THE MOLECULAR BASIS OF SINGLE GENE DISORDERS

Prof. Arjumand Warsy
Department of Biochemistry, College of Science, King Saud University, Riyadh

---

Introduction

- Genetic disorders result from mutations in the genes or the chromosomes.

- Mutation → Altered protein structure → Abnormal protein Function
  ↓
  Genetic Disease

- Alterations caused by mutations in different classes of proteins disrupt cell and organ function and hence lead to genetic disease.

  Over 6000 single gene defects and over 1960 chromosomal defects are known which lead to genetic diseases of varying severity. Some are assymptomatic while others are lethal.

- Molecular pathology of genetic disorders: Mutation occur in or around genes and result in:
  (a) Single base changes - Nonsense mutations
  (b) Frameshift mutations
  (c) Insertions
  (d) Deletions
  (e) Chain termination mutations
  (f) mRNA processing mutation
  (g) Poly A addition site mutation
  (h) `Regulation box' mutation
  (i) Inversions
  (j) Fusion genes
  (k) Partial gene deletions
  (l) Complete gene deletions

These mutations may affect:

- Amino acid composition of proteins and hence its structure and function:
  - Increase activity
  - Decrease activity
  - Complete loss of activity

- Premature termination of proteins and hence altered structure and function.

- Protein elongation leading to an abnormal protein with altered functions.

- Decreased stability of mRNA and hence decrease in the level of protein.

- Decrease stability of the protein.
- Complete absence of the protein.

The various sites at which the defect may occur are presented in Figure 1. A graphic presentation of effect of mutation is presented in Figure 2.

Mutations may lead to:

(a) enzyme defects
(b) defects in transport function of proteins
(c) defects in structure of cells and organs
(d) altered extracellular homeostasis
(e) defect in control of growth and differentiation
(f) altered intercellular metabolism and communication

Examples of diseases due to mutations in different classes of proteins.
Enzyme defects: Hundreds of examples are known, affecting almost all areas of metabolism. As a result of an enzyme defect, a substrate may accumulate or the product is decreased e.g.

Metabolism	Deficiency	Disease	Inheri-tance
Amino acid	Phenylalanine hydroxylase	Phenylketonuria	AR
Carbohyd-rate	Galactose-1-phosphate uridyl transferase	Galactosemia	AR
Organic acid	Methylmalonyl CoA mutase	Methylmalonyl aciduria	AR
Lipid	Medium chain acyl - CoA mutase	Methylmalonic aciduria	AR
Purine	Adenosine deaminase	Severe combined immune deficiency	AR
Purine	Hypoxanthine-guanine phosphoribosyl transferase	Lesch Nyhan syndrome	XR
Porphyrins	Porphobilinogen deaminase	Acute intermitten porphyria	AD

Defects in Receptor Proteins

Several diseases result due to mutations in the receptor proteins

Defective Receptor	Disease	Inheritance
LDL receptor	Familial hyper-cholesterolaemia	AD
Insulin receptor	IDDM	?

Defects in Membrane Transport

Defect	Disease	Inheritance
CF transmembrane conductance regulation (CFTR gene)	Cystic fibrosis	AR 1 in 2000 Caucasians
Glucose transporter gene	NIDDM	?

Defects in Control of Growth and Differentiation

Defect	Disease	Inheritance
Tumor suppressor gene	Cancer (Retinoblastoma Osteosarcoma)	AD
Oncogenes	Cancer Chronic myelogenous leukaemia (CML)	Somatic Muation

Defects in structure of cells and organs

Defect	Specific defect	Disease	Inheritance
Protein structure defect	Dystrophin defect	Duchenne Muscular Dystrophies	XR
	Collagen defect	Osteogenesis imperfecto	AD; AR
	Spectrin defect	Hereditary spherocytosis	AD

Prof. Dr. Mohsen A.F. El-Hazmi
Department of Medical Biochemistry & W.H.O. Collaborating Centre for
Haemoglobinopathies, Thalassaemias and Enzymopathies
College of Medicine & King Khalid University Hospital, Riyadh

---

Chromosomal Disorders

- Mutations (Deletions, translocations, additions, formation of ring chromosomes) result in structural abnormalities of chromosomes and cause genetic defect:
  - The amount of genetic material may remain the same yet the change in chromosome structure may have serious consequences.
- Nondisjunction at meiosis and other mechanisms results in altered number of chromosome.
  - Both increase or decrease in number may occur:
    e.g. Trisomy 21, 18, 13 etc. have one extra 21, 18 or 13 chromosome, respectively.
  - Decrease in the number of chromosome.
  - Increase in the whole set of chromosomes 2n+n.

These mutations are serious and affect development, maturation and growth and may cause mental retardation (Discussed in another lecture on chromosomal anomalies).

Defect in Extracellular Homeostasis

Defect	Disease	Inheritance
Complement gene	Complement deficiency	AD, AR
Factor VIII deficiency	Lemophelia A	XR
α1AT	Lung and liver disease	AR

Defect in Transport Proteins

Defect	Disease	Inheritance
Single point mutation in: - α-globin gene - ß-globin gene	 Haemoglobinopathies or Thalassaemias	 AR AR
Deletion of: - α-globin gene	 α-thalassaemia	 AR

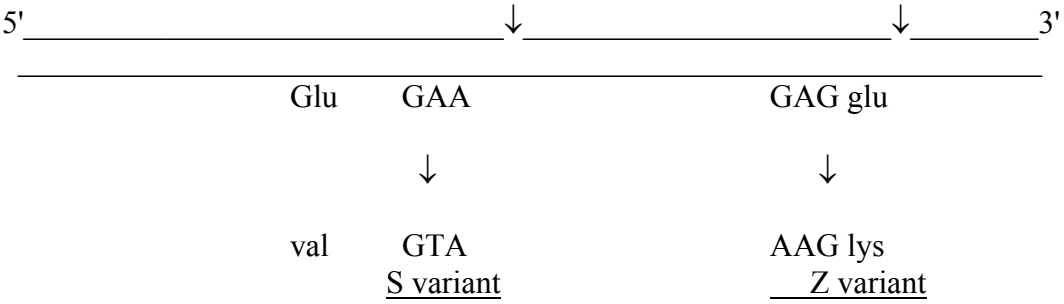
- ß-globin gene	ß-thalassaemia	AR
Copper transport protein	Menekes syndrome	XR

Single Gene Disorders

Introduction

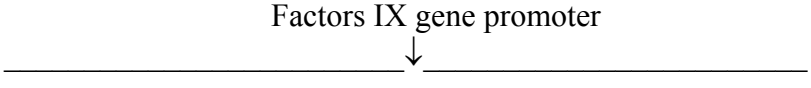
- Single gene disorders are caused by mutations in or around the gene.
- As stated in the 11th edition of the McKusick's Mendelian Inheritance in man, published in 1994, over 6000 single gene disorders are known to affect mankind and together they affect almost 1% of the populations.
- The many different types of mutations that affect the genes or their surrounding areas were discussed in the lecture on mutations and a few are listed in Table 1. A few examples of the molecular pathogenesis of some single gene disorders are listed in Table 2. In this presentation, we will pick up a few well studied examples of single gene disorders and will explain the molecular basis of these disorders.
- An interesting point to remember is that a single disease may be caused by different types of mutations, some are substitutions, others are additions, deletions or frameshift mutations e.g. ß-thalassaemia are caused by >130 different point mutations in different parts of the gene. The severity of the disease depends on the nature and location of the mutation and extent of reduction in ß-globin chains synthesis.

(i) <u>α1-Antitrypsin - deficiency:</u> Caused by missense mutation α1-AT gene is located on chromosome 14q32. It shows extensive polymorphism and exists in the



form of several variants, resulting generally from point mutations (substitution). The Z variant which gives severe α1-AT deficiency has a point mutation G → A which changes the codon GAG → AAG and substitutes Lys for glu. The S variants has GAA to GTA mutation which substitutes val for glu.

(ii) <u>Hemophilia caused by missense mutation on X chromosome</u>



--- CTAATCGACCTTA CCACTTTCACAATCTG------

↓
G

The A → G mutation in the promoter site in the 5' untranslated region of factor IX gene affects the expression of the gene and hence factor IX mRNA is not transcribed and factor IX is not produced thus leading to hemophilia.

(iii) Neurofibromatosis type I caused by a nonsense mutation

A mutation C → T converts a codon for arginine CGA to TGA, a stop codon. This causes premature termination with the production of the neurofibromatosis type I allele.

	Asp	Asp	Ala	Lys	Arg	Glu
Normal allele	GAT	GAT	GCC	AAA	CGA	CAA
					↓	
NFI allele	GAT	GAT	GCC	AAA	TGA	CAA
	Asp	Asp	Ala	Lys	Stop	

(iv) Tay-Sachs disease results from RNA splice junction mutations

A specific nucleotide sequence is located on the intron/exon (acceptor site) and exon/intron (donor site) boundaries and are required for normal RNA splicing.

Acceptor site Donor site
Intron ↓ Exon ↓ Intron

Any mutation in these sites results in abnormal RNA splicing and hence affects protein synthesis.

A G → C mutation in the exon/intron junction of hexosaminidase A gene results in Tay-Sachs disease.

		Normal splice	
Normal	Exon	↓	Intron
Hexosaminidase A allele	A G G C T C T G	g t a a g g t	
		↓	

Tay-Sachs allele	A G G C T C T G	c t a a g g g t
		No splicing

Similar splice site mutations produce phenylketonuria, hemophilia B, ß-thalassaemia.

(v) Substitutions produce several single gene disorders, including hemophilia A&B. A specific type of mutation referred to as "hotspot" mutation occur in CG doublets giving rise to C → T or G → A transition. Mutations in CG doublets are several times more than in other sequences. These are referred to as the "hotspots" for mutation in human genome.

(vi) Single base deletion cause frameshift

Single base delition changes the whole amino acid sequence in the polypeptide chain from the point of deletion e.g. the ABO locus (glycosyltransferase), the A allele has a single base deletion which leads to the formation of the O allele.

	Leu	Val	Val	Thr	Pro
'A' allele	C T C G T G G T G A C C C C T T				
'O' allele	C T C G T G G T - C C C C T T				
	Leu	Val	Val	Pro	

(vii) Tay-Sachs disease results from insertion of four base pairs

An insertion of four base pairs in the hexosaminidase A gene results in frameshift mutation and premature termination. There is a complete enzyme deficiency and a severe phenotype of Tay-Sachs disease

Normal	Arg	Ile	Ser	Tyr	Gly	Pro	Asp	
Hexosaminidase gene CGT	ATA	TCCT	TAT	GCC	CCT	GAC		
Tay-Sachs allele	CGT	ATA	TCT	ATC	CTA	TGC	CCC	TGA
	Arg	Ile	Ser	Ile	Leu	Cys	Pro	Stop

(viii) Other small deletions or insertions in the gene results in several diseases.

Cystic fibrosis results from deletion of three base-pairs which does not cause frameshift, but deletion of only one amino acid.

	Ile	Ile	Phe	Gly	Val
Normal DNA	TATC	ATC	TTT	GGT	GTT
CF DNA	TATC	AT		TGGTGTT	

Ile Ile Gly Val

This mutation is found in 70% of all CF cases so far investigated.

(ix) Large deletions and insertions result in several diseases.
Majority of the α-thalassaemias results from deletion of part or whole of α-globin gene in the α-globin gene cluster.

- Some ß-thalassaemias result from deletion of different lengths of DNA in the ß-globin gene cluster.
- Total deletion of steroid sulfatase locus occurs in 90% of X-linked ichthyosis.
- Total deletion of X-linked ornithine transcarbamylase deficiency occurs in 10% of the cases.
- Deletions within the large dystrophin gene on the X-chromosome occur in Duchennes Muscular Dystrophy in almost 60% of cases.
- Large insertions are rare. In two sporadic cases out of 200 of hemophilics, large insertions have been detected in an exon of factor VIII genes interrupting the coding sequence and inactivating the gene. The copies inserted are those of L1 sequence. This is known as "insertional mutagenesis".

(x) Haemoglobinopathies and Thalassaemias as models of molecular diseases

- Haemoglobin (Hb) is a protein in the red cells with a quaternary structure and is made up of 4 subunits known as α and globins and non α-globin chains.
- Types of Hb in adults

Hb A	-	Major (~95%)	$\alpha_2\beta_2$
Hb F	-	(<1%)	$\alpha_2\gamma_2$
Hb A_2	-	(2.5-3.5%)	$\alpha_2\delta_2$

- Location of α genes and non α-genes
The α-globin genes are located on chromosome 16 in the α-globin gene cluster 5'-ξ-_ξ-_α1-α2-α1-3' and the ß-globin genes are located on chromosome 11 in the ß-globin gene cluster. 5'-ε-Gγ-Aγ-ψß-δ-ß-3'
- Different globin genes are switched on at different stages of development.
- In the adult life only α, ß, δ, γ (at a very low level) are switched on.
- Mutations in and around the globin genes lead to haemoglobinopathies.
- Two major groups have been recorded

(i) Disorder of haemoglobin structure
(ii) Disorder of haemoglobin biosynthesis.

(i) Disorders of Haemoglobin Structure

When the mutation occurs in the exons, it may affect the amino acid sequence, and alter the structure and hence function of haemoglobin

- Over 500 Hb variants have been reported to date. Several of these can be separated on electrophoresis and isoelectric focussing as they differ in their charge and isoelectric pH, respectively.
- The mutations identified leading to haemoglobin structural variants include:
  - Single point mutations
  - Deletions
  - Insertions
  - Frameshift
  - Chain termination mutations
  - Chain fusion

Sickle cell Haemoglobin (Hb S)

Hb S results from a single point mutation GAG → GTG which results in substitution of valine for glutamic acid at position 6 of ß-globin chains of Hb. HbS in heterozygotes (HbAS) is asymptomatic, but in homozygous state (Hb SS) it results in sickle cell anaemia an inherited haemolytic anaemia with several complications.

(ii) Disorders of Haemoglobin Biosynthesis - The Thalassaemias.

- The thalassaemias are the most common single gene disorders in humans.
- Occurring at a high frequency in the population of the Middle East, the Mediterranean area, the Indian sub-continent and South-East Asia.
- They are heterogenous group of disorders in terms of both mutations and clinical presentations.
- They are classified according to the globin chain which is synthesized in reduced amount.

α-thal.	Decreased α-chain synthesis
ß-thal.	Decreased ß-chain synthesis
δß-thal.	Decreased δ & ß chain synthesis
γδß-thal.	Decreased γ, δ and ß chain synthesis

- As a result of an imbalance of the α/non-α-chains, the free globin chains accumulate in red cells and often precipitate leading to haemolytic anaemia, with consequent compensatory hyperplasia of the bone marrow.

Mutations producing α-thalassaemia.

- α-thal. generally results from deletion, though non-deletion type of α-thal. also exist.
- The deletion of one, two, three or all four α-globin genes is shown below

Normal		α-thal.2 (hetero)	α-thal.2 (homo)
αα/αα	-α/αα		-α/-α
α-thal.1 (hetero)		Hb H disease	α-thal.1 (homo)

--/αα --/-α --/--

- Deletion of various length fragments in the α-globin gene cluster on chromosome 16 have been reported.
- Some cases of point mutation leading to α-thal. phenotype are also known i.e. the α-globin chain is either synthesized in decreased amount or what is synthesized is unstable and leads to globin chain imbalance.

Mutations producing Beta-thalassaemia

- Underproduction of ß-globin chains increases α/non-α globin chains ratio and leads to ß-thal.
- Majority of the ß-thal. result from point mutations.
- Over 130 mutations have been reported occurring as:
  - . mRNA splice junction mutations
  - . Frameshift mutations
  - . Cap site mutations
  - . RNA cleavage mutations
  - . Initiation codon mutations
  - . Nonsense mutations
  - . Small deletions
  - . Mutations producing unstable globin
  - . Mutations affecting rate of transcription
  - . Mutations producing unstable mRNA

  The position of mutation varies and is in or around the ß globin gene cluster. In each case the synthesis of ß-globin chain is either completely stopped (i.e. ß°-thal.) or is reduced (i.e. $ß^{+}$-thal.).

  Deletion of part of whole of the ß-globin gene cluster resulting in ß°-thal. have been reported, but are more rare.
- Unequal crossing over between two homologous chromosome 11 results in production of Hb Lepore and Hb anti-lepore.

In summary, the single gene disorders are a large group which result from mutations in or around the genes and alter either the structure or the stability or the rate of synthesis of a protein or enzyme and thus results in a genetic disease.

Table 1: Types of mutations producing single gene disorders

1. Nucleotide substitutions (Point mutations) - Missense mutation . A.A. substitution - Nonsense mutation . Premature termination - RNA splicing mutations

. Intron/exon splice site or cryptic site mutation

2. Small deletions and insertions

- Codon deletion or insertion
  . Multiple of 3 bases deleted or added
- Frameshift mutation
  . No. of base deleted or added are not multiples of 3.

3. Gene deletion and duplications

- By unequal crossing-over or other mechanisms

4. Insertion of repeated elements

- Interrupts coding sequences

Table 2: Molecular pathology of some single gene disorders

Disorder	Molecular pathology
- Growth hormone deficiency	Deletion
- Factor VIII deficiency	Deletion, nonsense mutation
- Factor IX deficiency	Deletion, single point mutation
- Antithrombin III deficiency	Deletion
- Osteogenesis imperfecta	Partial deletion of collagen gene
- α1-AT deficiency	Single base change
- Lesch-Nyhan Syndrome	Deletion of HGPRT gene
- PKU	Point mutation
- Familial hypercholesterolaemia	Partial deletion of LDL receptor gene
- Structural variants of Hb	Single base change, frameshift, insertion, deletion, chain termination
- Thalassaemia	Nonsense, frameshift, deletion, poly A site mutation, mRNA processing mutation, inversion, fusion, etc.