Biochemical Genetics of Collagen

Collagen Structure and Genes

COLLAGENS

Structural proteins in extracellular matrix

Composed of three polypeptide chains - α-chains

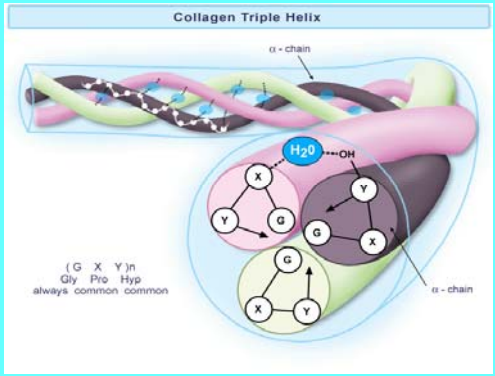
27 collagen types - 42 collagen genes

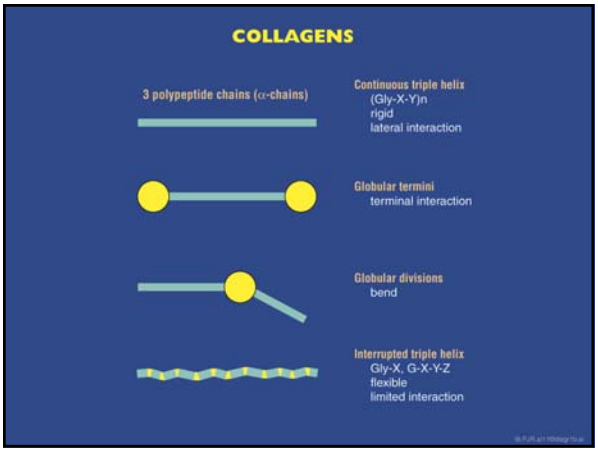
Exist as homotrimers or heterotrimers

Form triple helix along part of length

Triple helix may be continuous or interrupted

Form aggregates by association of triple helices



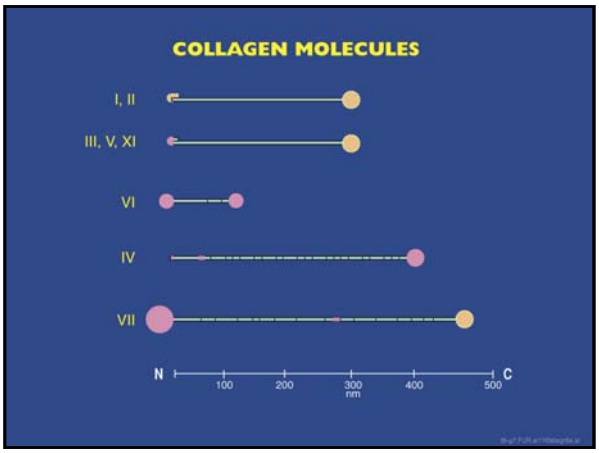


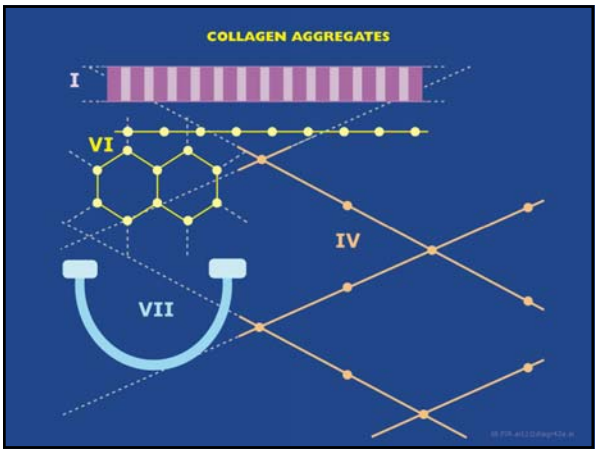
Collagen - Structure/Distribution

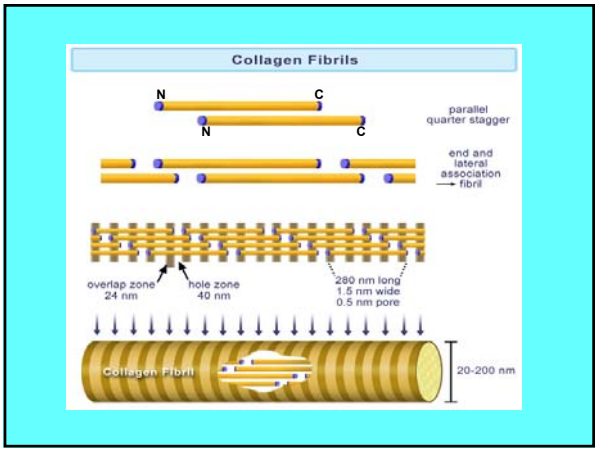
Type	Structure	Tissues
I	Fibril	Bone, dermis, ligament, tendon, aorta, sclera, cornea …….
II	Fibril	Cartilage, intervertebral disc, vitreous
III	Fibril	Aorta, intestine, uterus, dermis, ligament ……
IV	Network	Basement membrane
V	Fibril	Cornea, placenta, dermis, ligament ……
VI	Beaded filament	Ubiquitous
VII	Anchoring fibril	Basement membrane (epidermal, mucosal epithelial)

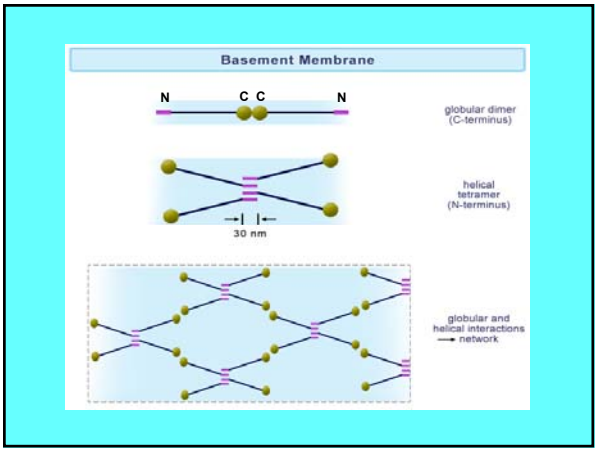
Collagen - Composition

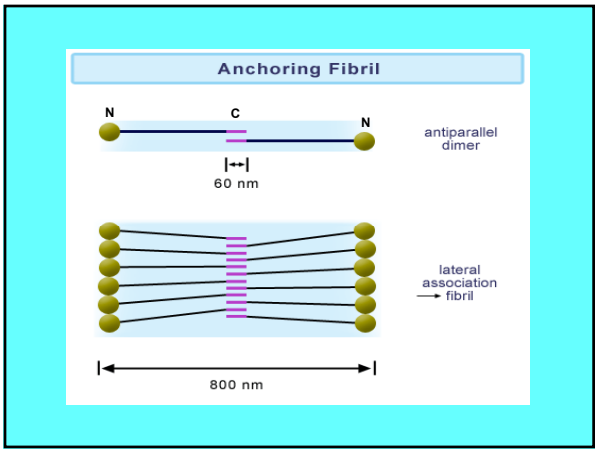
Type	Composition	Genes
I	$[\alpha1(I)]_2.\alpha2(I)$	2
	$[\alpha1(I)]_3$	
II	$[\alpha1(II)]_3$	1
III	$[\alpha1(III)]_3$	1
IV	$[\alpha1(IV)]_2.\alpha2(IV)$	6
	$[\alpha3(IV)]_2.\alpha4(IV)$	
	$[\alpha5(IV)]_2.\alpha6(IV)$	
V	$[\alpha1(V)]_3$	3
	$[\alpha1(V)]_2.\alpha2(V)$	
	$\alpha1(V).\alpha2(V).\alpha3(V)$	
VI	$\alpha1(VI).\alpha2(VI).\alpha3(VI)$	3
VII	$[\alpha1(VII)]_3$	1

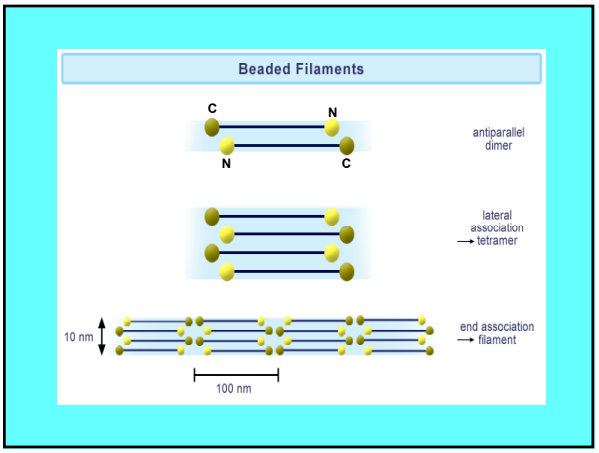

COLLAGEN MOLECULES
I, II
III, V, XI
VI
IV
VII
N
C
100
200
300
400
500
nm


COLLAGEN AGGREGATES
I
VI
IV
VII


Collagen Fibrils
N
C
N
C
parallel quarter stagger
end and lateral association → fibril
overlap zone 24 nm
hole zone 40 nm
280 nm long 1.5 nm wide 0.5 nm pore
Collagen Fibril
20-200 nm

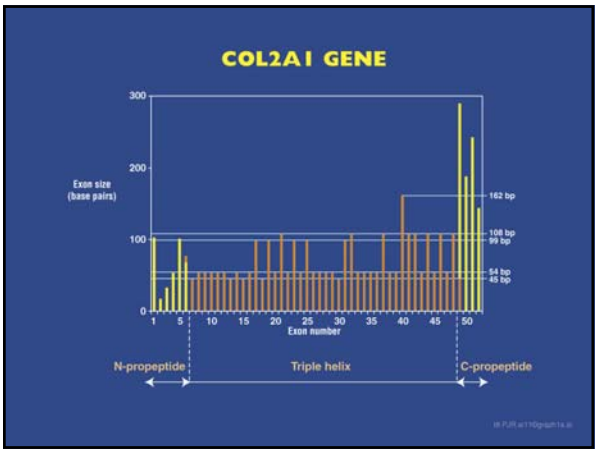

Basement Membrane
N
C C
N
globular dimer (C-terminus)
helical tetramer (N-terminus)
30 nm
globular and helical interactions → network

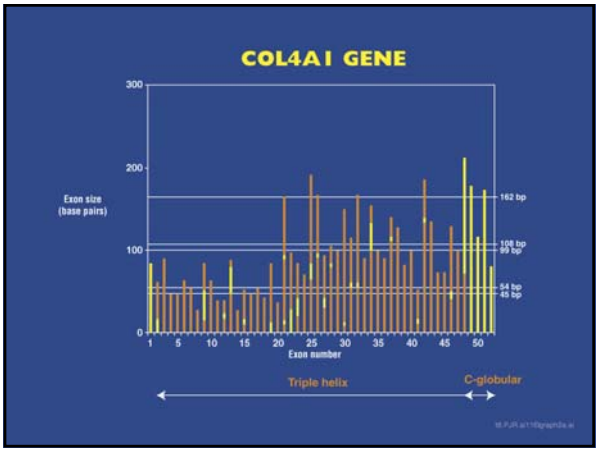

Anchoring Fibril
N
C
N
antiparallel dimer
60 nm
lateral association → fibril
800 nm


Beaded Filaments
C
N
N
C
antiparallel dimer
lateral association → tetramer
10 nm
end association → filament
100 nm

Collagen - Genes

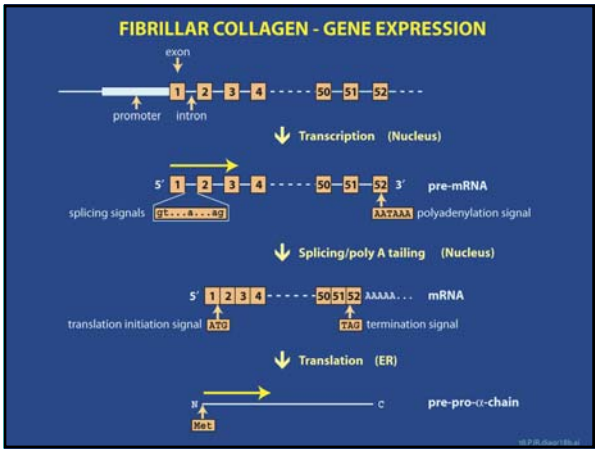
Type	Gene	Chromosome	Exons
I	COL1A1	17q21-22	51
	COL1A2	7q21-22	52
II	COL2A1	12q13-14	52
III	COL3A1	2q24-31	52
IV	COL4A1, COL4A2	13q33-34	52 / 47
	COL4A3, COL4A4	2q35-37	52 / 48
	COL4A5, COL4A6	Xq22	51 / 46
V	COL5A1	9q34	66
	COL5A2	2q24-31	54
	COL5A3	19p13	67
VI	COL6A1	21q22	36
	COL6A2	21q22	30
	COL6A3	2q37	43
VII	COL7A1	3p21	118

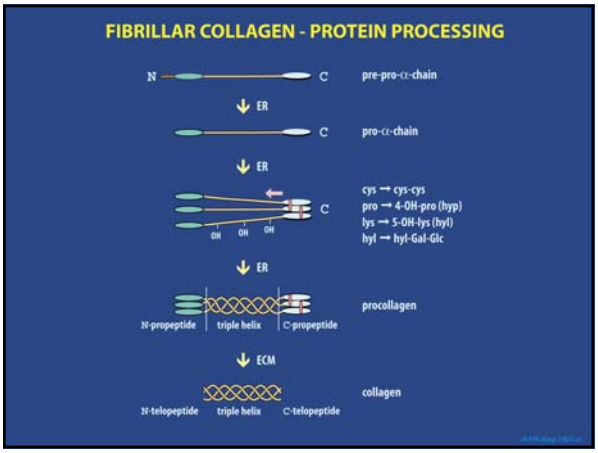


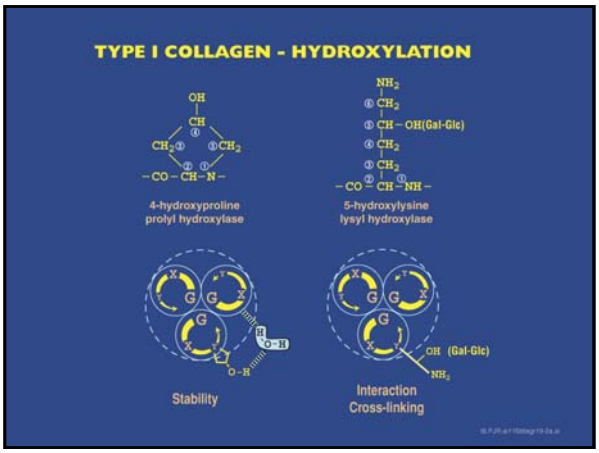


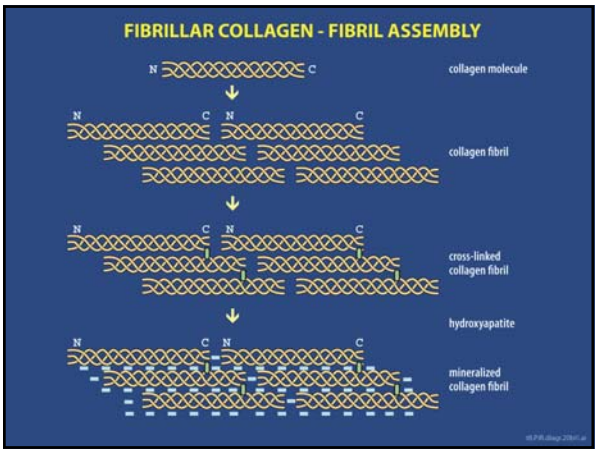
Biochemical Genetics of Collagen

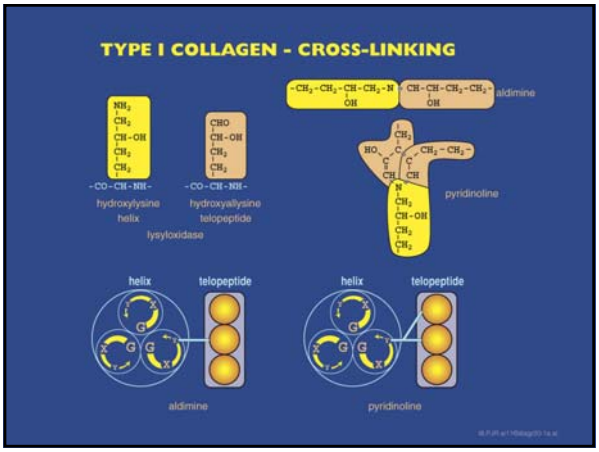
Collagen Synthesis and Disorders











Collagen - Gene Disorders

Type	Disorder	Tissue defect	Inheritance
I	Osteogenesis Imperfecta	Bone fragile	AD
	EDS type VII	Ligaments lax	AD
	Caffey syndrome	Bone hyperostosis	AD
II	Chondrodysplasias	Cartilage abnormal	AD
III	EDS type IV	Blood vessel rupture	AD
		Aortic aneurysm	
IV	Alport syndrome	Glomerlular BM	AD, AR, XL
		Hematuria	
V	EDS type I/II	Ligaments lax	AD
VI	Bethlem myopathy	Muscle dystrophic	AD
VII	Epidermolysis bullosa	Skin blisters	AD, AR

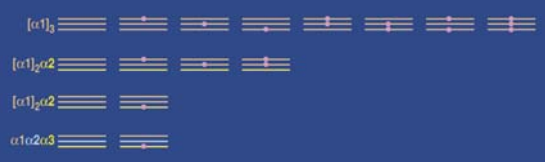
Collagen - Synthesis Disorders

Enzyme/Cofactor	Disorder	Inheritance
Prolyl 4-hydroxylase		
Prolyl 3-hydroxylase / CRTAP	OI type VIII / VII	AR
Lysyl hydroxylase	EDS type VI	AR
Vitamin C	Scurvy	Acquired
Procollagen-C-proteinase		
Procollagen-N-proteinase	EDS type VII	AR
Lysyl oxidase	EDS type V	AR
Copper/transporter	Menkes syndrome	Acquired / XR
Lathrogen	Lathyrism	Acquired

Types of Inheritance

Type	Features
Dominant	Mutation in 1 allele causes disease - heterozygous
	Common for structural proteins
	Haploinsufficiency – no production of abnormal protein, but only half of normal protein
	Dominant negative effect - production of abnormal protein that can impair function of normal protein
Recessive	Mutation in both alleles needed for disease
	Common for enzymes, rare for structural proteins
	Homozygous (consanguineous) or compound heterozygous
	Absence of normal protein, or only abnormal protein

MUTANT COLLAGEN TRIMERS



Dominant negative effect
- On collagen molecule due to Gly mutation in α chain
- On collagen fibril due to presence of mutant molecules

Disease Inheritance

Affected	Sibling	Offspring
Dominant inherited	50%	≥ 50%
Germ line mosaic	low	≥ 50%
Recessive inherited	25%	low
X-linked inherited	M → M, 0% F, 100% (carrier if R, mild if D) F → M, 50% F, 50% (carrier if R, mild if D)	

Biochemical Genetics of Collagen

Osteogenesis Imperfecta and Gene Mutations

Osteogenesis Imperfecta

Type	Features
I	Mildest form No fractures at birth, begin with ambulation/falls No skeletal deformity, normal height
II	Most severe form, perinatal lethal Multiple intrauterine fractures
III	Most severe form compatible with life Fractures at birth, frequent throughout life Progressive skeletal deformity, very short stature
IV	Between I and III Most variable phenotype, moderate to severe

Osteogenesis Imperfecta

Type I

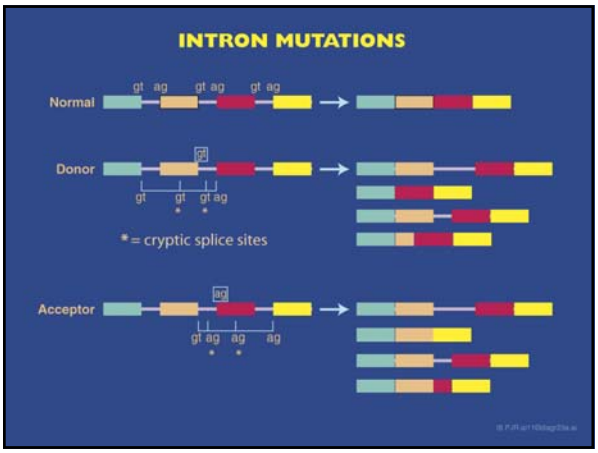
Type III

POLYMORPHISMS

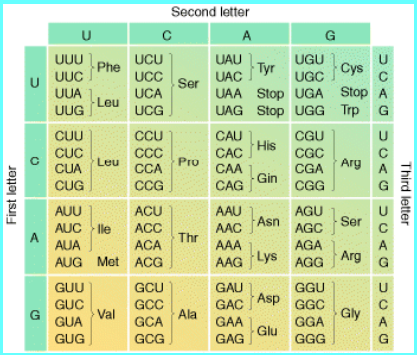
promoter:	not influence gene expression
intron:	not influence splicing
exon:	not change AA (Gly, GGN; Pro, CCN) conservative AA change (CTA → ATA; Leu → Ile)
variation:	alter expression/function but acceptable
intergenic:	90% of genome

Exon Mutations

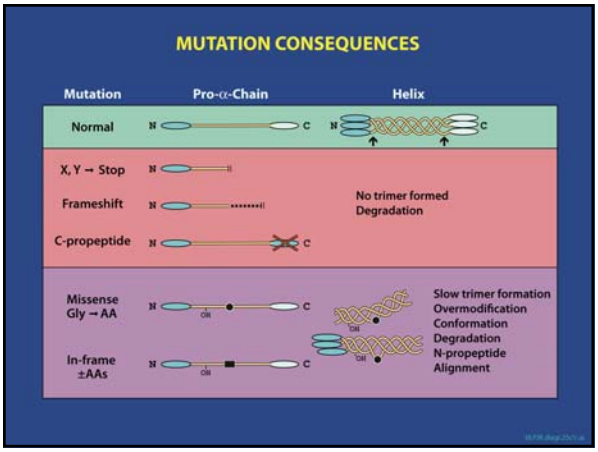
Normal		GGA AAG CCT	
Substitution	Missense	CGA AAG CCT	Change AA
	Nonsense	GGA TAG CCT	AA → stop
	Polymorphic	GGT AAG CCT	No AA change
Insertion	In-frame	GGA CAT AAG CCT	Add AA
		GGA AAC ATG CCT	Add/change AA
		GGA ACA TAG CCT	Generate stop
	Frameshift	GGA CAA AGC CT	Change AA sequence
Deletion	In-frame	GGA AAG CCT	Remove AA
		GGA AAG CCT	Remove/change AA
	Frameshift	GGA AAG CCT	Change AA sequence

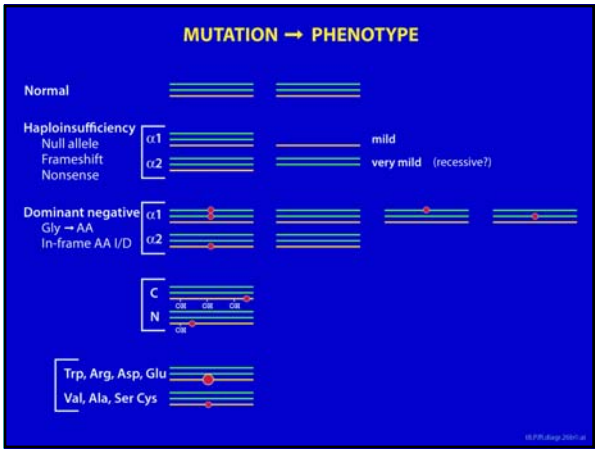


Glycine and Nonsense Mutations



From An Introduction to Genetic Analysis





Phenotype Modulation

Milder Phenotype

- Mutant mRNA unstable
- Mutant pro-α-chain unstable
- Helix tolerant of mutation (Gly-Pro-Hyp v Gly-X-Y)
- → Less abnormal collagen

More Severe Phenotype

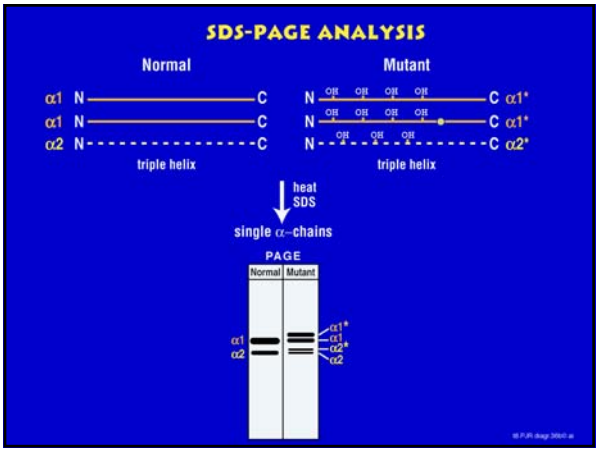
- Protein suicide
- Helix essential for interaction in ECM
- Other gene/metabolic abnormality
- → Less bone

Biochemical Genetics of Collagen

Analysis of Gene Mutations

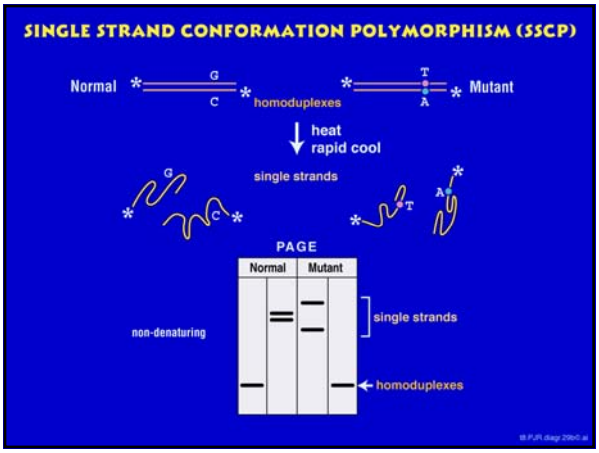
Mutation Analysis

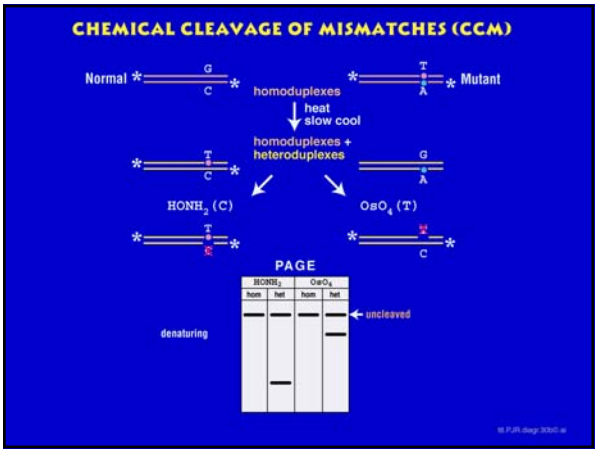
	gDNA	mRNA	Protein
Stability	Stable	Unstable	Stable
Abundance	1:1	<1:1	<<1:1
Source	Blood	Bone/skin	Bone/skin
Analysis	Easy	Easy	Difficult
Splicing	Deduced	Proven	Proven

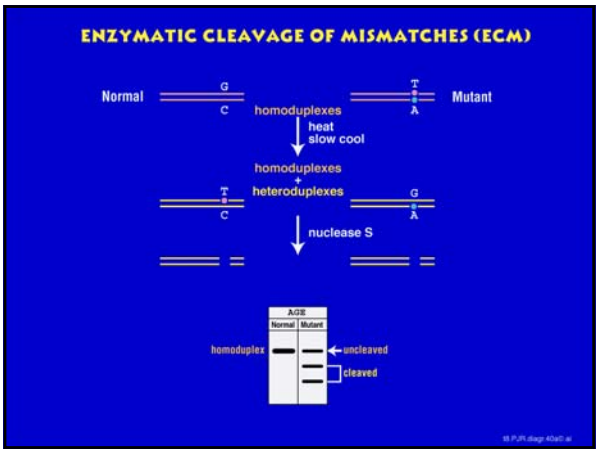


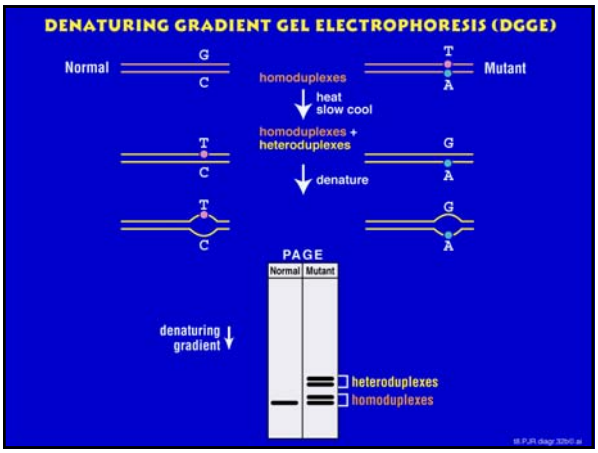
Mutation Screening

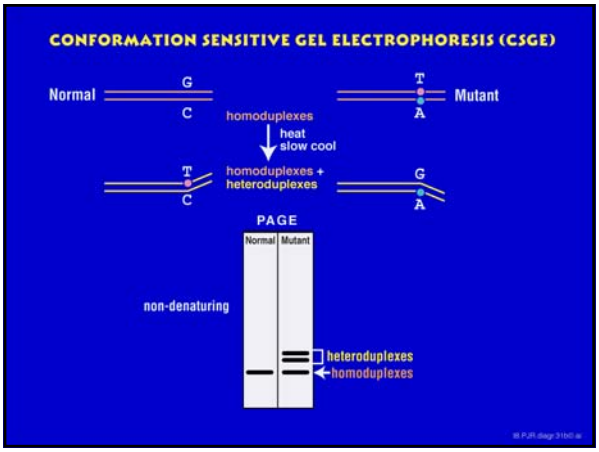
	SSCP	CCM	ECM	DGGE	CSGE
Size	200 bp	2 kbp	500 bp	500 bp	500bp
DNA	single	duplex	duplex	duplex	duplex
Detection	radioactive	radioactive	UV	UV	UV
Ease	simple	tedious	simple	complex	simple
Cost	cheap	cheap	expensive	expensive	cheap
Reliability	<90%	100%	100%	100%	<90%





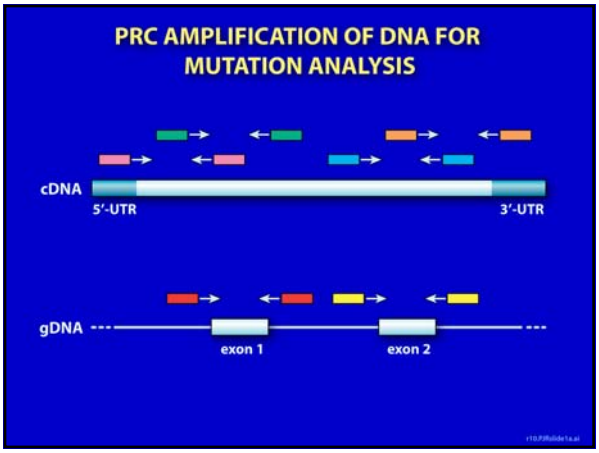


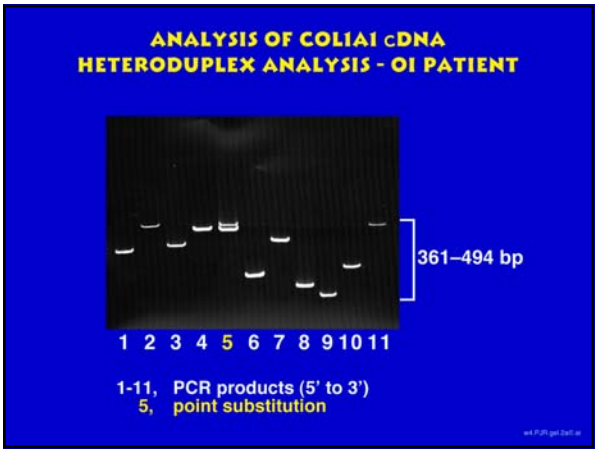


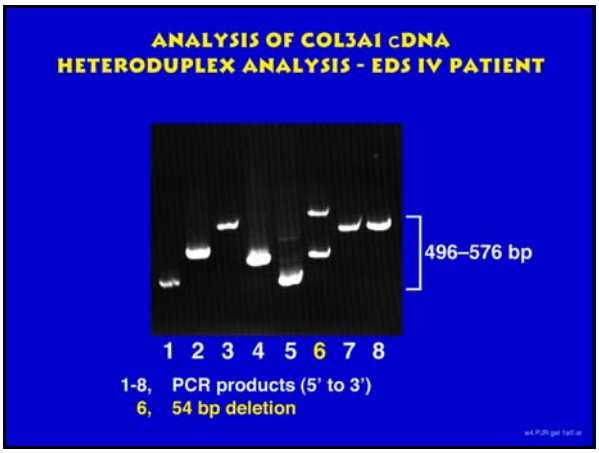


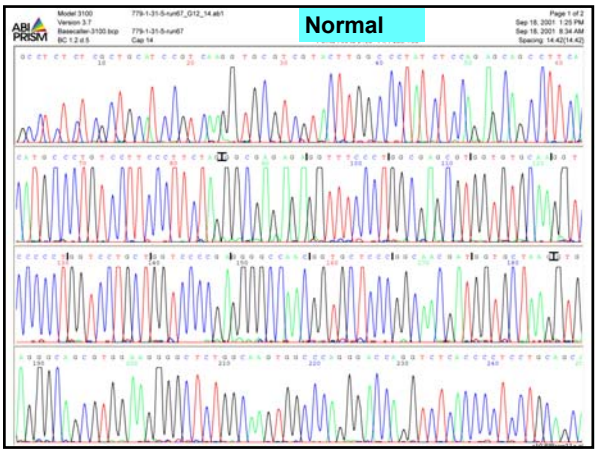
OI Mutation Analysis

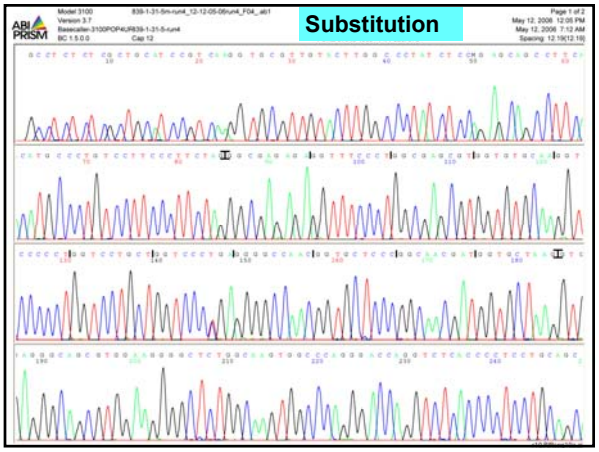
Blood	200 μl, fresh anticoagulated (EDTA)
	10 μg gDNA, less if frozen/coagulated blood
Skin	4 mm punch biopsy, 10^6 fibroblasts
	10 ug RNA, reverse transcribe to cDNA
PCR	50 ng DNA/20 μl
CSGE	4 μl, 300-500 bp
	~70% mutations identifiable by CSGE
Sequence	0.5 μl (distinguish mutation from polymorphism)

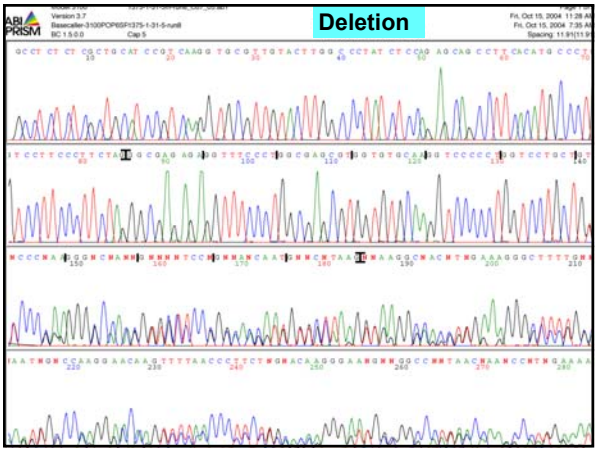












Molecular Diagnosis

Affected child	blood/tissue	confirm diagnosis
		predict outcome
		genetic counselling
Siblings	blood/tisue	check if affected
Parents	blood/tissue	check inheritance
	sperm	germ line mosacism
Fetus	AF or CVS	prenatal diagnosis
Blastocyst		pre-implantation for IVF