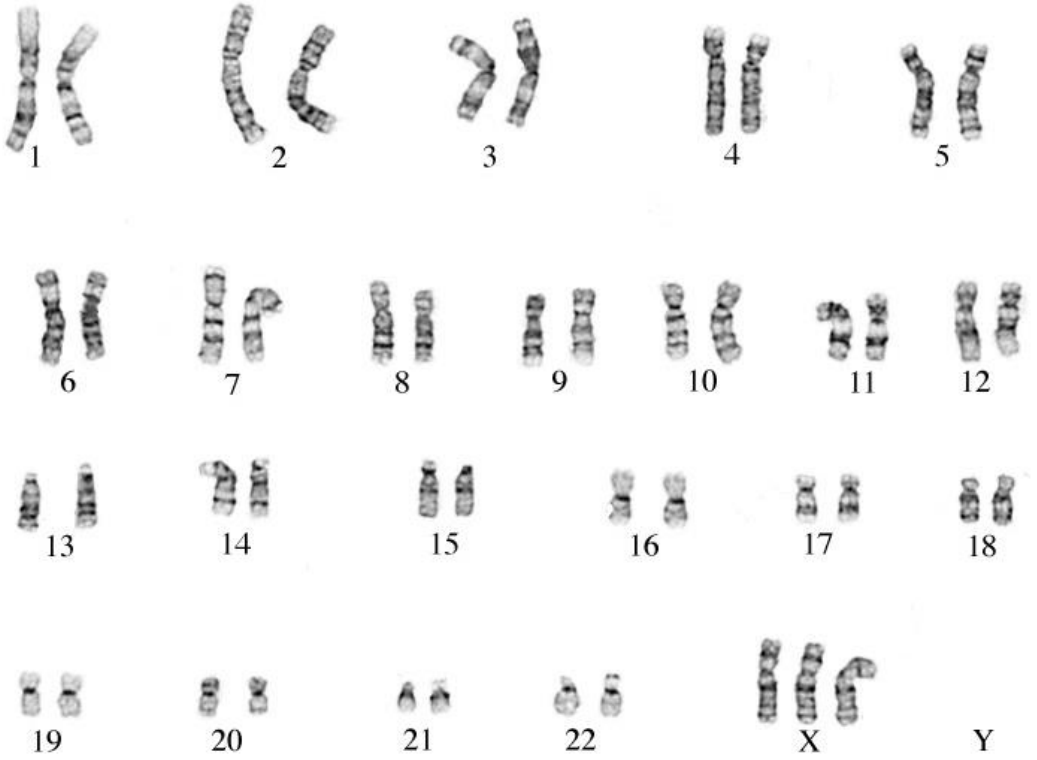
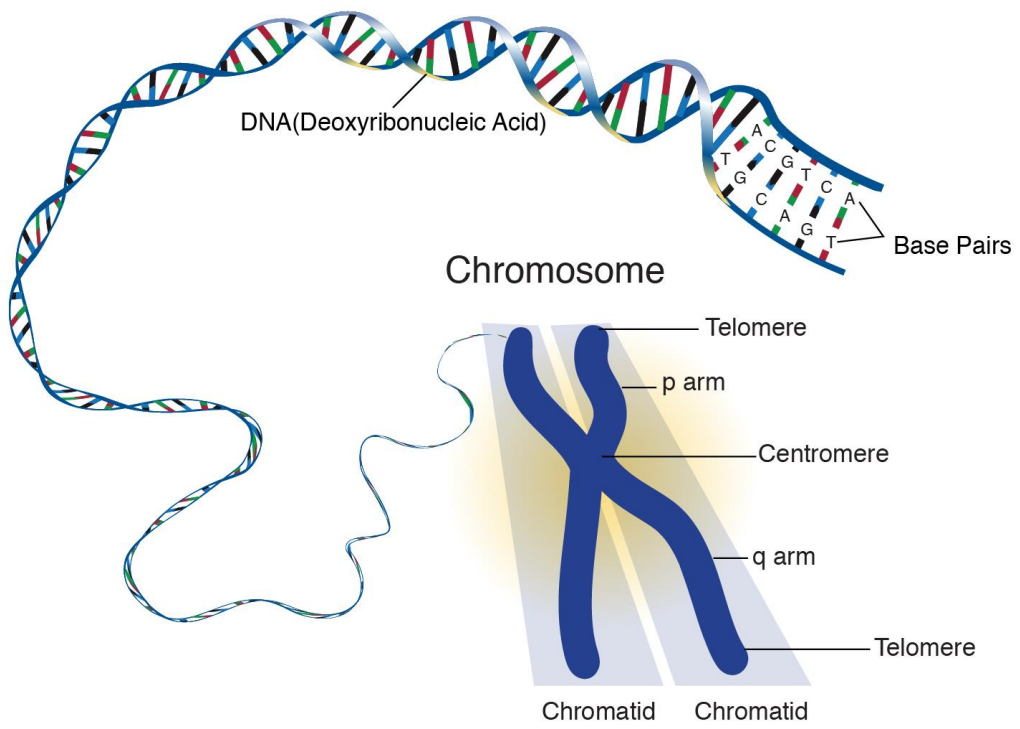
BIOL3120 –Human Genetics and Evolutionary Medicine

Nucleotide Mutations

4	Heritability and Polygenics Chromosomal Mutations	Problem Set 2	Problem Set 2 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
5	Nucleotide Mutations Human Genetic Diversity and Evolution	Problem Set 3	Problem Set 3 (5%)	Explain the principles of evolutionary biology and their role in human health and disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
6	Genetic Testing Techniques GWAS	Problem Set 4	Problem Set 4 (5%)	Explain the importance of new techniques in human genetics for understanding human disease Solve problems in human genetics using appropriate analytical methods and a variety of up to date resources
Recess				

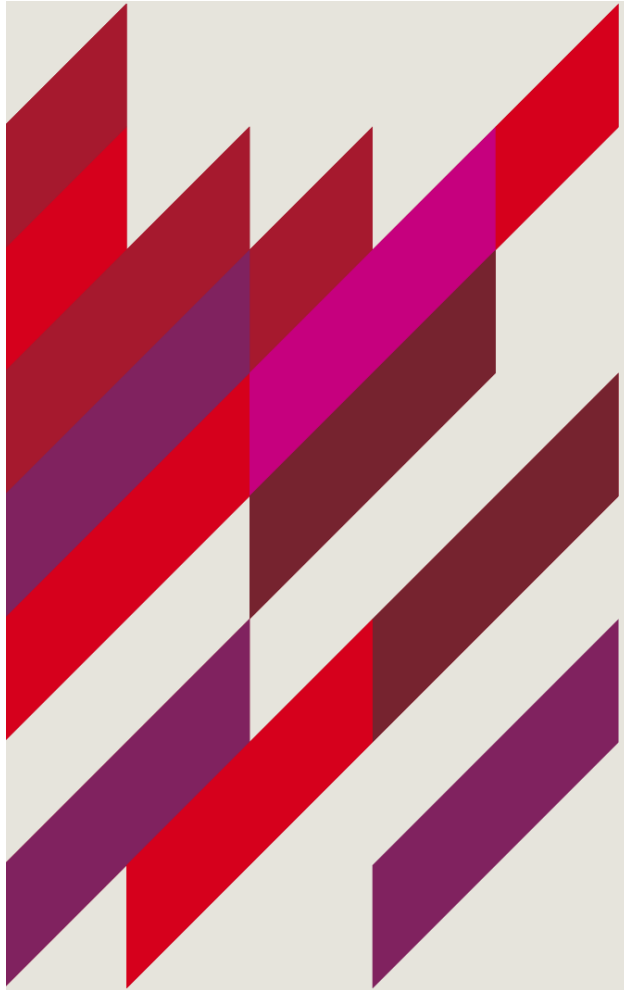
Quantity (chromosomes) vs quality (nucleotides)





BIOL3120 –Nucleotide Mutations

LEARNING OBJECTIVES



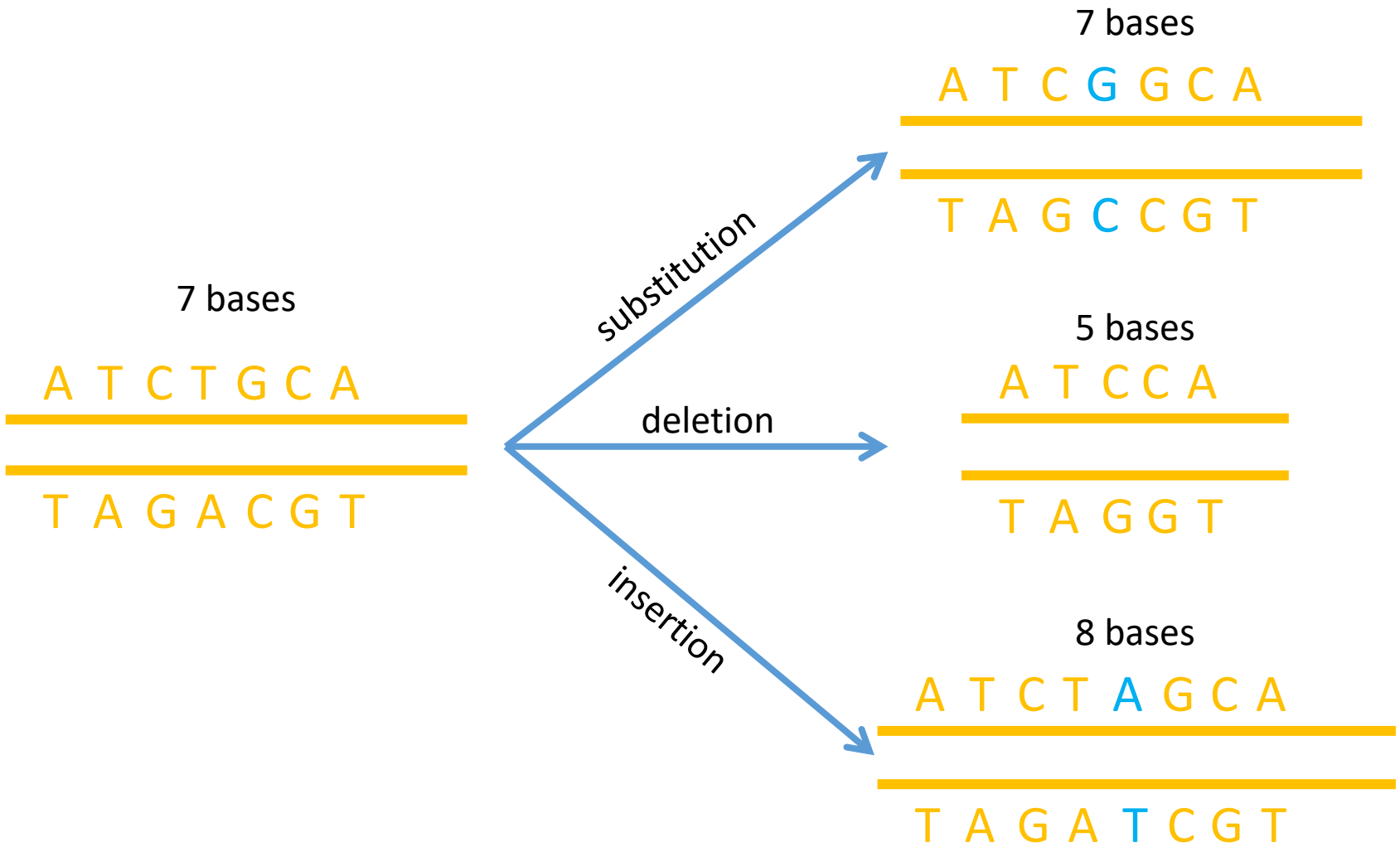
On successful completion of this lecture, you will be able to:

- Identify and name nucleotide mutations
- Interpret nucleotide variations
- Understand repeat expansions

Nucleotide Mutations

Nucleotide level variation

SINGLE NUCLEOTIDE AND INSERTION/DELETION VARIATION



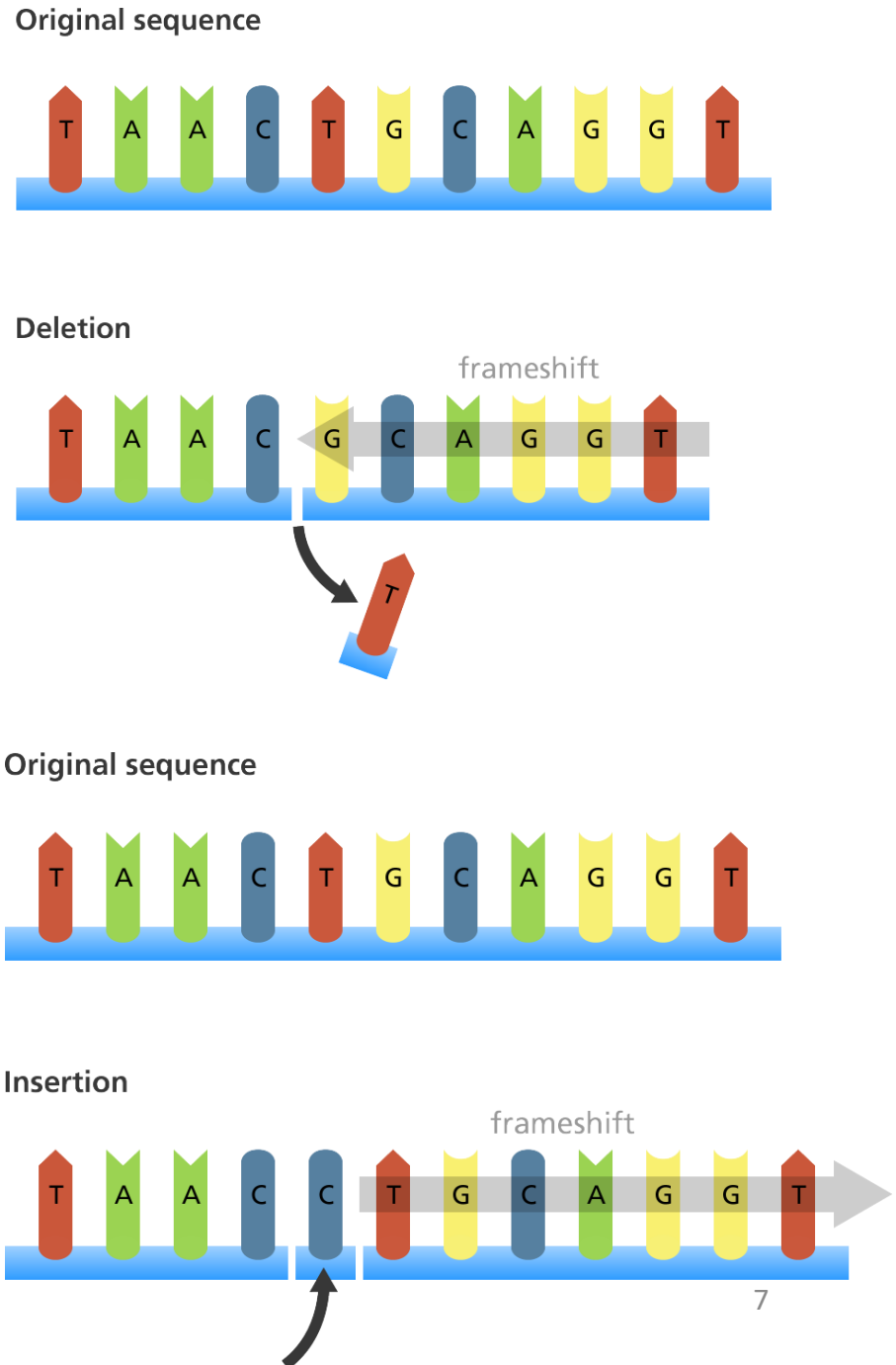
Deletions/Insertions

- Insertion or deletion of basepairs = Frameshift = all amino acids from there on affected

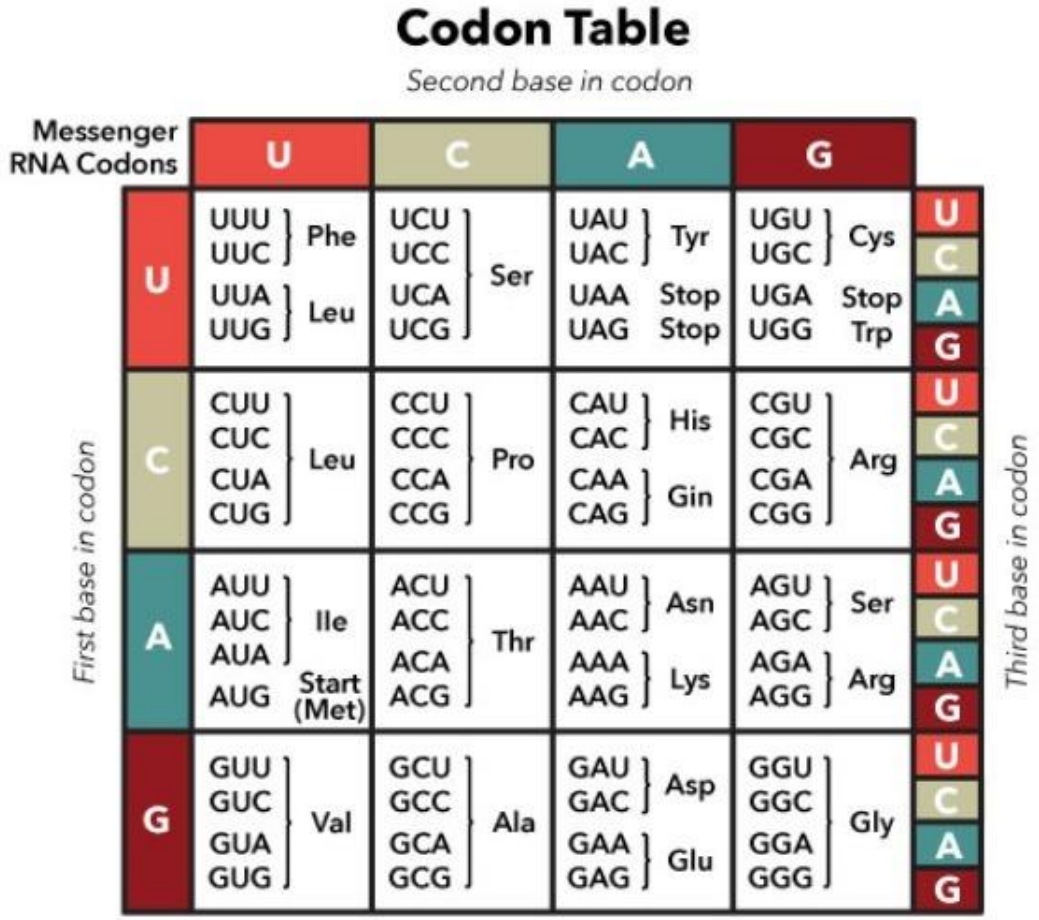
- Most likely a stop codon soon
  - 3/64 codons are stop codons = expect 1/23 codons to be a stop codon

- Very likely to impact function of protein

- Insertion/deletion of multiple of 3bp = no frameshift, extra or missing amino acids. May not impact function

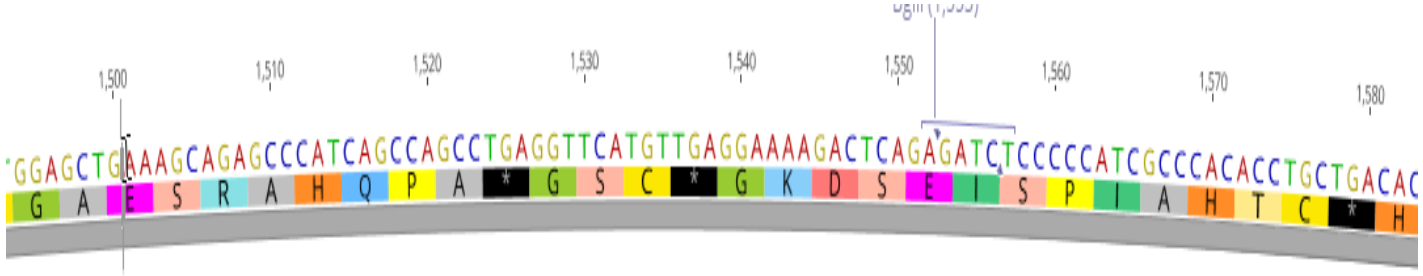


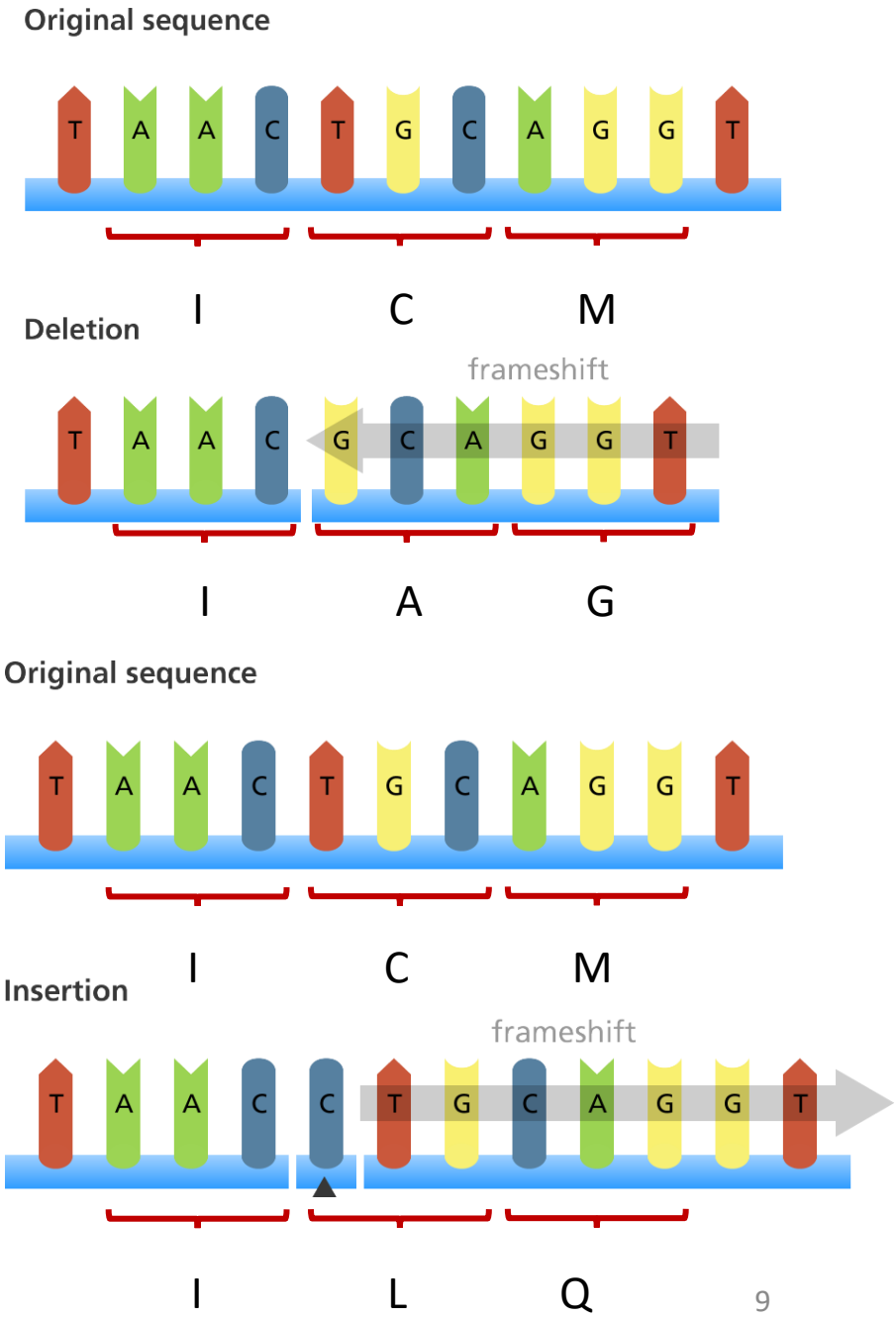
Frameshifts in coding region – what's changing?



Deletions/Insertions

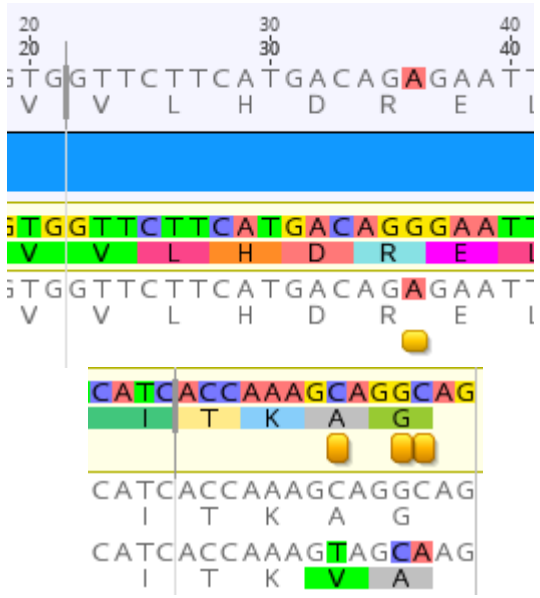
- Insertion or deletion of basepair = Frameshift = all amino acids from there on affected
- Most likely a stop codon soon
  - 3/64 codons are stop codons = expect 1/23 codons to be a stop codon
- Very likely to impact function of protein
- Insertion/deletion of multiple of 3bp = no frameshift, extra or missing amino acids. May not impact function

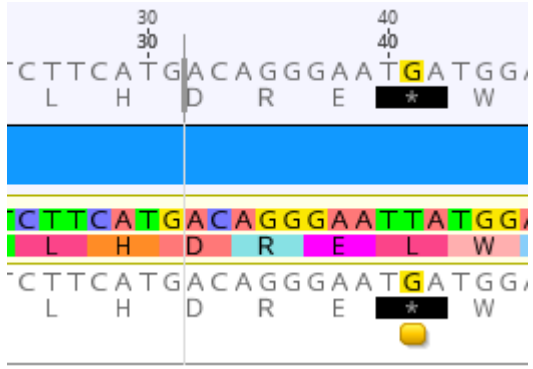




Substitutions/Point Mutations

- Silent or synonymous mutation
  - No change to amino acid
  - In some cases can cause disease
- Missense mutation
  - Changes to another amino acid
  - Many possibilities
- Nonsense mutation
  - Changes to stop codon
  - How much is the protein getting shortened?





Nomenclature for the description of sequence variations

- First, describe the reference sequence
  - DNA, RNA or Protein
- Then describe the location
  - State the basepair or amino acid where the mutation has occurred
- Then describe the kind of mutation
  - Was is a substitution, deletion, stop codon, frameshift, etc.

Indicate the reference sequence:

DNA		
	coding DNA	c.
	genomic DNA	g.
	mitochondrial DNA	m.
RNA		r.
Protein		p.

Code:

substitution (for bases)	>
range	_
more change in one allele	;
more transcripts / mosaicism	,
uncertain	()
allele	[ ]
deletion	del
duplication	dup
insertion	ins
inversion	inv
conversion	con
extension	ext
stop codon	X
frame shift	fsX
opposite strand	o
translocation	t

http://atlasgeneticsoncology.org/Educ/NomMutID30067ES.html

Type of variation/mutation:

Substitution	
c.123A>G	on cDNA, A in 123 is replaced by G
p.P252R	on protein, proline (P) replaced by arginine (R)
Deletion	
c.546delT	deletion of T in 546
c.586_591del	for six bases deleted
p.F508del	deletion of phenylalanine (F) in 508
Duplication	
c.546dupT	duplication of T in 546
c.586_591dup	duplication of the segment 586 to 591
p.G4_Q6dup	duplication of the segment from glycine (G) in 4 to glutamine (Q) in 6
Insertion	
c.546_547insT	insertion of T between 546 and 547
c.1086_1087insGCGTGA	insertion of GCGTGA
p.K2_L3insQS	insertion of glutamine serine between lysine (K) in 2 and leucine (L) in 3
Inversion	
c.546_2031inv	segment 546 to 2031 inverted
Frameshift	
p.R83SfsX15	arginine (R) is the first amino acid changed, it is in position 83, it makes serine (S) instead, the length of the shift frame is 15, including the stop codon (X)

http://atlasgeneticsoncology.org/Educ/NomMutID30067ES.html

Interpreting nucleotide variations

Biggest problem in clinical genetics now: Is this variant likely to cause a problem? ‘variant classification’

- In the past -> phenotype first
  - problem first, find the shared mutation
  - All research

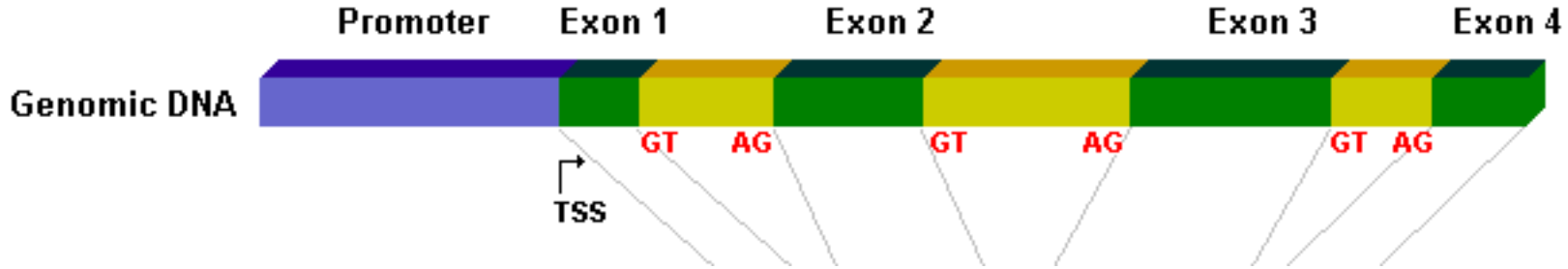
- Now -> lots of sequence data
  - find variants first, try to determine likelihood of any of them causing the problem
  - Mutations may be seen first in the clinic

Variant classification

Questions to ask:

- Where in the gene is the change?
- If in an exon, what is the amino acid change? Bioinformatics predictions of the outcome on protein function
- Has this variant been reported in literature? Has a similar variant been reported (same residue?)
- Is this variant seen in healthy populations? What's the frequency?
- Is this residue conserved across other species' version of this protein?

Substitution – where?



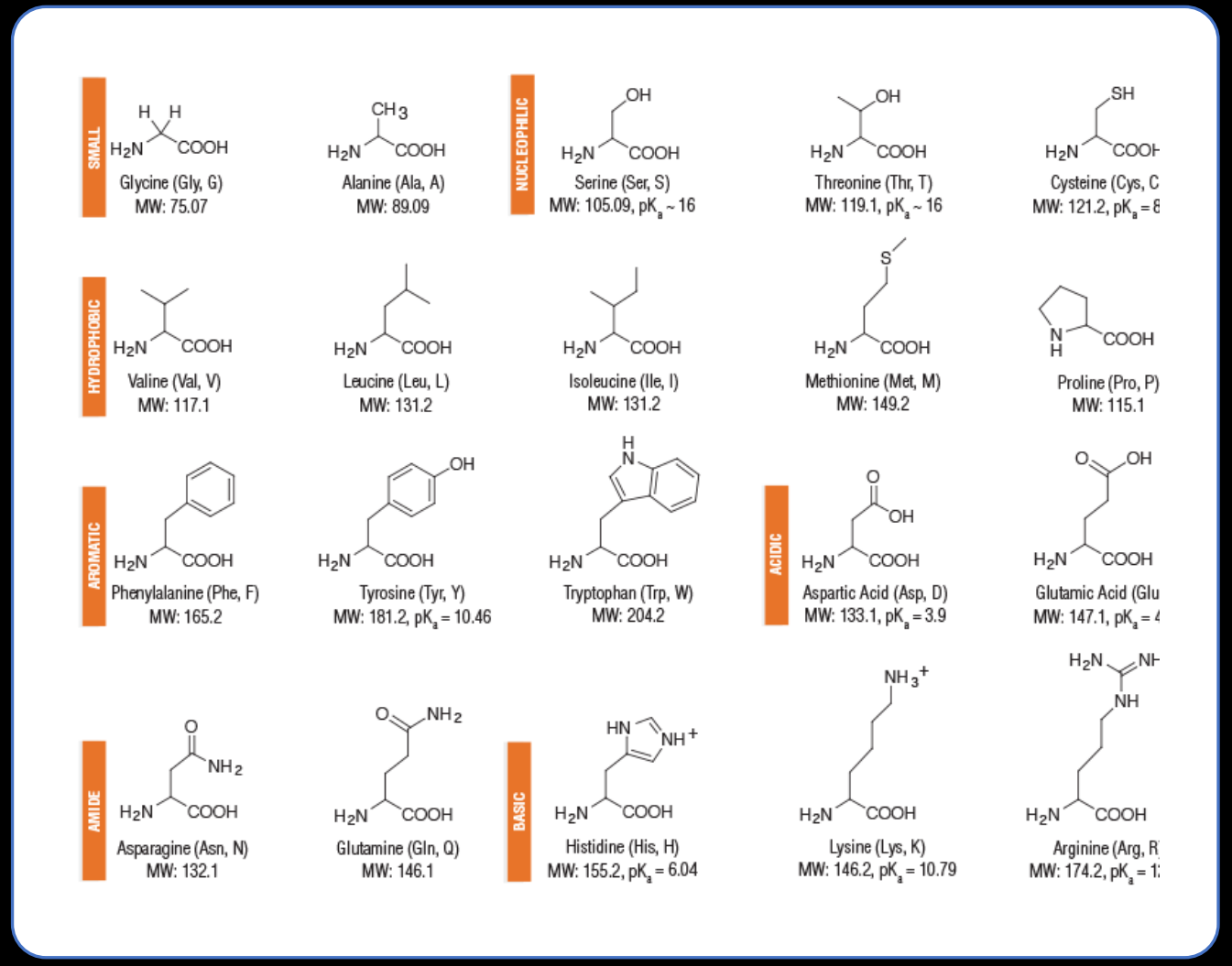
- Exon = can effect amino acid sequence
- Intron = within 8-10bp of exon may cause splicing issues
- Mutations within an exon near intron may affect splicing, even if it doesn't affect the amino acid (silent mutation)
- Promoter = maybe but far less likely
  - Might impact expression level

- Further upstream, promoter?

Missense mutations: How similar are the amino acids?

Protein prediction algorithms:

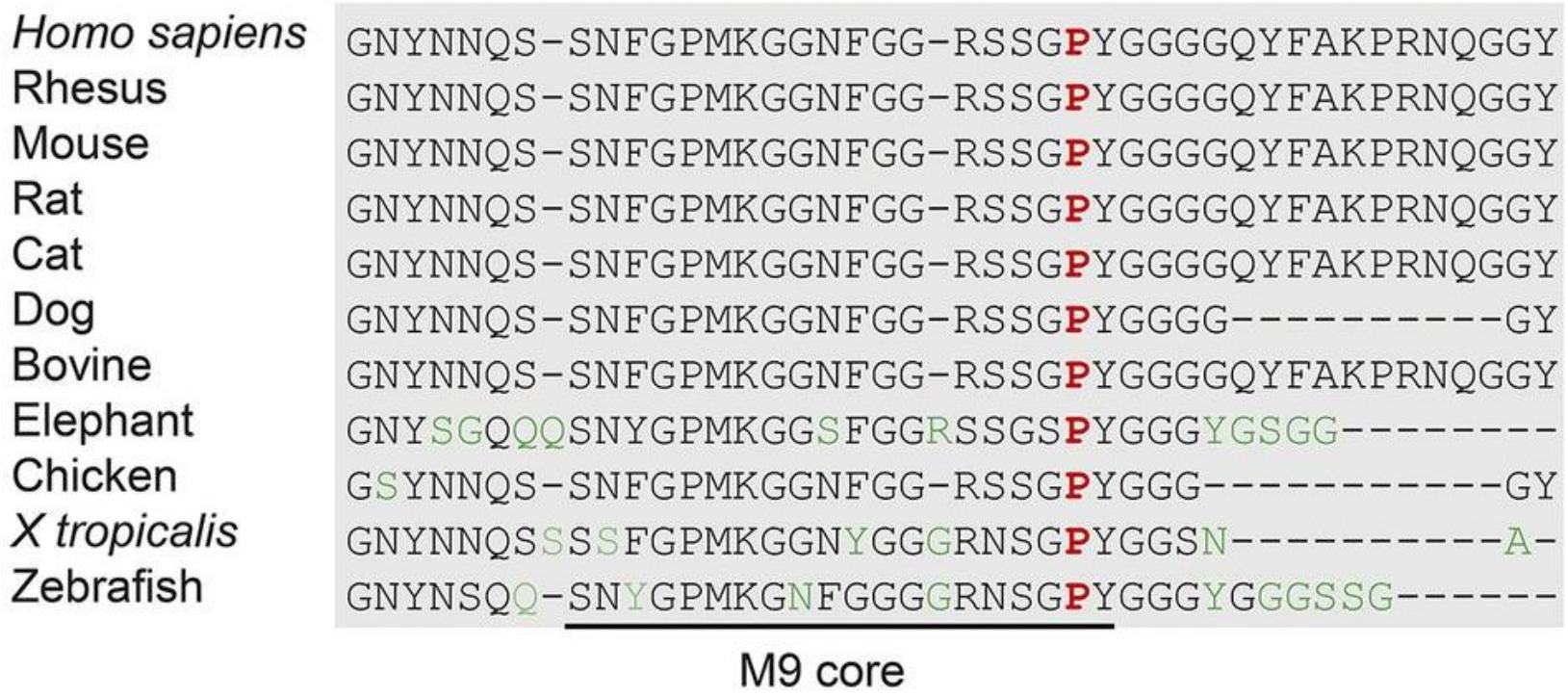
- SIFT
- Polyphen



Variant seen in healthy populations?

- Databases can provide frequency of this variant in healthy population
- ExAC (Exome Aggregation Consortium)
  - http://exac.broadinstitute.org/
- gnomAD
  - https://gnomad.broadinstitute.org/
- Consider dominant vs recessive ‘allowed’ frequencies

Missense mutations:
Is this residue conserved in other species' homologues of this protein?



Other things to consider:

- Functional studies?
- De novo (new mutation) or in parents?
- Segregate throughout family with disease?

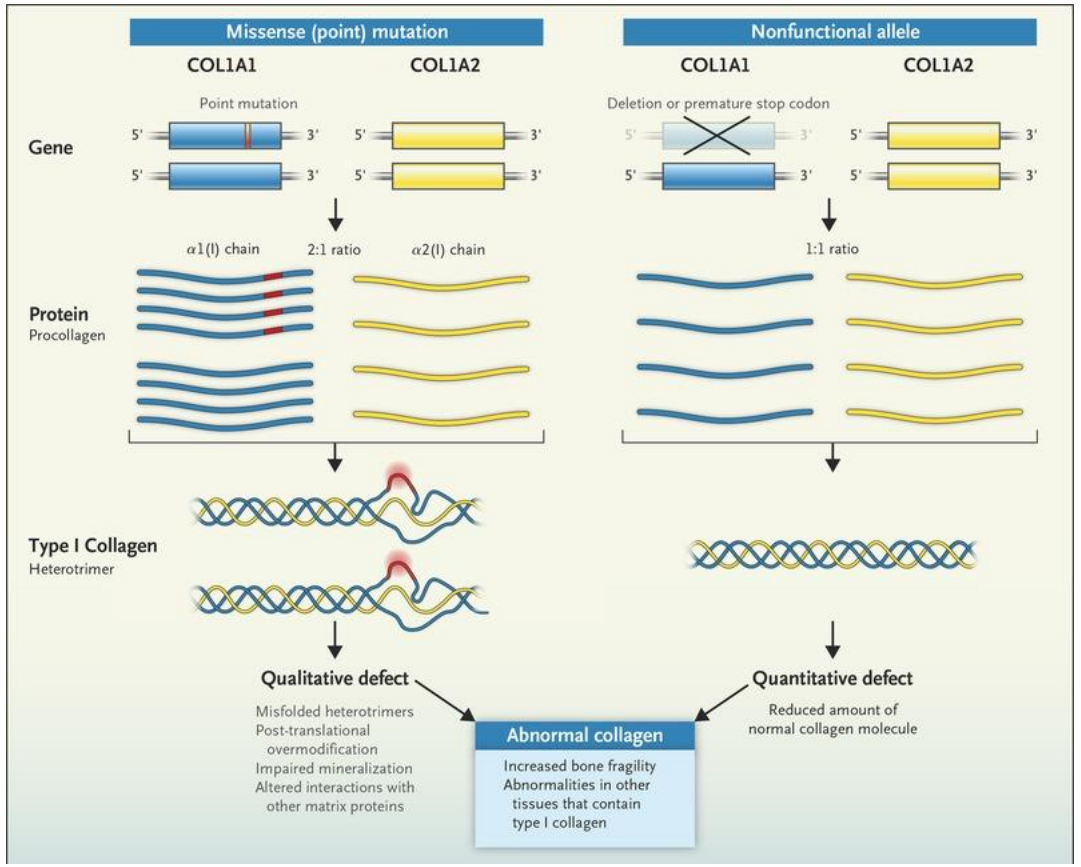
Why are diseases dominant or recessive?

Recessive: Having 1 normal allele gives normal phenotype

- 2 mutant alleles **= loss of function =** having no functioning protein

Dominant: 1 mutant allele causes condition

- **Gain of function** = new toxic function with mutant allele (majority of dominant conditions)
- **Haploinsufficiency** = the problem/condition is caused by not enough protein being made e.g. Marfan syndrome
- **Dominant negative** = the mutant protein product interferes with the function of the natural/normal protein product (eg osteogenesis imperfecta)



Repeat Expansions

Repeat expansions

- Some genes contain short repetitive sequences
- These repetitive sequence can expand and cause disease
  - >40 diseases, primarily affecting the nervous system
- Expanded trinucleotide repeat diseases were the first discovered and most frequent
  - Tetra-, penta-, hexa- and dodeca-nucleotide repeat expansions exsist
- Can experience anticipation
  - Decreasing age of onset or increased severity of disease across generations

Repeat expansions

Huntington's Disease (HD)

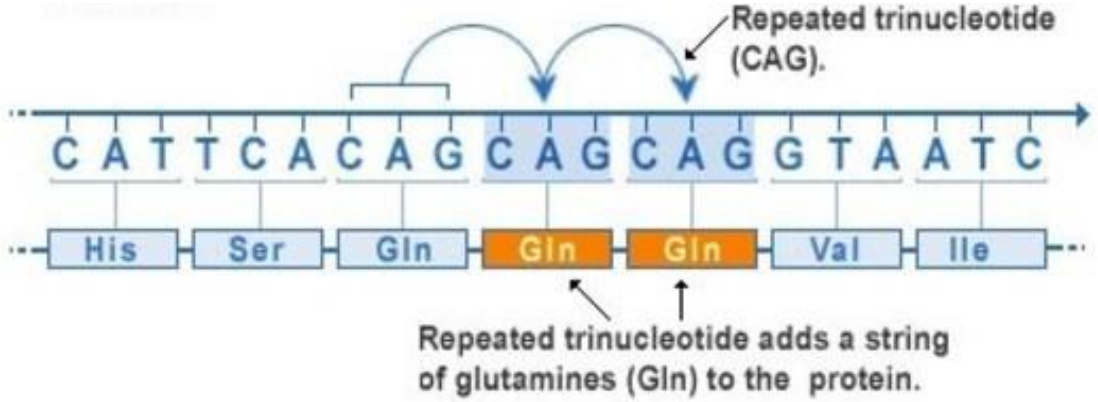
- Dominant inheritance
- HTT gene

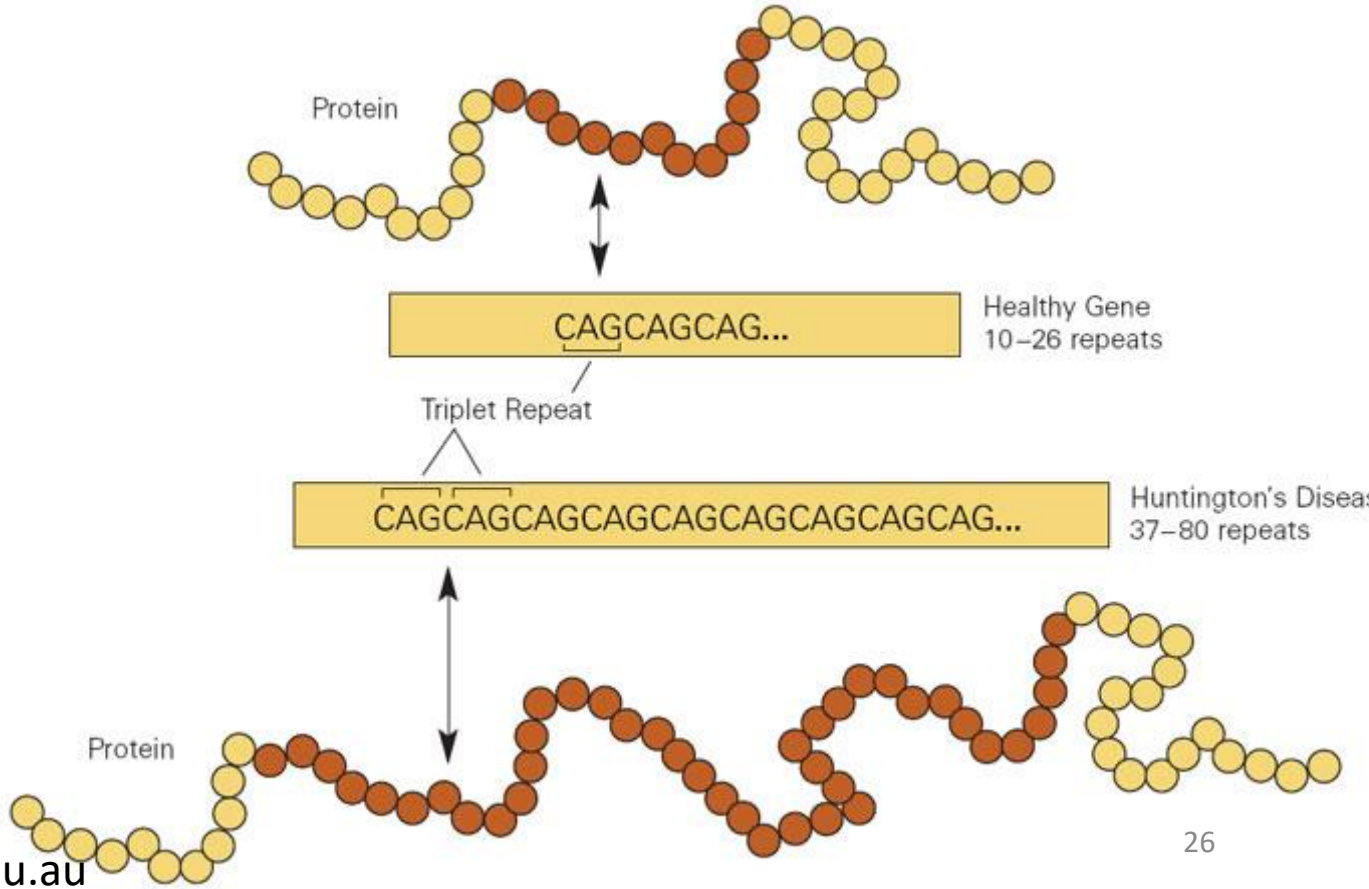
- Onset 30s or 40s
- Progressive loss of motor abilities, psychological deterioration, cognitive decline
- Death 15~20 years after first symptoms

- Variable repeat of 3 nucleotides in 1st exon - CAG

Repeat expansions in HD

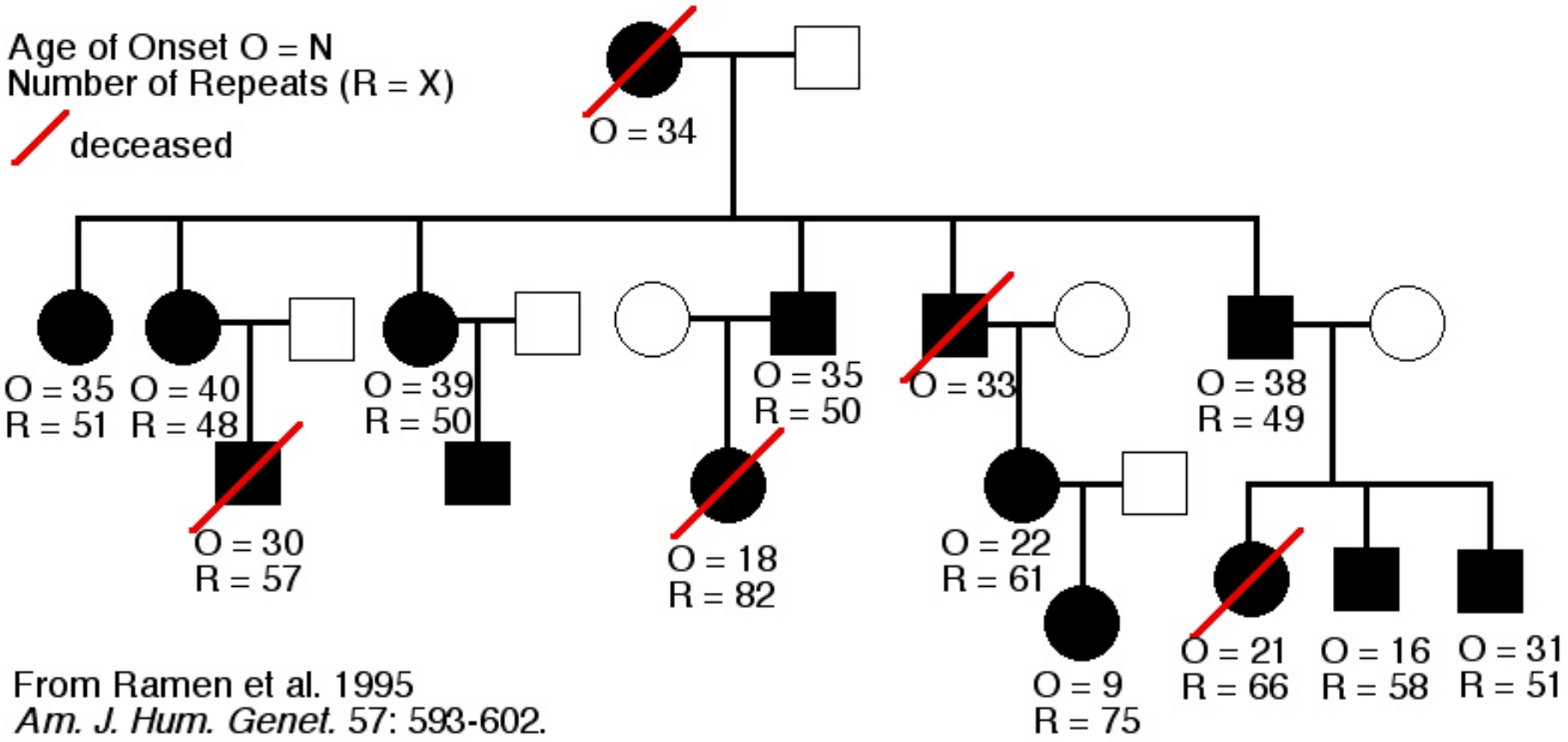
Number of CAG repeats*	What does it mean for you?
26 or less Normal range	You will not develop HD.
27 – 35 Intermediate range	You will not develop HD.
36 – 39 Increased risk range (reduced penetrance range)	You are likely to develop HD in your lifetime. However: • you might develop it at a late age • the condition might be less severe • you might not develop HD at all.
40 or more faulty gene HD range (full penetrance range)	You will almost certainly develop HD if you live a normal life expectancy.





From the centre for genetics education – www.genetics.edu.au

Anticipation in Huntington's Disease (HD)

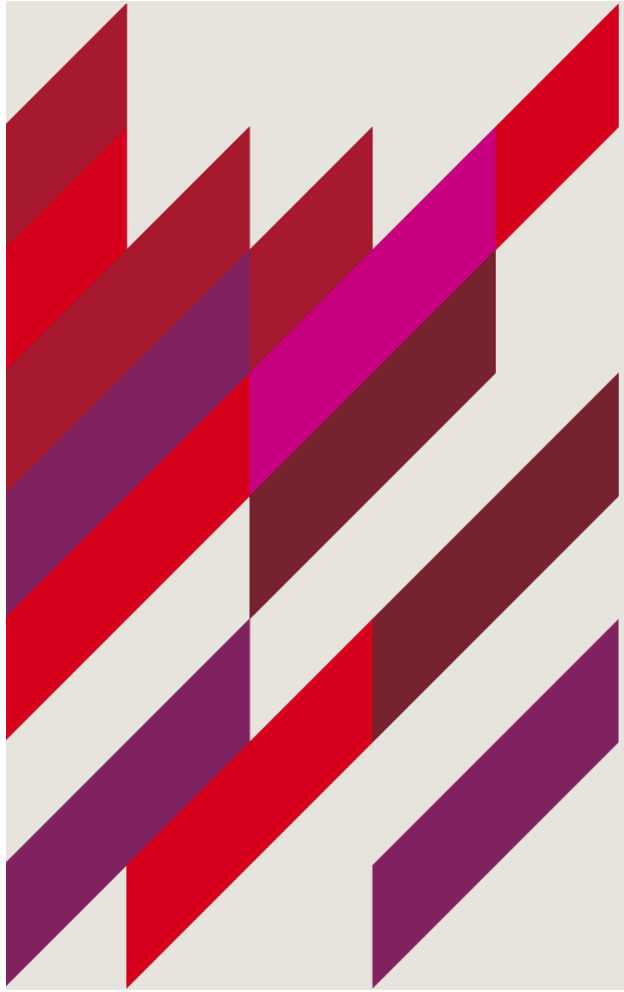


From Ramen et al. 1995
Am. J. Hum. Genet. 57: 593-602.

- Larger expansion = earlier onset

BIOL3120 –Nucleotide Mutations

LEARNING OBJECTIVES



On successful completion of this lecture, you will be able to:

- Identify and name nucleotide mutations
- Interpret nucleotide variations
- Understand repeat expansions