecular formula is $\alpha^{A}{}_{2}\delta_{2}$, indicating that the two α chains are identical with those of HBA, but that the other two identical chains are sufficiently different from β chains to warrant the designation δ. In fact, the δ chain has the same number of amino acid residues as the β chain and there are only 10 differences in the sequence. Another example of normal haemoglobin is HBF. It has a concentration of less than 0.5% after the first few years of life in normal subjects, but it is the major haemoglobin component during fetal development. Like HBA and HBA2, HBF contains two $\alpha^{A}$ chains, but a different pair of chains, called γ, completes the tetramer, so the molecular formula is $\alpha^{A}{}_{2}\gamma^{F}{}_{2}$. The γ chain differs from the β chain in 39 amino acid residues, although the total number of residues is the same.

The abnormal haemoglobins differ from normal haemoglobins in molecular structure. They have structure similar to that of normal haemoglobin except slight alteration in the sequence of amino acids (usually in the β chain) and hence may be designated as mutant or variant haemoglobins. Most of the abnormal haemoglobins appear to be products of point mutations with a single amino acid substitution in the α, β, γ or δ peptide chain. The resultant change in the whole haemoglobin molecule may be so benign that no physiological effect is detectable and even the rate of synthesis remains normal. A large number of the apparently 'benign' abnormal haemoglobins have so far been detected in humans. In some instances, these variants have been observed in combination with thalassaemia, which depresses the synthesis of either α or β chain product of the homologous locus. Detection of structurally abnormal haemoglobin is usually demonstrated by electrophoresis.

The term haemoglobinopathies covers a group of hereditary abnormalities in which either (a) the haemoglobin structure is altered (like in HBS, HBC, HBD, HBE etc.) or (b) there is a defect in globin chain synthesis of one of the normal haemoglobin chains (i.e. α or β) (thalassaemias). The term thallasaemia is applied to a group of inherited disorders in which there is a variable decrease in net synthesis of a particular globin chain without an associated change in the structure of that chain. There are two principal types of thalassaemia namely alpha (α) and beta (β). β thalassaemia is of two subtypes viz., β-thalassaemia major (homozygous) and β-thalassaemia minor (heterozygous). α thalassaemia is present in two forms – haemoglobin Barts (four γ chains, no α chain produced) and haemoglobin H (excess β chains form unstable tetramer). Haemoglobinopathies are the most frequent genetic disease, affecting approximately 7 per cent of the world population.

## 2.5 STRUCTURAL VARIATION IN HAEMOGLOBIN (HAEMOGLOBIN VARIANTS)

Variant haemoglobins such as HBS, HBC, HBD and HBE, among others, are the products of point mutation and their detection is dependent on a difference in their altered electrophoretic mobility. Worldwide more than 470 genetically controlled haemoglobin variants, mostly of the β chain, are known (Honig and Adams III, 1986). Haemoglobin variants may occur independently or in association with thalassaemia.

### Haemoglobin S (HBS)

One of the most interesting and widespread abnormal human haemoglobin, first described by Pauling et al. (1949), is designated haemoglobin S (HBS)($\alpha^{A}_{2}\beta^{S}_{2}$), more specifically, $\alpha^{A}_{2}\beta_{2}^{6Glu?Val}$. The notation signifies that HBS differs from its normal counterpart HBA by a single amino acid substitution i.e. valine replaces glutamic acid at position six in the β polypeptide chain (Ingram, 1957). It may be noted that the mutation has only affected the β chain while the α chain is normal. Haemoglobin S may be separated from haemoglobin A by electrophoresis or by performing sickling test on fresh blood. Individuals who have one sickle mutant gene and one normal beta gene i.e. heterozygote (HBAS) are referred as sickle cell trait, which is benign. The sickling of red cells of homozygote (HBSS) is more severe than that of heterozygote (HBAS). Haemoglobin S is also known to occur in combination with other abnormal haemoglobins, for example, sickle cell HBC (HBS/HBC), sickle cell HBD (HBS/HBD), sickle cell HBE (HBS/HBE) and sickle cell thalassaemia (HBS/HBTh), among others.

Sickle cell trait occurs with highest frequency in tropical Africa (0.1-0.4), with high frequency in India, Greece and Southern Turkey (0.05-0.1) and less than 0.1 frequency in populations inhabiting Palestine, Tunisia, Algeria and Sicily, situated around the Mediterranean Sea.

### Haemoglobin D (HBD)

Like HBS, HBD (also known as HBD-Punjab) is a β chain variant of haemoglobin in which glutamine replaces glutamic acid at position 121 in the β polypeptide chain ($\alpha^{A}_{2}\beta_{2}^{121Glu?Gln}$). Both homozygote (HBDD) and heterozygote (HBAD) phenotypes have been reported among individuals with reasonable health. Haemoglobin D occurs mainly in North-West India, Pakistan and Iran. The electrophoretic mobility of HBAD is identical to that of HBAS at alkaline pH in cellulose acetate, but to distinguish these variants separation may be carried out in agarose gel using a different buffer system.

### Haemoglobin E (HBE)

Like HBS and HBD, HBE is also an abnormal haemoglobin with a single point mutation in the β polypeptide chain i.e. at position 26 in the chain there is a change of one amino acid from glutamic acid to lysine ($\alpha^{A}_{2}\beta_{2}^{26Glu?Lys}$). The mutation is estimated to have arisen within the last 5,000 years resulting in the second most common variant of normal haemoglobin in the world. Persons with homozygote haemoglobin E (HBEE) have a mild haemolytic anaemia and mild splenomegaly; heterozygote haemoglobin E (HBAE) is benign. Haemoglobin E

is extremely common in Southeast Asia (Thailand, Myanmar, Cambodia and Laos, among others) where its prevalence can reach 30-40% and in some areas equals haemoglobin A frequency. In Thailand the mutation can reach a frequency as high as 50-70%. In India, this abnormal haemoglobin is most prevalent in North-East region, where in certain areas carrier (HBAE) incidence reaches as high as 60% of the population.

## 2.6 QUANTITATIVE VARIATION IN HAEMOGLOBIN (THALASSAEMIAS)

Quite a number of inherited abnormalities are known in man in whom the central defect is the deficiency of a particular protein, and many of these are probably due to mutations which result in a gross but specific reduction in the rate of synthesis of one or more polypeptide chains. The most extensively studied of such conditions are those which involve defects in the synthesis of haemoglobin. Among them are a series of chronic haemolytic anaemias collectively known as the thalassaemias. They appear to be determined by a series of distinct abnormal genes in different heterozygous and homozygous combinations. It is useful to classify the various thalassaemias according to the polypeptide chain primarily concerned in causing the haemoglobin deficit (Ingram and Stretton, 1959). Thus in the β-thalassaemia there is defective synthesis of β-chains whereas in α-thalassaemia, there is defective synthesis of α-chains. β-thalassaemia is characterized in the heterozygote by variable depression in β chain synthesis, a moderate increase in HBA2 (6-7%) and in about half the cases, a slight increase in HBF (1-5%). Defective synthesis of α chains i.e. complete or nearly complete absence of α chain characterizes α-thalassaemia which is incompatible with life because α-chains are required for all of the normal haemoglobins and their structural variants.

Thalassaemia (thalassa - the Greek word for the sea), appears in two clinical states – thalassaemia major and thalassaemia minor. The former, also known as Cooley's anaemia or Mediterranean anaemia, is a severe haemolytic anaemia and afflicted individuals cannot survive unless they receive blood transfusions regularly or undergo bone marrow transplantation. Thalassaemia major is a clinical entity found in individuals homozygous for α or β chain defect i.e. $HB\alpha^{Th}/HB\alpha^{Th}$ or $HB\beta^{Th}/HB\ \beta^{Th}$, where $^{Th}$ denotes thalassaemia. Thalassaemia minor occurs in individuals heterozygous for either α or β chain defect i.e. $HB\alpha^{N}/HB\alpha^{Th}$ or $HB\beta^{N}/HB\ \beta^{Th}$, where $^{N}$ denotes normal production of α or β chain.

In India, the occurrence of both α and β thalassaemias has been reported. β-thalassaemia major has been reported predominantly from East India while cases of it, both major and minor, have been reported from other parts of India, including Uttar Pradesh, Punjab, Gujarat and Kerala. Various haemoglobin variants, more frequently HBE, occur in combination with β-thalassaemia. In α thalassaemia, haemoglobin Barts was detected in 2.4% Bengali population while haemoglobin H (HBH) cases have been reported from Calcutta in East India and West India. Sukumaran (1974) observed that β-thalassemia is probably the commonest inherited haemoglobin disorder in the Indian subcontinent.

## 2.7 GEOGRAPHIC DISTRIBUTION OF *HB*S*, *HB*D* AND *HB*E* IN INDIAN POPULATIONS

***HB*S***

It is the most common variant haemoglobin allele in India and is present mostly in South, West and Central India which inhabit sizeable autochthonous tribal population of the country. The allele has not been reported from North India, except in Chandigarh with a frequency of 0.011. *HB*S* shows polymorphic proportions in the states of Karnataka (0.082), Andhra Pradesh (0.038), Tamil Nadu (0.047), Kerala (0.045), Gujarat (0.065), Maharashtra (0.023), Dadra and Nagar Haveli (0.08) and Madhya Pradesh (0.061). The incidence of the allele for entire India has been estimated to be 0.031 (Bhasin et al., 1994).

***HB*D***

Though a distinct characteristic of populations originating from North-West border Indian state of Punjab, this variant allele of haemoglobin occurs sporadically almost all over North India. Bird et al. (1956) found 5 out of 279 Sikh subjects and one out of 13 Punjabi Hindu subjects to carry haemoglobin D. Five of these had haemoglobins A and D in heterozygous form (HBAD) and one had homozygous haemoglobin D (HBDD), giving a value 2.05% for haemoglobin D variants and a frequency of 0.012 for *HB*D* in Punjab. The authors concluded "It is clear, however, that haemoglobin D, though not common, is something more than a sporadic mutant as had been suspected when it had been found only in one Sikh" (Bird et al., 1955). In addition, this variant haemoglobin has also been reported in Audich Brahmin of Gujarat in 2 out of 200 subjects, giving a value of 1% for HBD variant and a frequency of 0.05 for *HB*D* (Parikh et al., 1969).

***HB*E***

It is the second most common variant haemoglobin allele present in people of India. But the distribution of the allele in the country is essentially limited to East India, where it has been reported with a frequency of 0.303 in Assam, 0.066 in Manipur, 0.363 in Meghalaya, 0.006 in Sikkim and 0.007 in West Bengal; the allele being polymorphic in the former three states. Thus, in general, *HB*E* is confined to people of East India, particularly the Eastern Himalayan populations with Mongoloid affinities. The allele is totally absent from West, Central and South India and from North India there is a solitary report from Uttar Pradesh.

Like any other tropical country of the world, India is a "great reservoir for abnormal haemoglobins" (Saha and Banerjee, 1971) and therefore further studies are desirable to find out the true incidence of haemoglobin variants in this mega diversity country. Needless to mention electrophoresis is essentially the technique of choice for such investigations.

## 2.8 SUMMARY

Isozymes refer to the existence of different molecular forms of an enzyme that catalyze the same biochemical reaction but which differ in their electrophoretic properties. Electrophoretic investigations have led to the discovery of a large number of structurally variant forms of enzymes/proteins attributable to specific

gene mutations. The first example of electrophoretically detectable biochemical variation in humans was that of the well-known blood protein haemoglobin (HB), followed by polymorphisms of serum proteins HP, TF, GC, C3 and BF, and red cell enzymes ACP1, PGM1, AK1, ADA, GPI, ESD and GLO1, among others. The term haemoglobinopathies comprises a group of hereditary abnormalities in which either (a) the haemoglobin structure is altered (e.g. HBS, HBD, HBE) or (b) there is a defect in globin chain synthesis of one of the normal haemoglobin chains ($\alpha$, or $\beta$) resulting, respectively, in $\alpha$ or $\beta$ thalassaemia. Most variant haemoglobins are products of point mutations and their detection is dependent on a difference in their altered electrophoretic mobility. *HB*S* is the most common variant haemoglobin allele in India, followed by *HB*E* and *HB*D.*

## References

Bhasin, M. K., Walter, H. and Danker-Hopfe, H. 1992. *The Distribution of Genetical, Morphological and Behavioral Traits among the Peoples of Indian Region*. Delhi: Kamla-Raj Enterprises.

Chahal, S. M. S., Singh, P., Singh, H., Bansal, R. and Bansal, I. J. S. 2008. *Genetic variation and structure of the people of Uttarakhand, Central Himalayas, India. Hum. Biol*., 80: 409-434.

Pauling, L., Itano, H. A., Singer, S. J. and Wells, I. C. 1949. *Sickle cell anaemia, a molecular disease. Science*, 110: 543-548.

Shaw, C. R. 1969. *Isozymes: classification, frequency, and significance*. *Int. Rev. Cytol*., 25:297-332.

## Suggested Readings

Bhasin, M. K., Walter, H. and Danker-Hopfe, H. 1994. *People of India: An Investigation of Biological Variability in Ecological, Ethno-economic and Linguistic Groups*. Delhi: Kamla-Raj Enterprises.

Cavalli-Sforza, L. L. and Bodmer, W. F. 1971. *The Genetics of Human Populations*. San Francisco: Freeman.

Giblett, E. R. 1969. *Genetic Markers in Human Blood*. Oxford: Blackwell Scientific Publications.

Harris, H. 1977. *The Principles of Human Biochemical Genetics*. Amsterdam: Elsevier / North-Holland.

Mourant, A. E., Kopec, A. C. and Domaniewska-Sobczak, K. 1976. *The Distribution of the Human Blood Groups and Other Polymorphisms*. London: Oxford University Press.

## Sample Questions

1) List different serum protein and red cell enzyme polymorphisms discovered in man. What are the uses of these genetic markers in anthropology?
2) What are normal and abnormal haemoglobins and what techniques are used to study them?
3) Discuss the two major components of haemoglobinopathies in man.
4) Give an account of the distribution of variant haemoglobin alleles in Indian populations.

# UNIT 3 INBORN ERRORS OF METABOLISM

**Contents**



**Learning Objectives**

After reading this unit, you will be able to:

- define and explain with regard to metabolism;
- discuss inborn errors of metabolism; and
- discuss the examples for inborn errors of metabolism such as Phenylketonuria, Alkaptonuria, Galactosemia & Galactosuria; and Hypercholesterolemia.

## 3.1 INTRODUCTION

Metabolism broadly refers to the sum total of all the chemical changes occurring in a cell, a tissue or the body, and these chemical changes are performed through organized multistep pathways involving a number of enzymes, proteins and bio-chemical ingredients. There are several pathways or cycles that intersect with each other forming a cohesive co-ordinated network of chemical reactions. The sum total of chemical reactions performed through different pathways can be classified as either Catabolic (degradative) or Anabolic (Synthetic). The breaking down of complex molecules such as proteins, polysaccharides and lipids etc. to simple molecules such as $Co_2$, $NH_3$ (ammonia) and water is the example of catabolic reaction. The synthetic or anabolic path way refers to the chain of events that lead to the formation of complex products from simple precursors; for example the synthesis of polysaccharide, glycogen from glucose. Investigation of metabolism involves elucidation of complex path ways. Each path way is characterized by action of multienzyme sequences, and each enzyme, in turn, may display important catalytic or regulatory features. As far as Catabolism is concerned, the degradation of complex molecules occurs in three stages and Adenosine Triphosphate (ATP) is released as chemical energy as a degradation of energy-rich molecules. Three stages of Catabolism are: a) Hydrolysis of Complex molecules; b) Conversion of building blocks to simple intermediates c) Oxidation of acetyl CoA. Anabolism refers to the process of formation of complex molecules such as protein, polysaccharides, lipids, nucleic acids from small molecules such as amino acids, sugars, fatty acids, nitrogenous bases etc.

Anabolic reactions require energy, which is generally provided by the breaking down of ATP to ADP and Pi. The figure 3.1 given below demonstrates a comparison of catabolic and anabolic pathways.

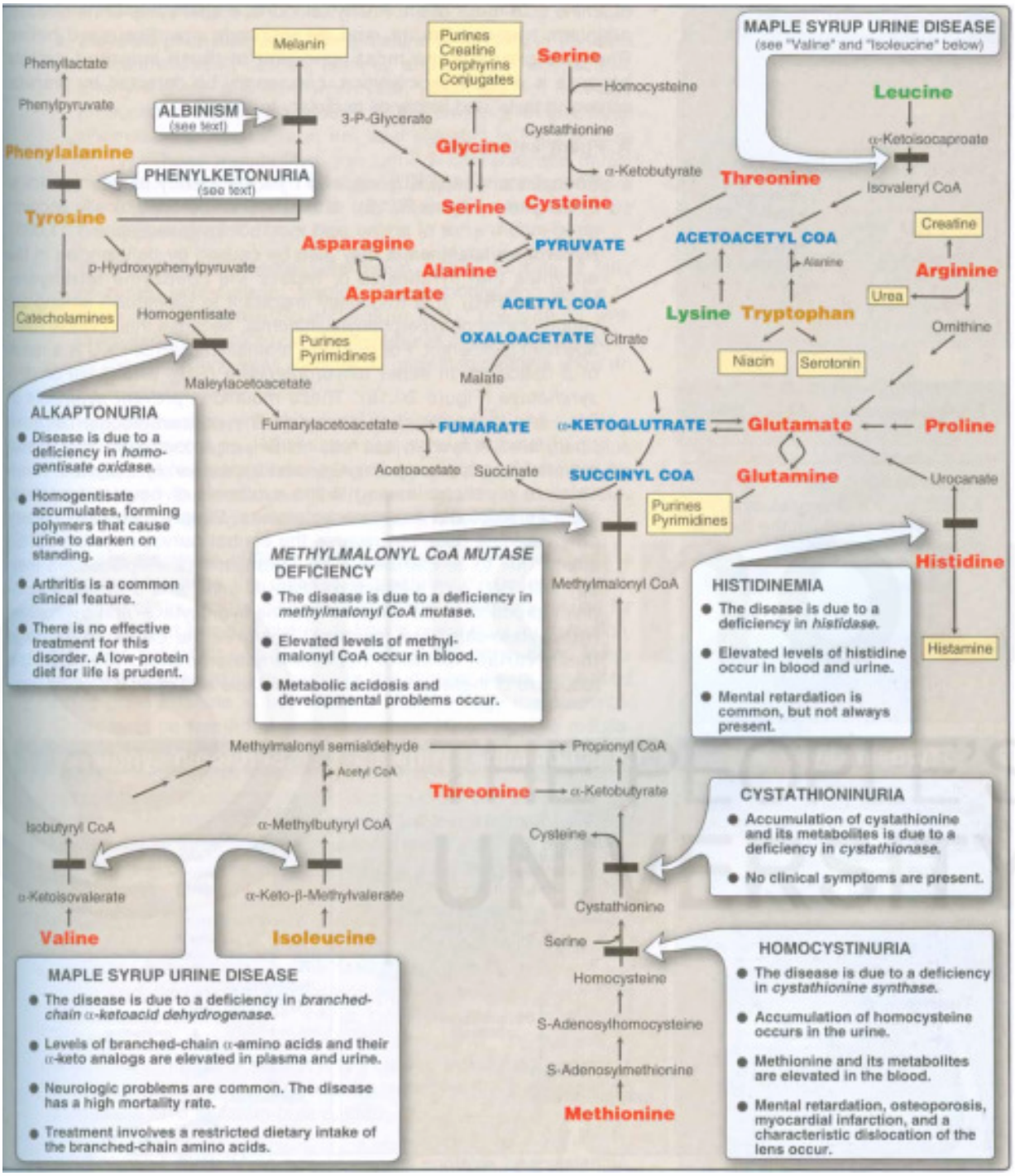



**Fig. 3.1: Enzymatic pathways of metabolism of amino acids in Humans (Courtesy : Champe, P.C, Harvey, R.A and Ferrier, D.R (Ed). 2005. Biochemistry).**

## 3.2 INBORN ERRORS OF METABOLISM: GENESIS

Sir Archibald Garrod's Paper on "The incidence of Alkaptonuria : A study of Chemical individuality" published in 1902, was a monumental paper that ushered in a new paradigm so far as the history of human genetics is concerned. The publication of the paper by Garrod opened up new perspectives for two reasons.

First, the results obtained by Garrod on genetics of Alkaptonuria reinforced the application of Mendel's paradigm on humans. Second, the findings of the paper subsequently led to the emergence of a new concept as well as an important area of study in human genetics under the rubric "Inborn Error Metabolism". With the accumulation of evidences on genetic basis of enzyme action in man, the study of inborn errors of metabolism in humans has become an important area of study in human bio-chemical genetics having tremendous application value in medical / clinical genetics.

Garrod's work gave a strong ground in favour of the concept of one gene-one enzyme hypothesis, which was established in the 1940s and 1950s by the pioneering work of Beadle and Tatum, who were awarded Nobel Prize.

**Principles of Inborn – errors of Metabolism**

Metabolism broadly refers to complex bio-chemical pathways functioning as an integrated system of chemical reactions controlled in some manner by genes and these genes regulate or control specific reactions in the system by acting directly as enzymes or by determining the specificities of enzymes, where as inborn error metabolism denotes inability on the part of humans to catalyze certain metabolic pathways completely due to alterations of enzymes, and this modification of the action of enzymes is attributed to genetic factor or an inborn factor. These defects or errors or blockages in the action of enzymes are due to mutations in the gene/ genes that code for the enzyme/enzymes. Normally, they usually affect only one in a series of iso-enzymes. If more than one iso-enzyme is affected, these variant iso-enzymes may share common polypeptide chain or secondary effects may occur on the structure of enzyme. The deficiency of an enzyme caused by mutation of a gene that is active only in one tissue affects the phenotypes of its carrier in different ways as compared to an enzyme deficiency affecting many tissues. This is analogous to pleiotropic effect of the gene. However, this phenomenon does not always occur in all humans. It may so happen that an enzyme defect that is found in all tissues some time lead to phenotypic anomaly in one tissue only - this is due to the fact that defect is ameliorated more early in other tissues.

## 3.3 INBORN ERRORS OF METABOLISM IN MAN

The following are the four examples for the Inborn Errors of Metabolism in Man. They are Phenylketonuria (PKU), Alkaptonuria, Galactosemia & Galactosuria and hypercholesterolemia.

### 3.3.1 Phenylketonuria (PKU)

This metabolic disorder was first described by Folling in 1934 in mentally retarded patients, who emitted a peculiar "mousy" odor. PKU is known as one of the best known inborn errors of metabolism of amino acid in humans. The enzymology of PKU established that L-phenylalanine is an essential amino acid with respect to protein synthesis. It is known that only a small proportion of L-phenylalanine is utilized in normal humans for protein synthesis. The larger proportion is oxidized primarily to tyrosine due to action of enzyme *phenylalanine hydroxylase.* The *hydroxylase* structurally contains two protein components: 1) one which is labile and found only in the liver. 2) the other with stable feature is found in other tissues. PKU is primarily caused by complete deficiency of action the enzyme – *phenylalanine hydroxylase* mostly found in liver. The non-action of

the enzyme *phenylalanine hydroxylase* leading to phenylketonuria is in fact a recessive mutation in the gene that codes for the enzyme *phenylalanine hydroxylase* and the gene for PKU is located on chromosome 12. PKU is an autosomal recessive disorder whose frequency in Northern European populations is about 1/10,000 live births. The Pathways of phenylalanine metabolism is presented in figures 3.2 and 3.3 and the pathways of phenylalanine metabolism in normal individuals and in patients with PKU is presented in figure 3.4.

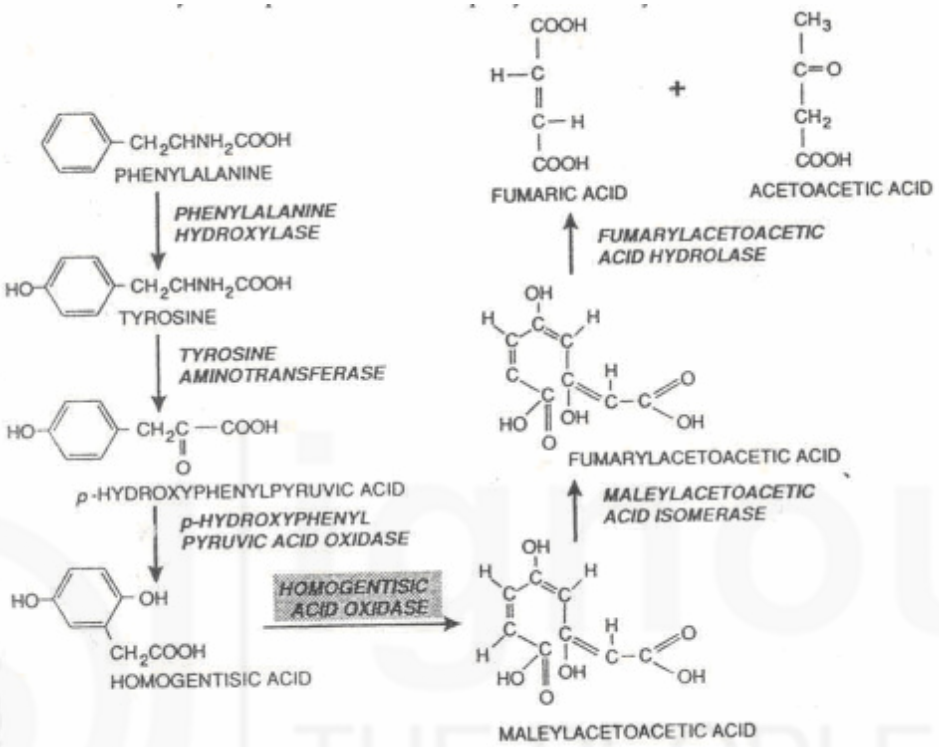



**Fig. 3.2: Enzymatic steps in the catabolism of phenylalanine and tyrosine to aceto-acetic acid (Courtesy: Gelehrter, T.D and Collins, F.S. 1998. Principles of Medical Genetics).**

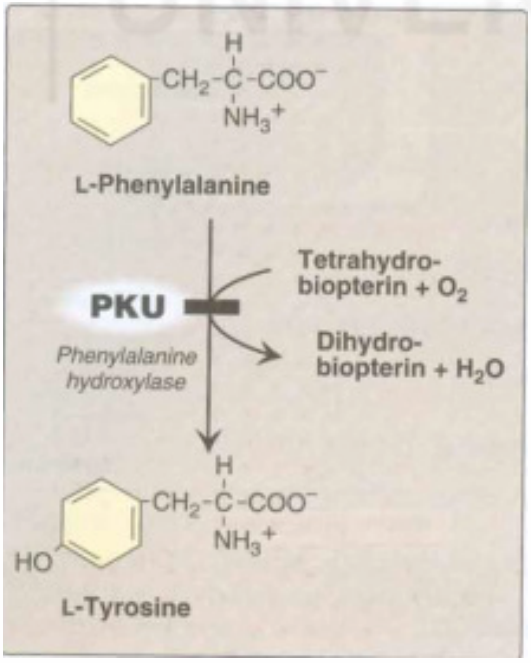



**Fig. 3.3 : A deficiency in Phenyllalanine Hydroxylase results in the disease Phenylketonuria (PKU) (Courtesy: Champe, P.C, Harvey, R.A and Ferrier, D.R (Ed). 2005. Biochemistry).**

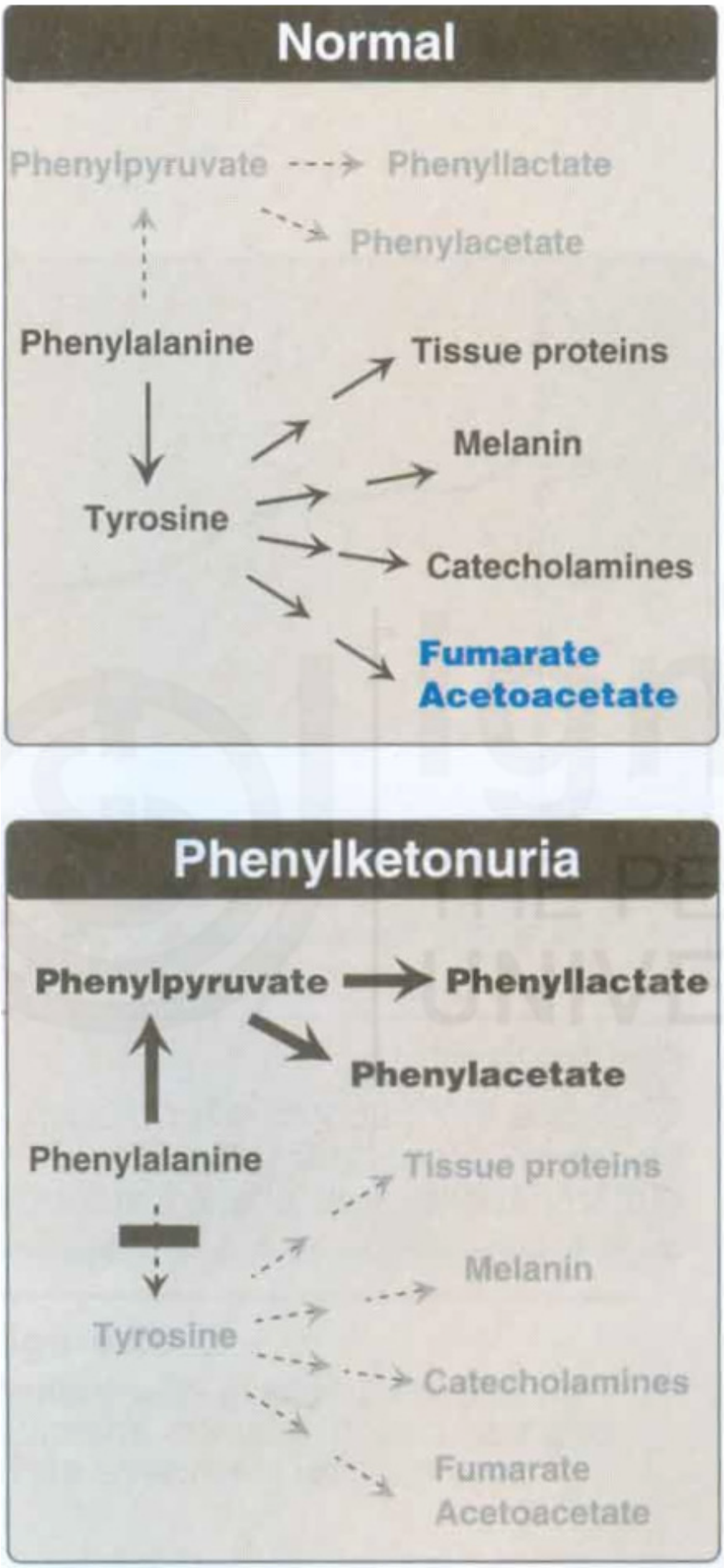



**Fig. 3.4:** **Pathways of phenylalanine metabolism in normal individuals (above) and in PKU patients (below) (Courtesy: Champe, P.C, Harvey, R.A and Ferrier, D.R (Ed). 2005. Biochemistry)**

## Characteristics of PKU

a) **Elevated phenylalanine**

The phenotypic–clinical feature of patients with PKU is characterized by microcephaly (small head) and profound mental retardation. As stated above the genetic mutation that blocks the action of the enzyme *phenylalanine-hydroxylase* in the hepatic system leads to low concentration of tyrosine and high concentration of phenylalanine in plasma. The concentration of phenylalanine is elevated in tissues, plasma and urine in persons with PKU. The reason is that three types of metabolites of phenylalanine such as phenyl lactate, phynyl acetate, and phenyl pyruvate, which are not formed in normal individuals due to action of enzyme phenylalanine hydroxylase, are produced in elevated concentration in PKU patients. Thus the name of the disease phenylketonuria (PKU) which signifies the presence of phenyl ketone in the urine in the form of phenyl pyruvate. Such high concentration phenylalanine metabolites such as phenyl-ketones are excreted through urine, giving a characteristic 'Mousy' odor. The high concentration phenylalanine in plasma is attributed to mental retardation among infants. It is known that the formation of tyrosine from phenylalanine is catalyzed by the action of enzyme phenylalanine hydroxylase. The reaction requires molecular oxygen and coenzyme - tetrahydrobiopterin (BH4) which is synthesized by the body. One atom of molecular oxygen becomes the hydroxyl group of tyrosine, the other atom is reduced to water. During reaction teatrahydrobiopterin is oxidized to dihydro hydrobiopterin in a separate reaction needing NADPH. Hyper-phenylalanimenia may also caused by deficiencies in the formation of enzyme, that synthesizes or reduces the hydroxylases co-enzyme - tetrahydrobiopterin.

b) **CNS Syndrome**: Mental retardation, failure to walk, talk, seizures, hyperactivity, tremor, microcephaly and failure to grow are characteristic findings in PKU.

c) **Hypo-pigmentation:** Patients with PKU with high levels of phenylalanine demonstrate a deficiency of pigmentation (fair hair, light skin colour, and blue eyes).

**Treatment of PKU** : As stated earlier, the PKU is a disease of metabolic disorder, caused by an autosomal recessive mutation, that inhibits the action of enzyme which is responsible for catabolizing phenylalanine to tyrosine. Since PKU has a genetic basis, both prevention as well as treatment of patients of varying age groups both at family level and community level are regarded as public health problem. The most pressing need is to address the issue of rise in phenylalanine metabolites in urine along with high incidence in the serum phenylalanine level. The therapy that was first administered for reducing phenylalanine intake through dietary supplementation was made by Bickel in 1953. Since then, several groups attempted on dietary supplementation programme and have achieved success. The evidence is now confirmed that dietary therapy used have the most ameliorating effect on the development of PKU patients. Two important guidelines have to be borne in mind while treating PKU.

i) In order to prevent brain damage the diet should be started as soon as possible within the first week of life.

ii) The metabolic status of the child especially phenylalanine levels should be monitored carefully.

Adults do not require life long treatment since adult frames are resistant to rise to metabolite concentration found in PKU.

### 3.3.2 Alkaptonuria (AKP)

Alkaptonuria is a rare metabolic disorder involving a deficiency in the enzyme *"homogentisic acid oxidase",* which is in fact responsible for the formation of maleylacetoacetate from Homogentisate. The bio-chemical pathway involving the Alkaptonuria is shown in figure 3.1.

Sir Archibald Garrod's first monumental study on Alkaptonuria in 1902 brought about a new perspective in human genetics. He could establish through his seminal paper that metabolic disorders have genetic basis. This was the first example that genes may register their effect by coding for enzymes genetically. Alkaptonuria was thus recognized as the first autosomal recessive human disease that emanates from the deficiency of an enzyme. The characteristic symptoms of the disease are:

a) Formation of homogentisic aciduria: In this condition, the patients urine contains elevated levels of homogentisic acid, which is oxidized to a dark pigment on standing leading to black colour (fig. 3.5)

b) Development of Arthritis, particularly in the spine and large joints due to the degenerative changes.

c) The deposition of endogenous homogentisic acid autooxidation products in the cartilage and collagenous tissue lead to blue or black ochronotic pigmentation of cartilage, and black deposits in the sclera of the eyes, other features of this condition can include heart problems, kidney stones and prostrate stone.

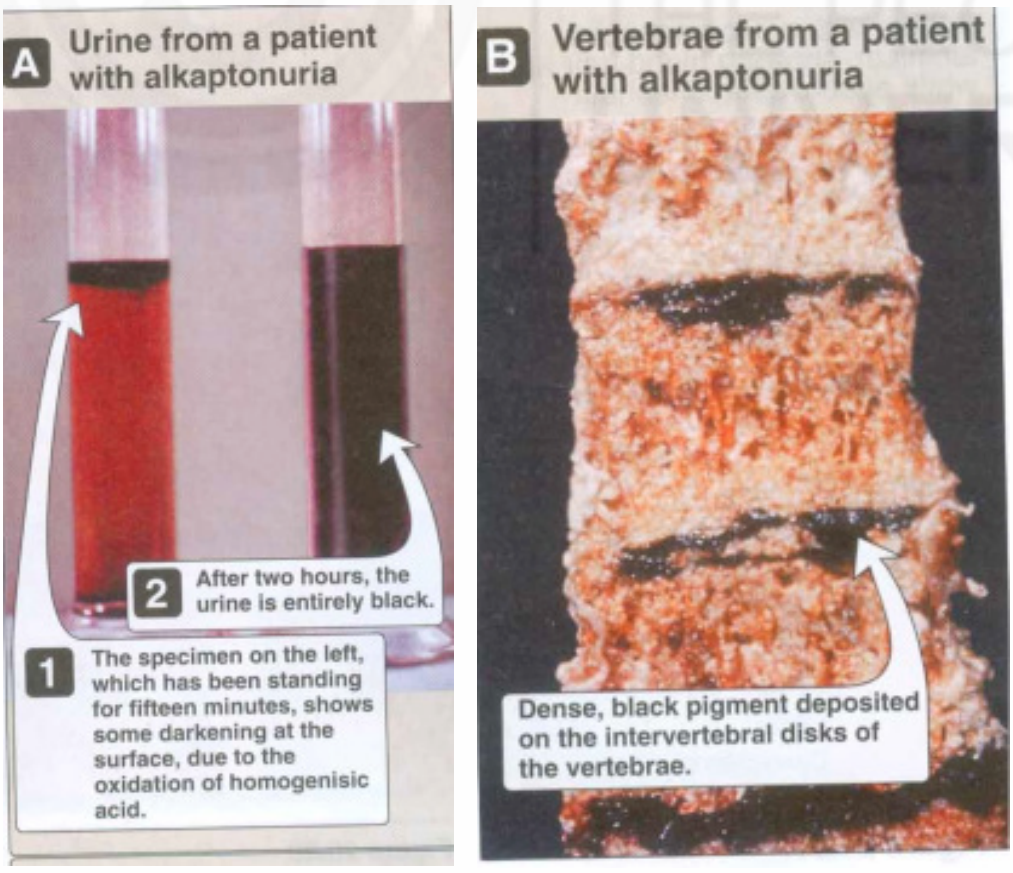



**Fig. 3.5. A patient with alkaptonuria A. Urine and B. Vertebrae.**

**How common is AKP?**: This condition is rare, affecting I in 250,000 to 1 million people world wide. AKP is more common in certain Areas of Slovakia (1 in 19,000) and in the Dominican Republic.

Genetics of Alkaptonuria : It is now clearly established that mutation in HGD gene causes Alkaptonuria. The gene encoding an enzyme called homogentisate 1, 2- dioxgenase (also known as *homogentisate oxidase*). The gene for HGD is located on chromosome 3 in region 3q21-q23.

The enzyme *Homogentisate Oxidase* is primarily active in the liver and kidney. It is chiefly responsible for breaking down the amino acid phenylalanine and tyrosine, into a molecule called homogentisic acid. Homogentisic acid oxidase adds two oxygen atoms to homogentisic acid converting homogentisic acid into another molecule called Maleylacetic acetate.

More than 40 mutations in HGD gene have been detected in people with Alkaptonuria. The substitution of the amino acid valine for methionine at position 368 is the most common form of HDG mutation in European population. The mutation in this gene inactivates the enzyme. Without HGD normal gene functioning, the enzyme homogentisic oxidase become ineffective accounting for building up homogentisic acid in the body. The gene for Alkaptonuria is inherited as autosomal recessive trait. The person with Alkaptonuria inherits both copies of the recessive genes from parents who are heterozygotes for Alkaptonuria.

### 3.3.3 Galactosemia and Galactosuria

Galactosemia & Galactosuria are identified as disorders of carbohydrate metabolism especially of Galactose metabolism.

The major source of Galactose is obtained from lactose – the principal source of milk and milk products. However, some amount of galactose can be obtained by lysosomal degradation of complex carbohydrates such as glycoproteins, and glycolipids – the membrane components of tissues.

**Metabolic Pathways**: The metabolic pathways that involve metabolism of Galactose are as follows:

a) Phosphorylation of Galactose

b) Formation of UDP-galactose

c) Use of UDP-galactose as a carbon source for glycolysis or gluconeogenesis

d) Role of UDP -Galactose in biosynthesis reaction.

The figure 3.6 showing metabolism of galactose is given below:

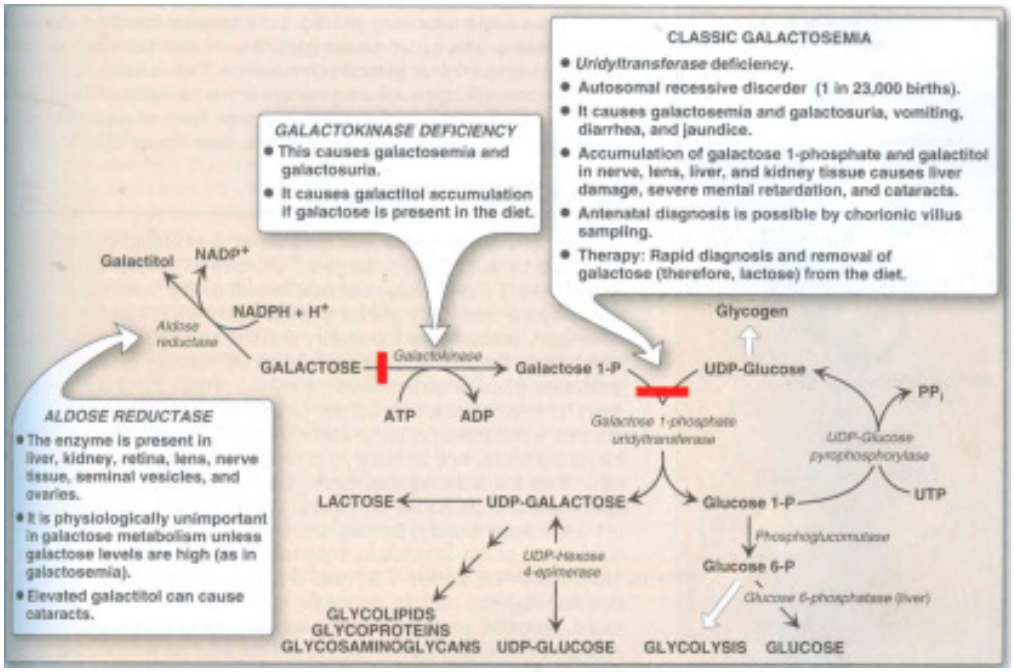



**Fig 3.6: Metabolism of Galactose (Courtecy: Champe, P.C, Harvey, R.A and Ferrier, D.R (Ed). 2005. Biochemistry)**

**Galactosemia and galactosuria** arise in individuals who are deficient in secreting enzymes– a) *Galacto kinase* that metabolizes Galactose to Galactose 1-P phosphate. b) *Galactose 1. Phosphate Uridyl transferase* that converts Galactose1-Phosphate to UDP-Galactose and Glucose1-Phosphate. Galactosuria: As stated earlier the cases of galactosuria and galactosemia are due to the deficiency of galactokinase enzyme that converts galactose to galactose1phosphate.

**Classic Galactosemia:** It is in fact a deficiency of the enzyme Uridyltransferase. The deficiency is due to autosomal recessive disorder. It causes both galactosemia and galactosuria . The frequency of this deficiency is 1 in 23,000 births. Individuals with Uridyl transference enzyme deficiency facilitate the accumulation of Galactose 1-Phosphate, and consequently excess of galactose in the body lead to vomiting, diarrohea and jaundice. In some cases, the excess accumulation of galactose 1-Phosphate also causes liver damage, severe mental retardation and cataract. Genetically this disorder is an autosomal recessive mutation in the gene that codes for the enzyme Galactose 1-Phosphate uridyl transferase.

### 3.3.4 Disorders of Lipid Metabolism: Familial Hypercholesterolemia

Familial Hypertension (FH) is an important cause of heart disease. It is one of the most common autosomal dominant disorder. In the individuals with FH the plasma cholesterol levels are twice higher than normal (i.e. about 300-400 mg/dl). The high cholesterol level is due to deficient or defective LDL receptors leading to enhanced levels of endogenous cholesterol synthesis.

All cells require cholesterol as a component of their plasma membrane. They can either synthesize their own cholesterol or they get it from the extra cellular environment, where it is carried principally by low density lipoprotein (LDL). In a process known as endocytosis (fig. 3.7) LDL-bound cholesterol is taken into

the cell via receptor on cell surfaces. In view of the lack of normal number of LDL receptors, cellular cholesterol uptake is reduced, and thereupon circulating cholesterol levels rise. The gene for FH is located on chromosome 19.

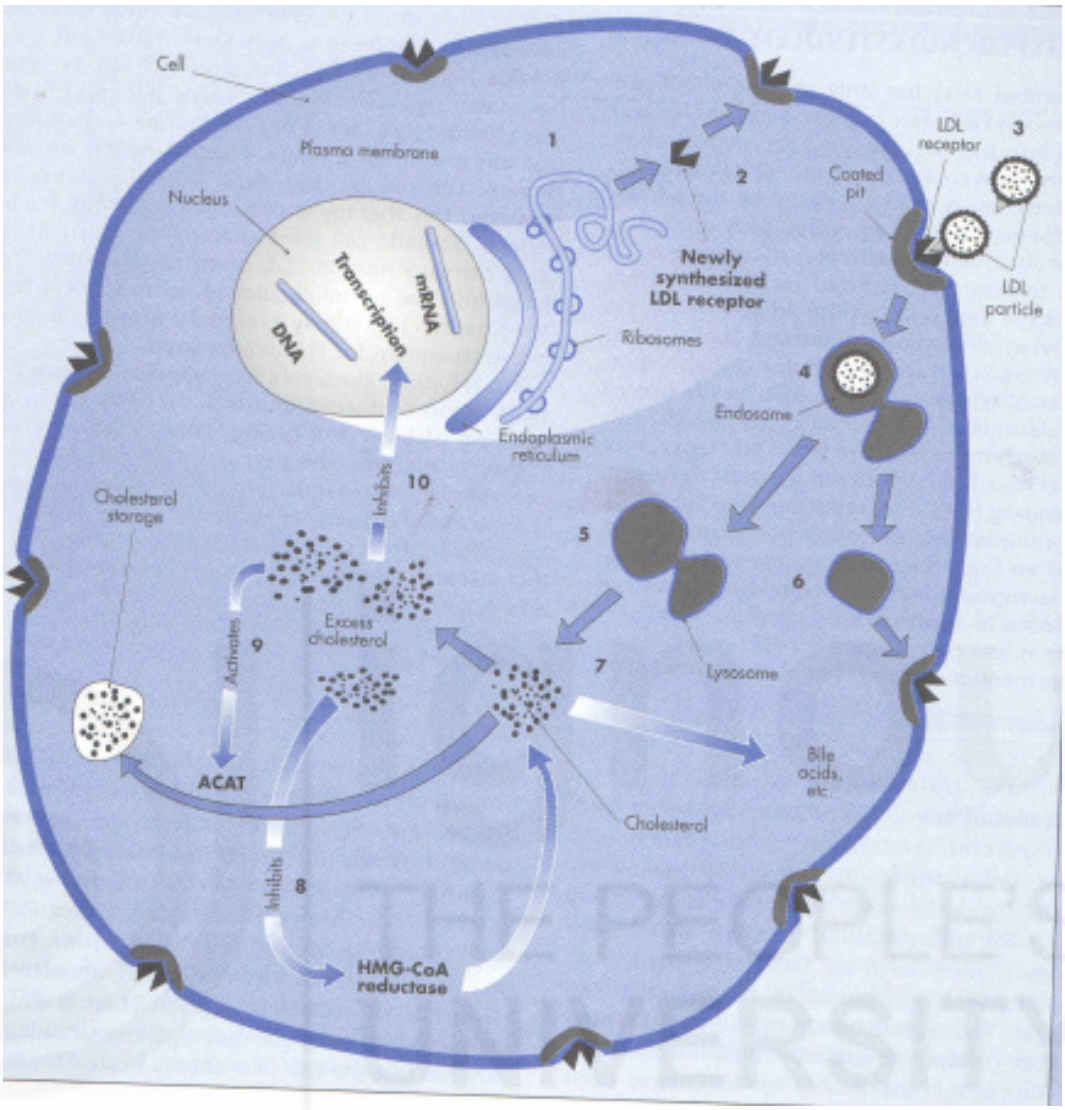



**Fig. 3.7: The process of receptor mediated endocytosis (Courtesy: Jorde, L.B, Carrey, J.C, Bamshad, J.C and White, R.L. 2000. Medical Genetics, 2nd edition).**

In case of homozygotes (cholesterol levels raising from 600 to 1200 mg/dl) the distinctive cholesterol deposits in the skin and tendons called xanthomas occur.

## 3.4 SUMMARY

The present unit describes few examples of inborn errors of metabolism in man such as phenylketonuria, alkaptonuria, galactosuria, and hypercholesterolemia with reference to biochemical and genetic aspects. These topics constitute the central theme of one important area of human biochemical genetics, covering the abnormal aspect of human genetic variation. The study of these metabolic diseases focuses on early diagnosis as well as treatment, aspects which are also important from clinical point of view.

## Suggested Reading

Champe, P.C, Harvey, R.A and Ferrier, D.R (Ed). 2005. *Biochemistry*, 3rd edition. MD: Lippincott Williams & Wilkins.

Gelehrter, T.D and Collins, F.S. 1998. *Principles of Medical Genetics,* Second edition, USA:Williams and Wilkins Co.

Jorde, L.B, Carrey, J.C, Bamshad, J.C and White, R.L. 2000. *Medical Genetics,* 2nd edition, St Louis: Mosby.

Vogel, F. and Motulsky, A.G. 1996. *Human Genetics: Problems and Approaches,* Berlin: Springer.

## Sample Questions

1) Discuss on example of error of amino acid metabolism. Can you treat it by any means?
2) Give a brief note on an example of error of lipid metabolism leading to heart disease.
3) What is the difference between galactosaemia and galactosuria? Comment
4) Write a note on the first metabolic disorders described by Garrod?