GENE MUTATIONS

Mutations

- Any change in the DNA sequence of an organism is a mutation.
- Mutations are the source of the altered versions of genes that provide the raw material for evolution.
- Most mutations have no effect on the organism, especially among the eukaryotes, because a large portion of the DNA is not in genes and thus does not affect the organism's phenotype.
- Of the mutations that do affect the phenotype, the most common effect of mutations is lethality, because most genes are necessary for life.
- Only a small percentage of mutations causes a visible but non-lethal change in the phenotype.

Gene mutations which affect only one gene

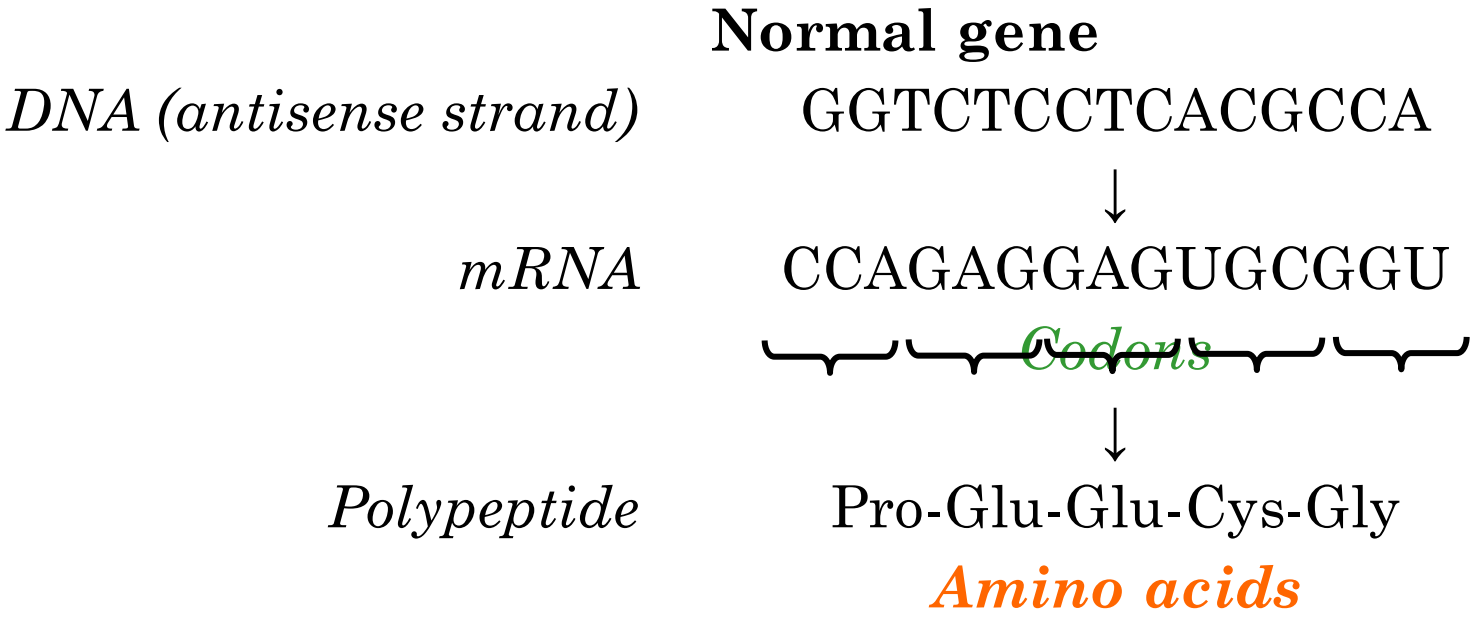
DNA sequence

Transcription ↓

mRNA sequence

Translation ↓

Polypeptide



The **antisense strand** is the DNA strand which acts as the template for mRNA transcription

Mutations: Substitutions

Normal gene

GGTCTCCTCACGCCA

↓

CCAGAGGAGUGCGGU

Codons

↓

Pro-Glu-Glu-Cys-Gly

Amino acids

Substitution mutation

GGTC**A**CCTCACGCCA

↓

CCAG**U**GGAGUGCGGU

↓

Pro-**Arg**-Glu-Cys-Gly

Substitutions will only affect a single codon

Their effects may not be serious unless they affect an amino acid that is essential for the structure and function of the finished protein molecule (e.g. sickle cell anaemia)

The genetic code is degenerate

A mutation to have **no** effect on the phenotype.

Changes in the third base of a codon often have no effect.

No change

Normal gene

GGTCTCCTCACGCCA

↓

CCAGAGGAGUGCGGU

Codons

↓

Pro-Glu-Glu-Cys-Gly

Amino acids

Substitution mutation

GGTCT**T**CTCACGCCA

↓

CCAGA**A**GAGUGCGGU

↓

Pro-**Glu**-Glu-Cys-Gly

Disaster

Normal gene

GGTCTCCTCACGCCA

↓

CCAGAGGAGUGCGGU

Codons

↓

Pro-Glu-Glu-Cys-Gly

Amino acids

Substitution mutation

GGTCTCCTCAC**T**CCA

↓

CCAGAAGAGUG**A**GGU

↓

Pro-Glu-Glu-**STOP**

Mutations: Inversion

Inversion mutations, also, only affect a small part of the gene

Normal gene

GGTCTCCTCACGCCA

↓

CCAGAGGAGUGCGGU

Codons

↓

Pro-Glu-Glu-Cys-Gly

Amino acids

Inversion mutation

GGTC**CT**CTCACGCCA

↓

CCAG**GA**GAGUGCGGU

↓

Pro-**Gly**-Glu-Cys-Gly

Mutations: Additions

A frame shift mutation

Normal gene

GGTCTCCTCACGCCA

↓

CCAGAGGAGUGCGGU

Codons

↓

Pro-Glu-Glu-Cys-Gly

Amino acids

Addition mutation

GGT**G**CTCCTCACGCCA

↓

CCA**C**GAGGAGUGCGGU

↓

Pro-**Arg-Gly-Val-Arg**

Mutations: Deletions

A frame shift mutation

Normal gene

GGTCTCCTCACGCCA

↓

CCAGAGGAGUGCGGU

Codons

↓

Pro-Glu-Glu-Cys-Gly

Amino acids

Deletion mutation

GGT**C/C**CTCACGCCA

↓

CCA**GG**GAGUGCGGU

↓

Pro-**Gly-Ser-Ala-Val**

Mutations of haemoglobin

- Haemoglobin is a tetramer = 2 α and 2 β-chains
- The genes for these polypeptides are found on different chromosomes
- The β-chain gene is found on chromosome 11
- The α-chain gene is found on chromosome 16
- The nucleotide sequences have been worked out
- Several inherited diseases occur on the β-chain, which contains 146 amino acids.

β HAEMOGLOBIN SENSE STRAND cDNA SEQUENCE

- cDNA (**complementary DNA**) is obtained by back-transcribing the mRNA used to translate the polypeptide
- So cDNA has no introns
- This is done using **reverse transcriptase** enzyme.

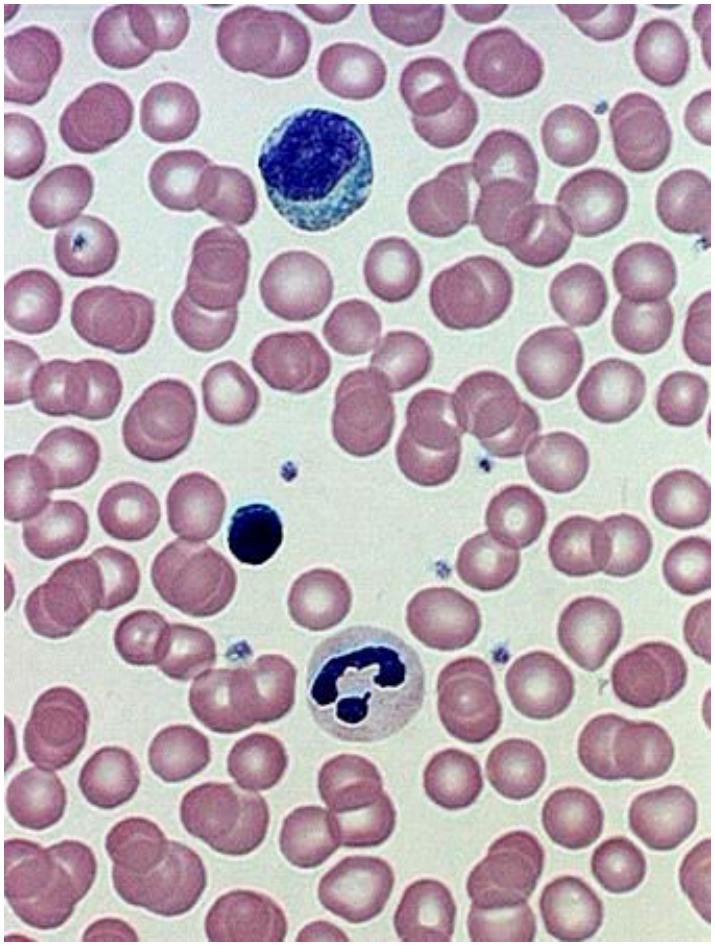
Methionine initiator

ATG GTG CAT CTG ACT CCT **GAG GAG** AAG TCT
GCC GTT ACT GCC CTG TGG GGC AAG GTG AAC GTG
GAT GAA GTT GGT GGT **GAG** GCC CTG GGC AGG
CTG CTG GTG GTC TAC CCT TGG ACC CAG AGG TTC
TTT GAG TCC TTT GGG GAT CTG TCC ACT CCT GAT
GCT GTT ATG GGC AAC CCT AAG GTG AAG GCT **CAT**
GGC AAG AAA **GTG** CTC GGT GCC TTT AGT GAT GGC
CTG GCT CAC CTG GAC AAC CTC AAG GGC ACC TTT
GCC ACA CTG AGT GAG CTG CAC TGT GAC AAG CTG
CAC GTG GAT CCT GAG AAC TTC AGG CTC CTG GGC
AAC GTG CTG GTC TGT GTG CTG GCC CAT CAC TTT
GGC AAA **GAA** TTC ACC CCA CCA GTG CAG GCT GCC
TAT CAG AAA GTG GTG GCT GGT GTG GCT AAT GCC
CTG GCC CAC AAG TAT CAC **TAA**

Nonsense terminator

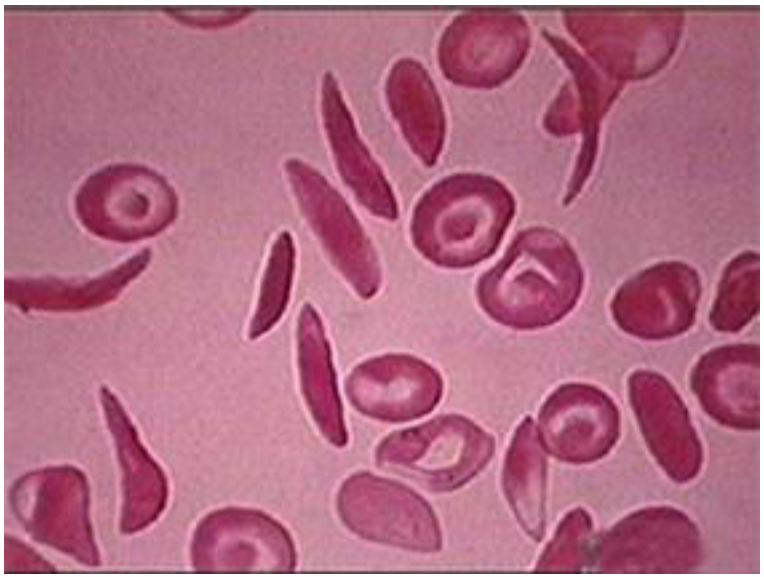
Mutation	Codon	Change to DNA sense strand	Change in Amino Acid
S (sickle cell anaemia)	**6**	**GAG to GTG**	**Glu to Val**
C (cooley's syndrome)	6	GAG to AAG	Glu to Lys
$G_{San\ Jose}$	7	GAG to GGG	Glu to Gly
E	26	GAG to AAG	Glu to Lys
$M_{Saskatoon}$	63	CAT to TAT	His to Tyr
$M_{Milwauki}$	67	GTG to GAG	Val to Glu
O_{Arabia}	121	GAA to GTA	Glu to Val

Sickle Cell Anaemia



Blood smear (normal)

Image Credit: http://lifesci.rutgers.edu/~babiarz/

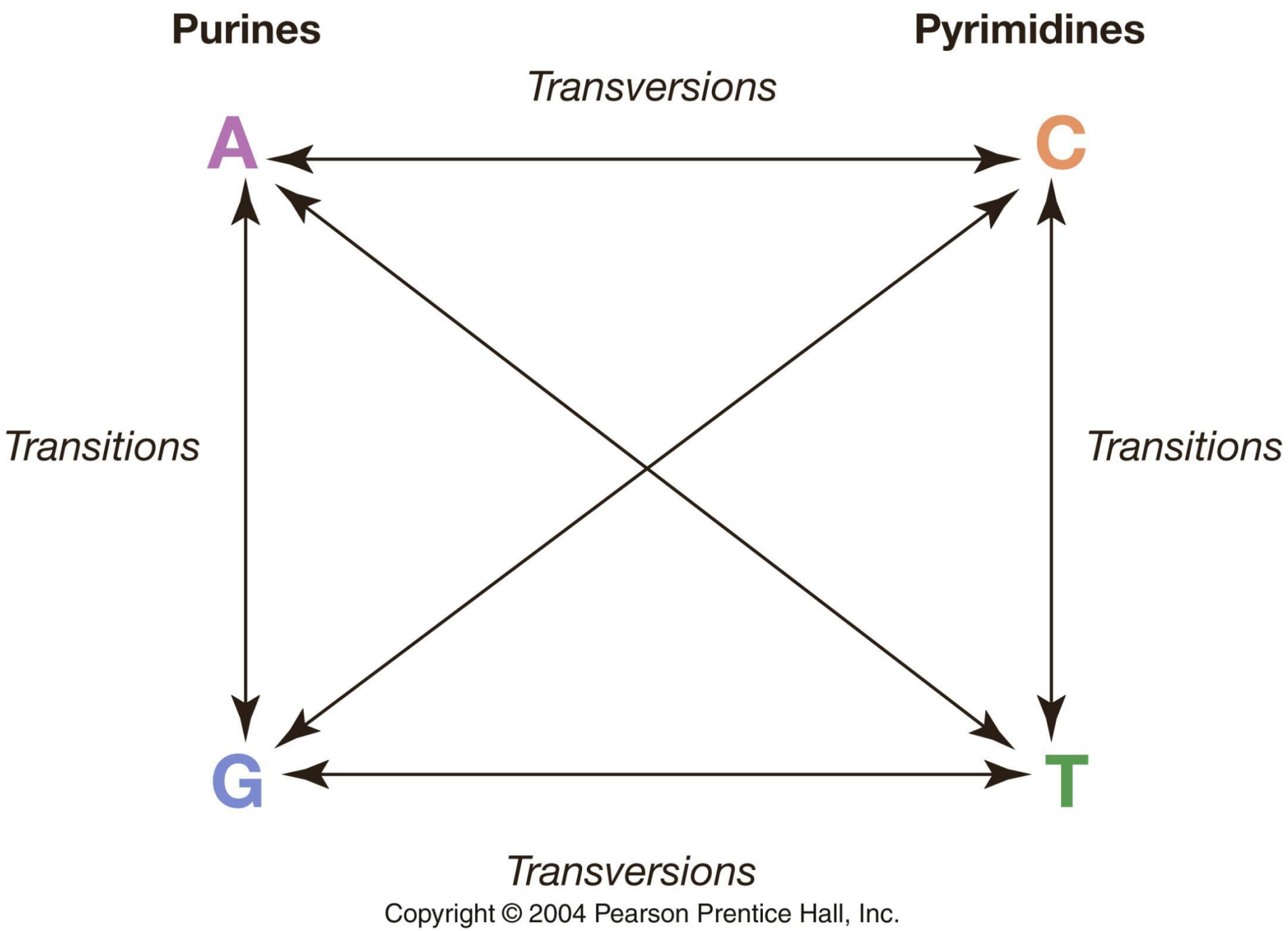


Sickle cell anemia

Image Credit: http://explore.ecb.org/

Types of DNA Change

- The simplest mutations are base changes, where one base is converted to another. These can be classified as either:
  - --"transitions", where one purine is changed to another purine (A -> G, for example), or one pyrimidine is changed to another pyrimidine (T -> C, for example).
  - "transversions", where a purine is substituted for a pyrimidine, or a pyrimidine is substituted for a purine. For example, A -> C.
- Another simple type of mutation is the gain or loss of one or a few bases.
- Larger mutations include insertion of whole new sequences, often due to movements of transposable elements in the DNA or to chromosome changes such as inversions or translocations.
- Deletions of large segments of DNA also occurs.



Types of mutations

- Not all mutations cause a change in amino acid coded for. These are called **silent mutations**.

- Mutations that do cause a change in amino acid are called **replacement mutations**.

(b)

First base	Second base: U	Second base: C	Second base: A	Second base: G	Third base
U	UUU Phenylalanine	UCU Serine	UAU Tyrosine	UGU Cysteine	U
	UUC Phenylalanine	UCC Serine	UAC Tyrosine	UGC Cysteine	C
	UUA Leucine	UCA Serine	UAA Stop	UGA Stop	A
	UUG Leucine	UCG Serine	UAG Stop	UGG Tryptophan	G
C	CUU Leucine	CCU Proline	CAU Histidine	CGU Arginine	U
	CUC Leucine	CCC Proline	CAC Histidine	CGC Arginine	C
	CUA Leucine	CCA Proline	CAA Glutamine	CGA Arginine	A
	CUG Leucine	CCG Proline	CAG Glutamine	CGG Arginine	G
A	AUU Isoleucine	ACU Threonine	AAU Asparagine	AGU Serine	U
	AUC Isoleucine	ACC Threonine	AAC Asparagine	AGC Serine	C
	AUA Isoleucine	ACA Threonine	AAA Lysine	AGA Arginine	A
	AUG Start (Methionine)	ACG Threonine	AAG Lysine	AGG Arginine	G
G	GUU Valine	GCU Alanine	GAU Aspartic Acid	GGU Glycine	U
	GUC Valine	GCC Alanine	GAC Aspartic Acid	GGC Glycine	C
	GUA Valine	GCA Alanine	GAA Glutamic Acid	GGA Glycine	A
	GUG Valine	GCG Alanine	GAG Glutamic Acid	GGG Glycine	G

Codon — Amino acid

Types of Mutation

- Mutations can be classified according to their effects on the protein (or mRNA) produced by the gene that is mutated.

1. Silent mutations (synonymous mutations). Since the genetic code is degenerate, several codons produce the same amino acid. Especially, third base changes often have no effect on the amino acid sequence of the protein. These mutations affect the DNA but not the protein. Therefore, they have no effect on the organism's phenotype.

2. Missense mutations. Missense mutations substitute one amino acid for another. Some missense mutations have very large effects, while others have minimal or no effect. It depends on where the mutation occurs in the protein's structure, and how big a change in the type of amino acid it is.
   - Example: Hb^S, sickle cell hemoglobin, is a change in the beta-globin gene, where a GAG codon is converted to GUG. GAG codes for glutamic acid, which is a hydrophilic amino acid that carries a -1 charge, and GUG codes for valine, a hydrophobic amino acid. This amino acid is on the surface of the globin molecule, exposed to water. Under low oxygen conditions, valine's affinity for hydrophobic environments causes the hemoglobin to crystallize out of solution.

Types of Mutation

3. Nonsense mutations convert an amino acid into a stop codon. The effect is to shorten the resulting protein. Sometimes this has only a little effect, as the ends of proteins are often relatively unimportant to function. However, often nonsense mutations result in completely non-functional proteins.
   - an example: Hb-ß McKees Rock. Normal beta-globin is 146 amino acids long. In this mutation, codon 145 UAU (codes for tyrosine) is mutated to UAA (stop). The final protein is thus 143 amino acids long. The clinical effect is to cause overproduction of red blood cells, resulting in thick blood subject to abnormal clotting and bleeding.

4. Sense mutations are the opposite of nonsense mutations. Here, a stop codon is converted into an amino acid codon. Since DNA outside of protein-coding regions contains an average of 3 stop codons per 64, the translation process usually stops after producing a slightly longer protein.
   - Example: Hb-α Constant Spring. alpha-globin is normally 141 amino acids long. In this mutation, the stop codon UAA is converted to CAA (glutamine). The resulting protein gains 31 additional amino acids before it reaches the next stop codon. This results in thalassemia, a severe form of anemia.

Frameshifts

- Translation occurs codon by codon, examining nucleotides in groups of 3. If a nucleotide or two is added or removed, the groupings of the codons is altered. This is a "frameshift" mutation, where the reading frame of the ribosome is altered.

- Frameshift mutations result in all amino acids downstream from the mutation site being completely different from wild type. These proteins are generally non-functional.

  - example Hb-α Wayne. The final codons of the alpha globin chain are usually AAA UAC CGU UAA, which code for lysine-tyrosine-arginine-stop. In the mutant, one of the A's in the first codon is deleted, resulting in altered codons: AAU ACC GUU AAG, for asparagine-threonine-valine-lysine. There are also 5 more new amino acids added to this, until the next stop codon is reached.

Reversions

A “reversion” is a second mutation that reverse the effects of an initial mutation, bringing the phenotype back to wild type (or almost).

Frameshift mutations sometimes have “second site reversions”, where a second frameshift downstream from the first frameshift reverses the effect.

> Example: consider Hb Wayne above. If another mutation occurred that added a G between the 2 C’s in the second codon, the resulting codons would be: AAU ACG CGU UAA, or asparagine-threonine-arginine-stop. Note that the last 2 codons are back to the original. Two amino acids are still altered, but the main mutational effect has been reverted to wild type.

mRNA Problems

- Although many mutations affect the protein sequence directly, it is possible to affect the protein without altering the codons.

- Splicing mutations. Intron removal requires several specific sequences. Most importantly, introns are expected to start with GT and end in AG. Several beta globin mutations alter one of these bases. The result is that one of the 2 introns is not spliced out of the mRNA. The polypeptide translated from these mRNAs is very different from normal globin, resulting in severe anemia.

- Polyadenylation site mutations. The primary RNA transcript of a gene is cleaved at the poly-A addition site, and 100-200 A's are added to the 3' end of the RNA. If this site is altered, an abnormally long and unstable mRNA results. Several beta globin mutations alter this site: one example is AATAAA -> AACAAA. Moderate anemia was the result.

Trinucleotide Repeats

- A fairly new type of mutation has been described, in which a particular codon is repeated.
- During replication, DNA polymerase can "stutter" when it replicates several tandem copies of a short sequence. For example, CAGCAGCAGCAG, 4 copies of CAG, will occasionally be converted to 3 copies or 5 copies by DNA polymerase stuttering.
- Outside of genes, this effect produces useful genetic markers called SSR (simple sequence repeats).
- Within a gene, this effect can cause certain amino acids to be repeated many times within the protein. In some cases this causes disease.
- The Huntington's disease gene normally has between 11 and 33 copies of CAG (codon for glutamine) in a row. The number occasionally changes. People with HD have 37 or more copies, up to 200.
- Interestingly, the age of onset of the disease is related to the number of CAG repeats present: the more repeats, the earlier the onset.

Germinal vs. Somatic Mutations

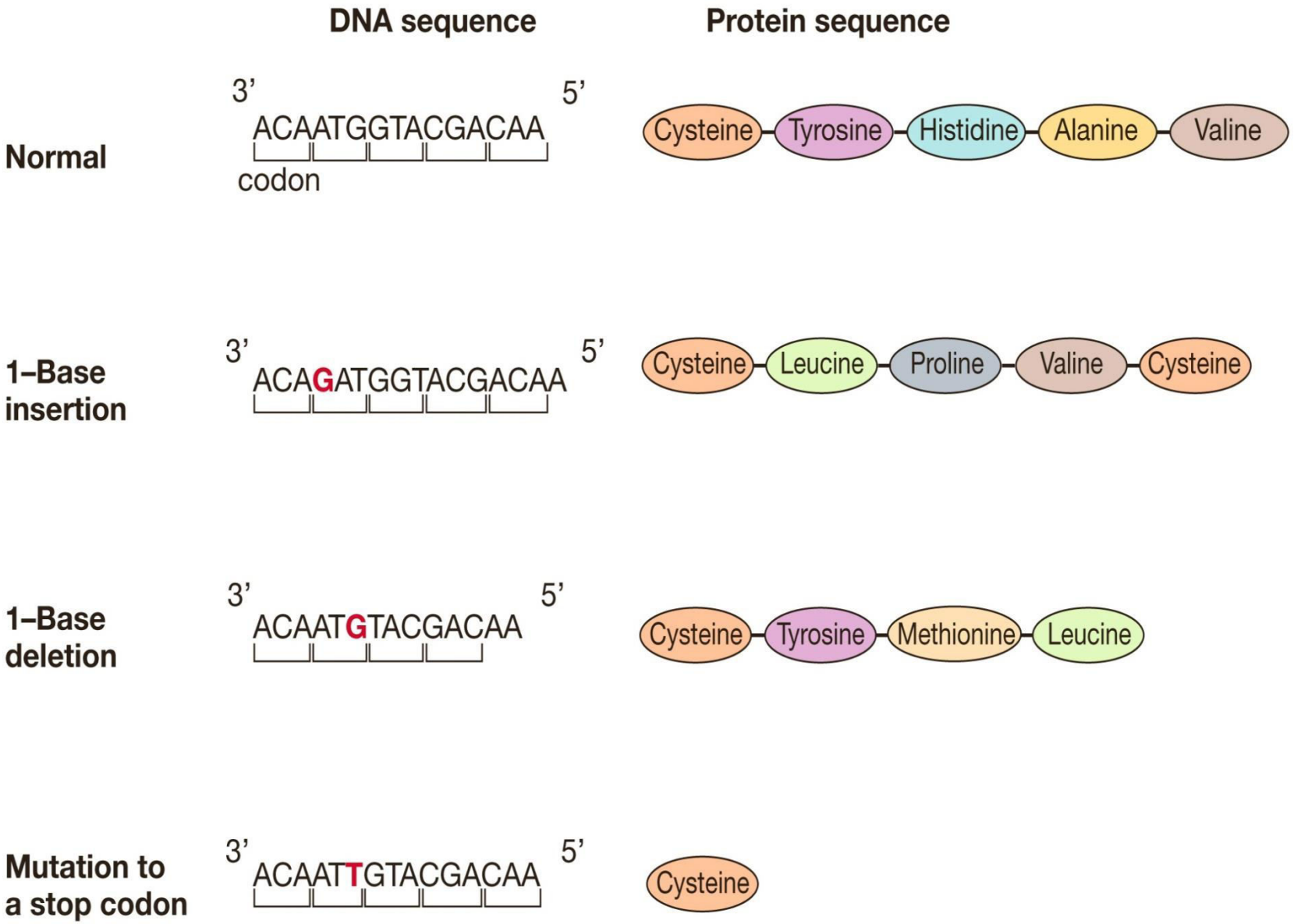
- Mutations can occur in any cell. They only affect future generations if they occur in the cells that produce the gametes: these are "germinal" or "germ line" mutations.

- Mutations in other cells are rarely noticed, except in the case of cancer, where the mutated cell proliferates uncontrollably. Mutations in cells other than germ line cells are "somatic" mutations.

- A human body contains 10^{13} - 10^{14} cells approximately. The average mutation rate for any given nucleotide is about 1 in 10^9. That is, on the average 1 cell in 10^9 has that particular nucleotide altered. This means that virtually every possible base change mutation occurs repeatedly in our body cells.

Mutation rates

- Most data on mutations comes from analysis of loss-of-function mutations.
- Loss-of-function mutations cause gene to produce a non-working protein.
- Examples of loss-of-function mutations include: insertions and deletions, mutation to a stop codon and insertion of jumping genes.



Mutation rates

- Some mutations cause readily identified phenotypic changes.

- E.g. Achrondoplastic dwarfism is a dominant disorder. An Achrondoplastic individual's condition must be the result of a mutation, if his parents do not have the condition.

Mutation rates

- Human estimate is 1.6 mutations/genome/generation.

- In *Drosophila* rate is only 0.14 m/g/g, but when corrected for number of cell divisions needed to produce sperm (400 in humans 25 in *Drosophila)* mutation rates per cell division are very similar.