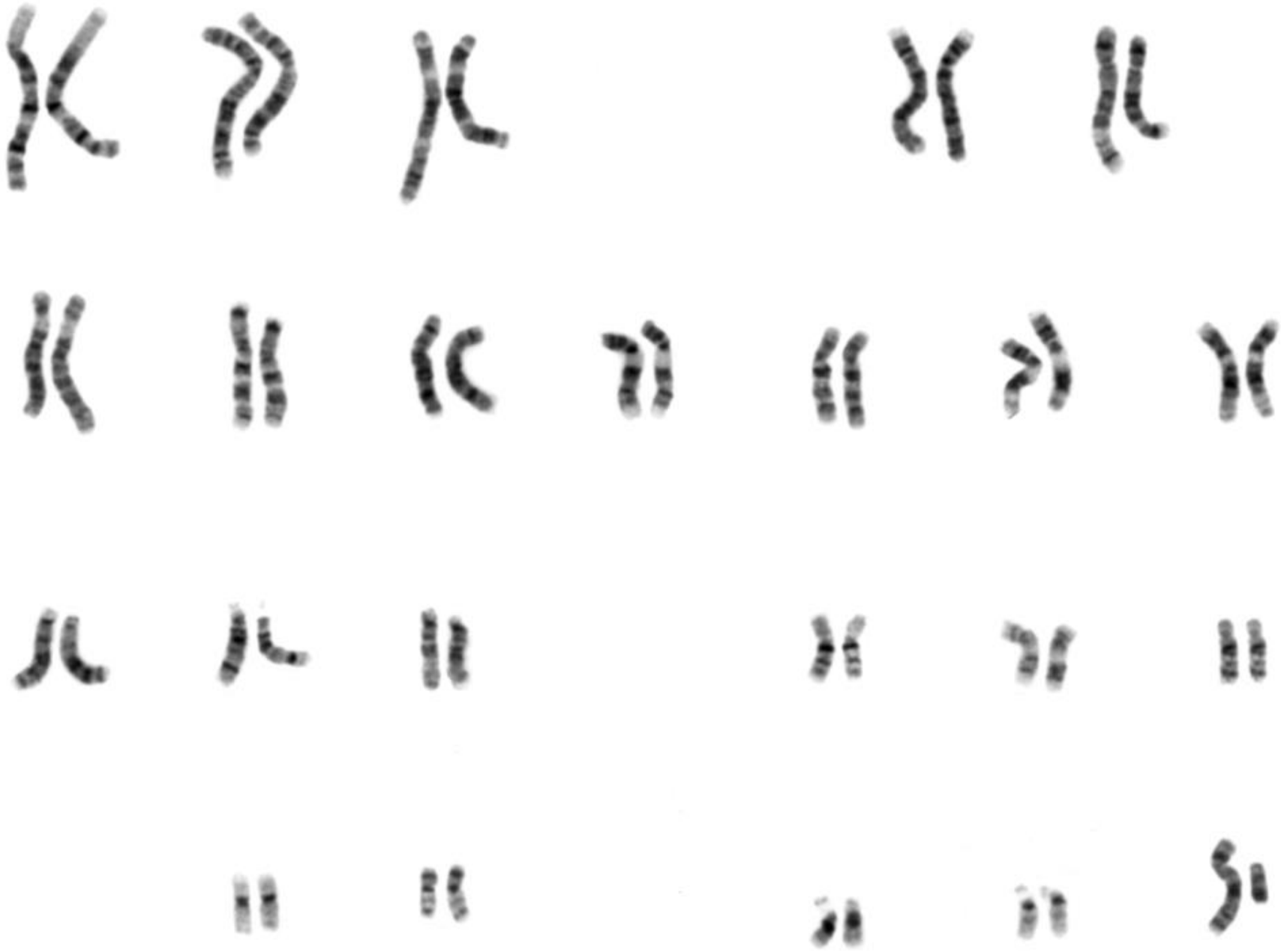
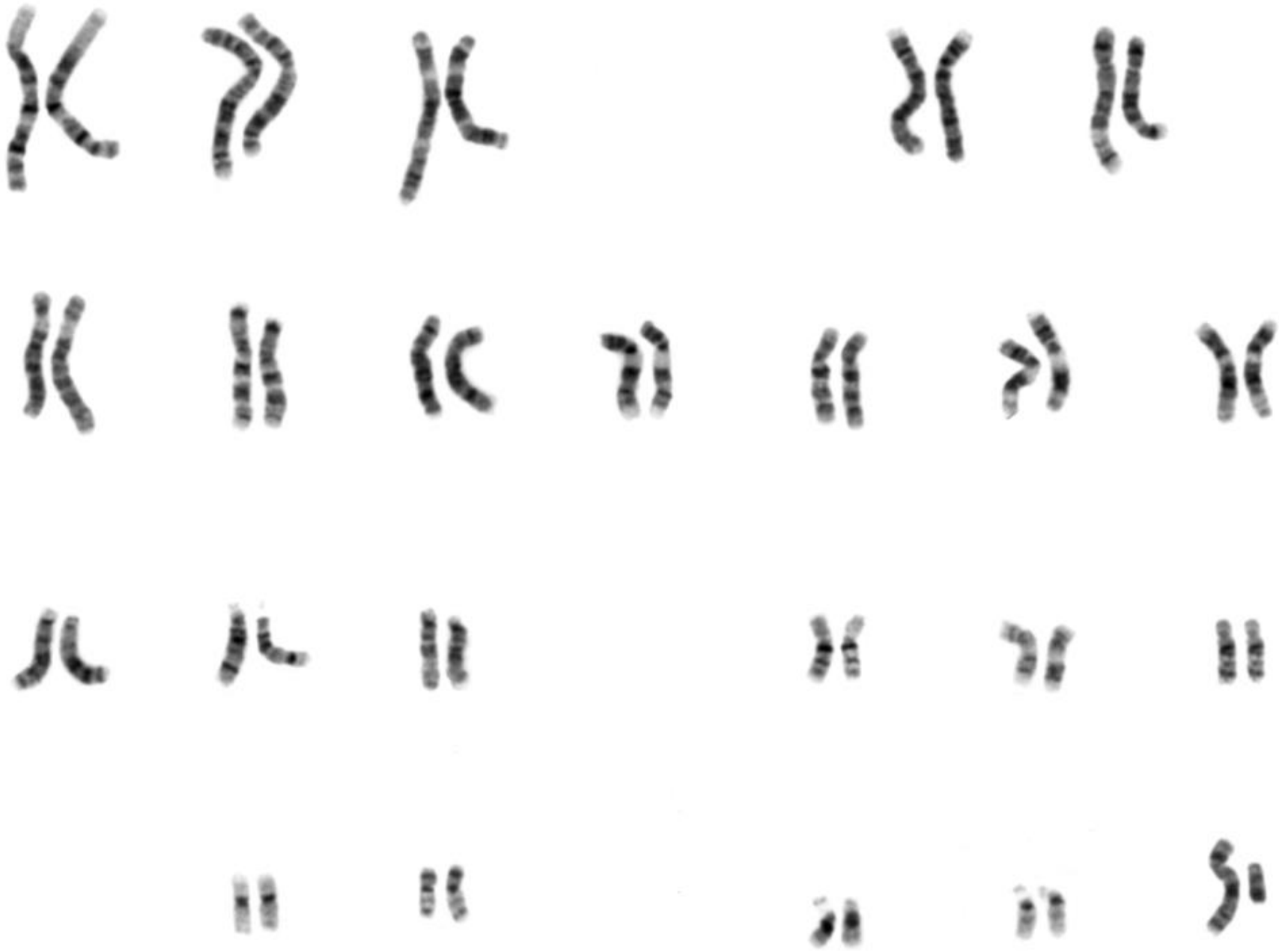
Human Genetic Variation

Section 3

Learning objectives

- Describe differences in types of genetic variation and how they affect phenotypes.
- Identify inheritance patterns of genotype-phenotype relationships.
- Describe the differences and the pros and cons of sequencing vs genotyping.





Zoom Poll

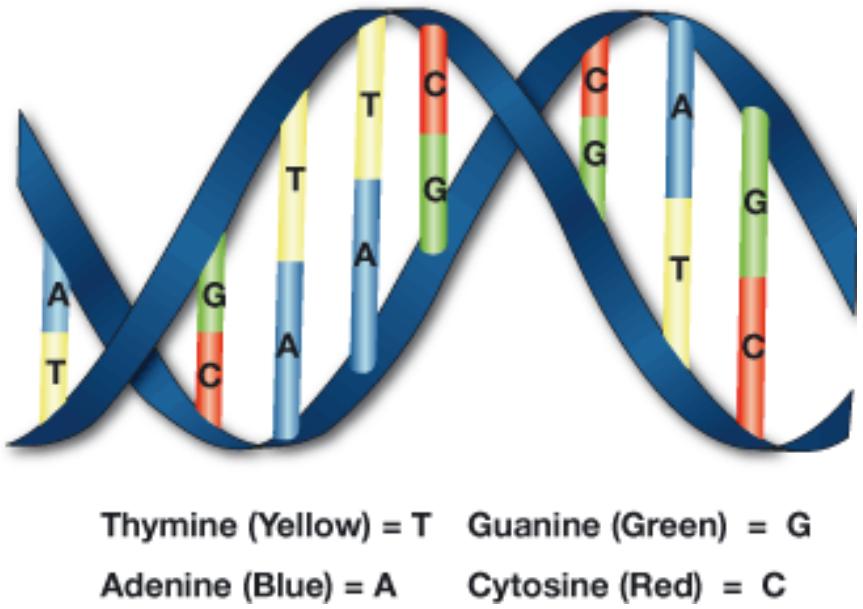
Our Genome in Numbers

23 chromosome pairs

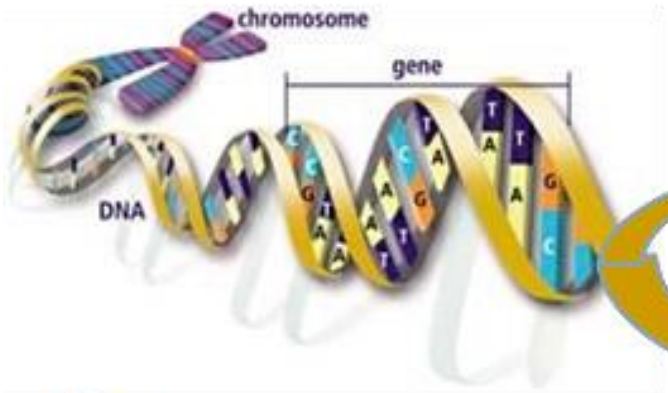
3.2 billion base-pairs (A,C,G,T)

~20,000 genes

~1.5% of the genome is coding DNA



Transcription and Translation



Coding Strand

ATGAAGCCGAGTCAT

UACUUCGGCUCAGUA ← mRNA

TACTTCGGCTCAGTA

Noncoding Strand

Nuclear Membrane

UACUUCGGCUCAGUA ← mRNA

Read as 3-letter "words" or triplets called "Codons"

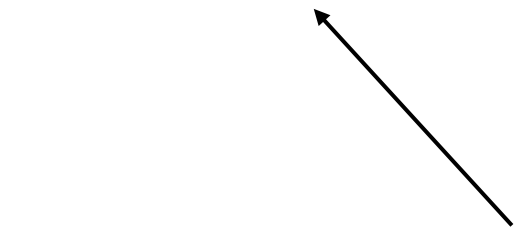
AUG | AAG | CCG | AGU | CAU

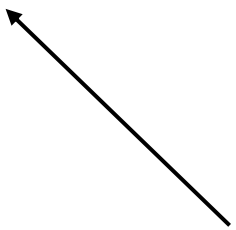
Met Lys Pro Ser Valine

"Anticodons" on tRNAs

Amino Acids

Genetic variation to phenotype variation



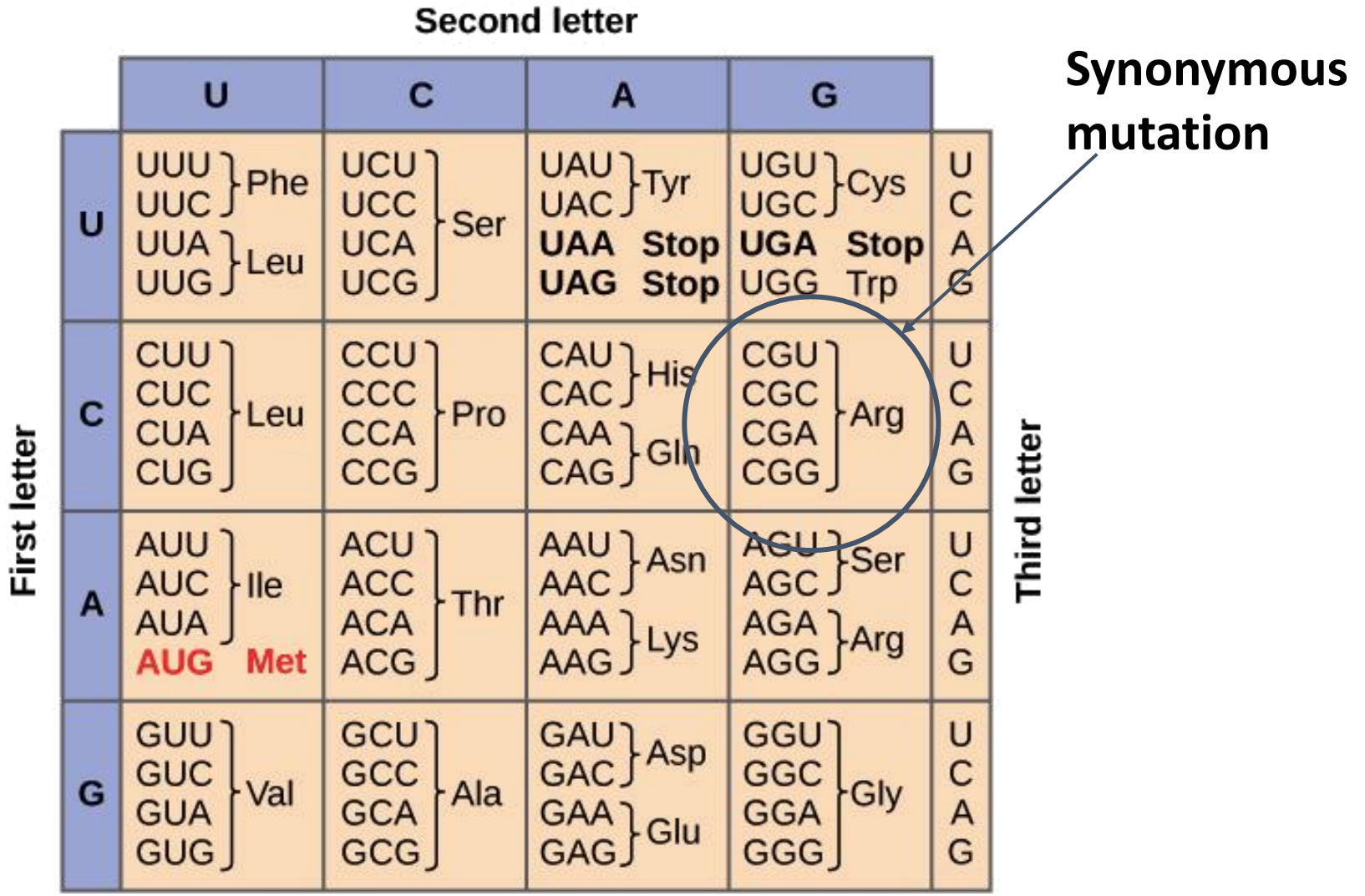


The changes in DNA

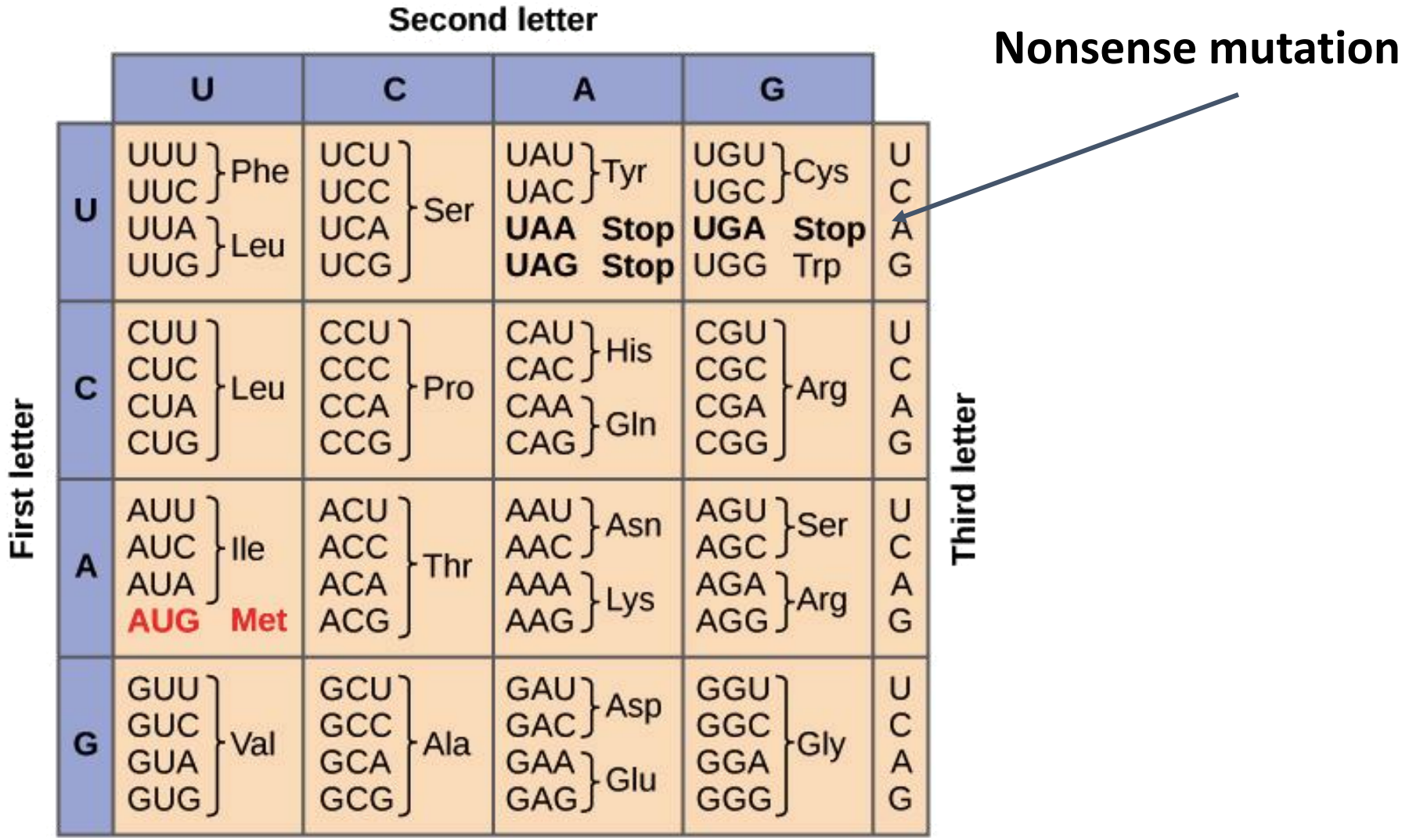
What we actually see
(disease, trait)

Single base change = Single Nucleotide Polymorphism/Variant (SNP/SNV)

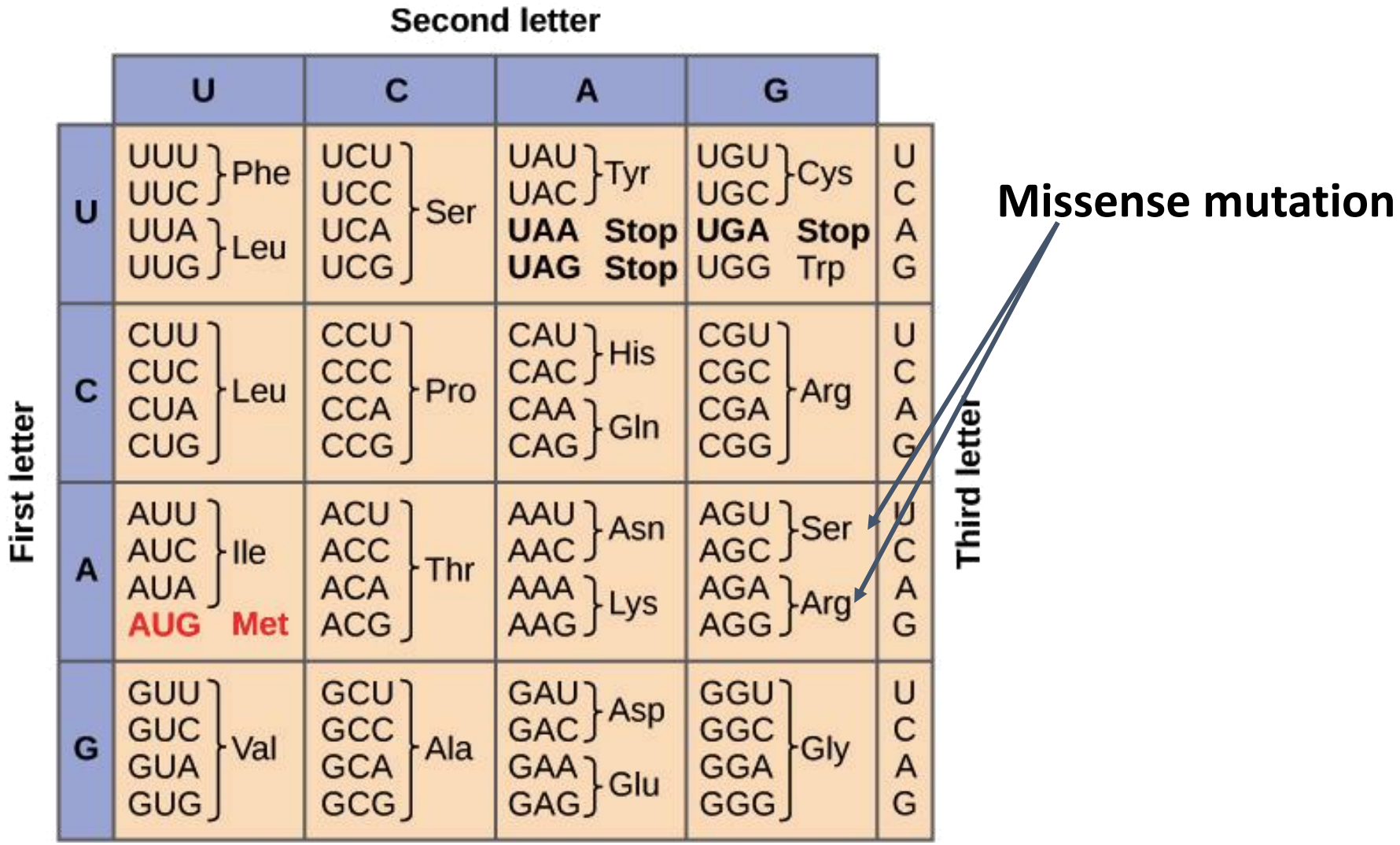
Genetic variant – changes in amino acid codons



Nonsynonymous is usually worse than synonymous

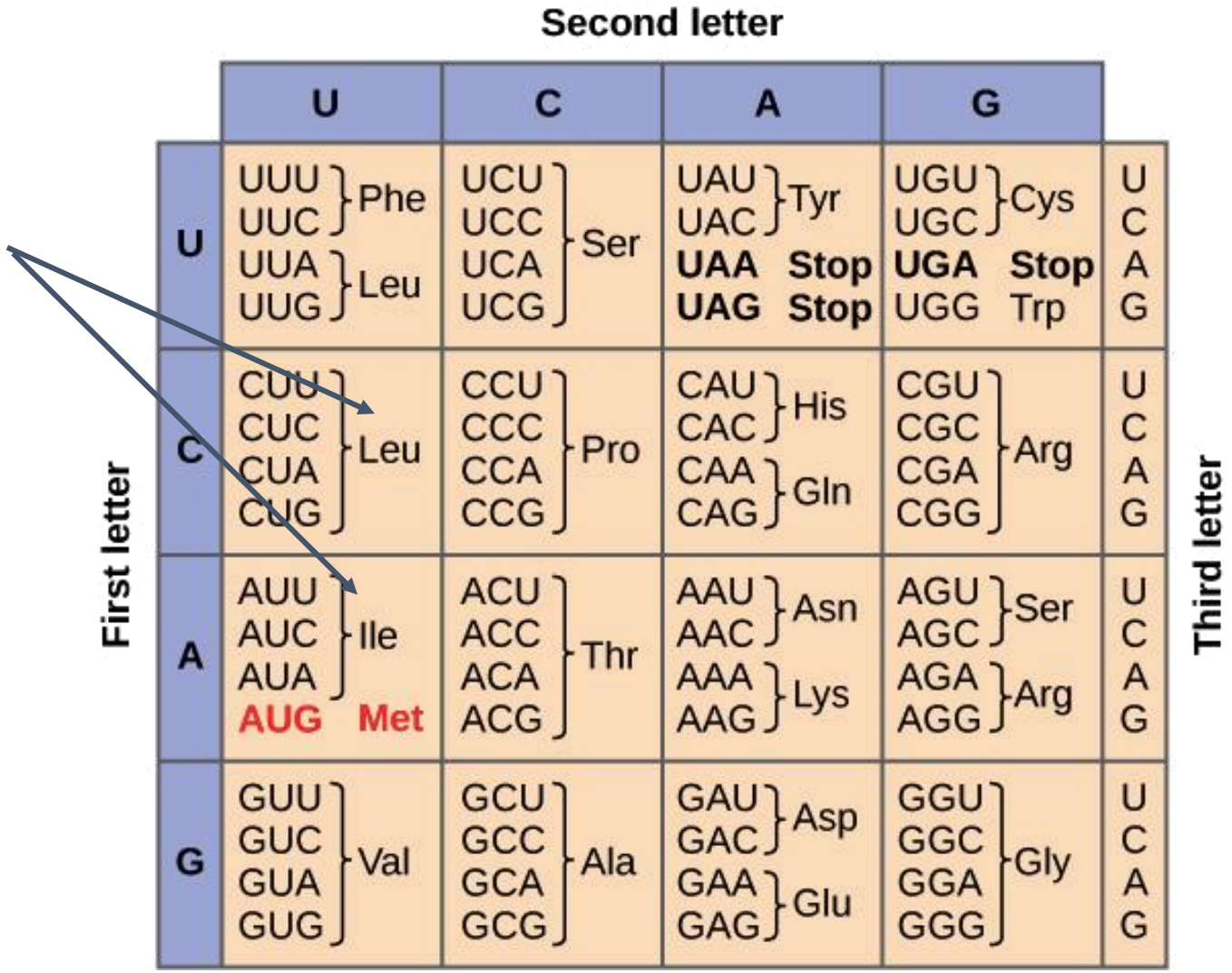


Nonsynonymous mutations



Some missense mutations can be less bad

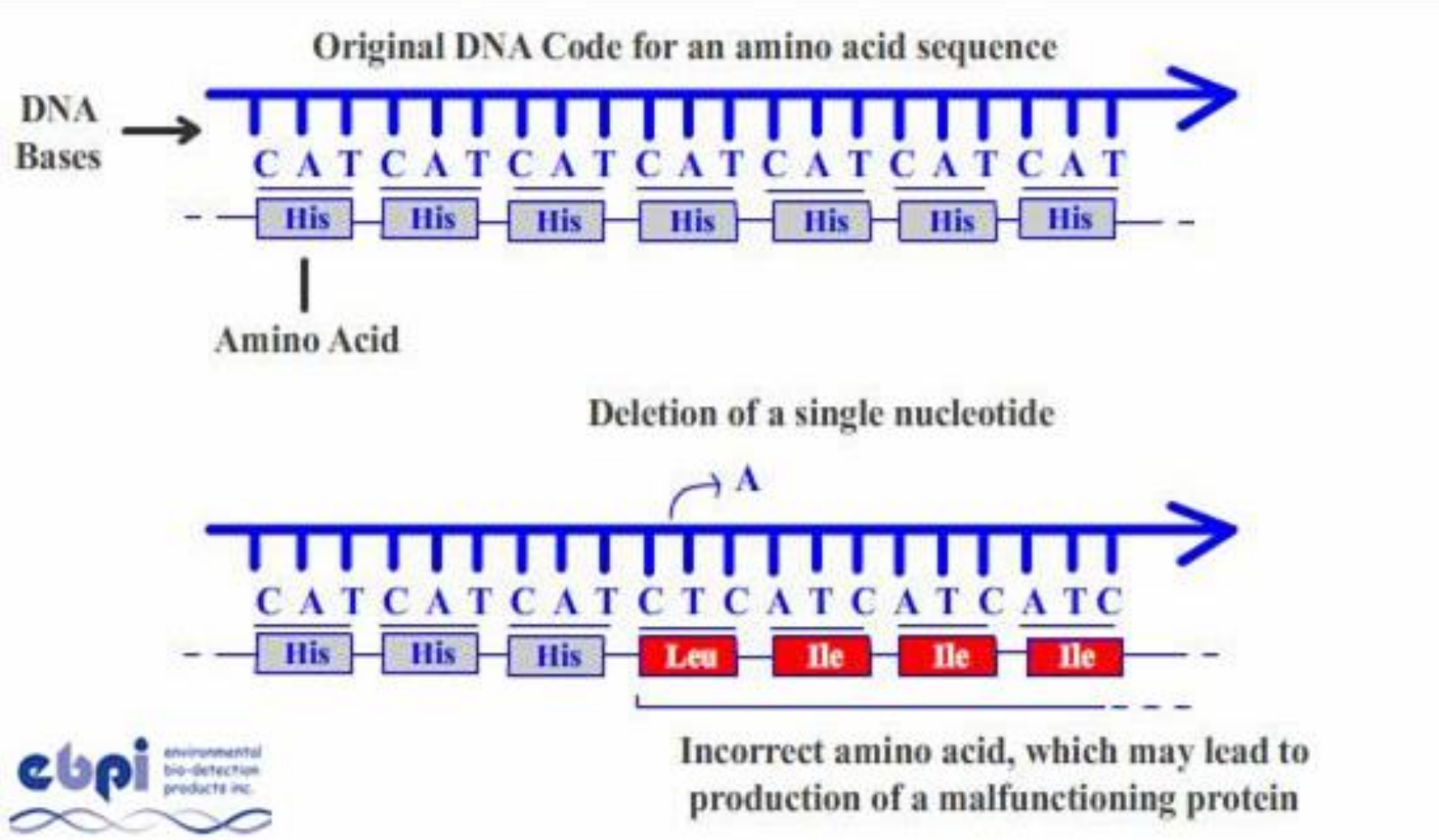
Both polar



First letter \ Second letter	U	C	A	G	Third letter
U	UUU, UUC: Phe UUA, UUG: Leu	UCU, UCC, UCA, UCG: Ser	UAU, UAC: Tyr **UAA Stop** **UAG Stop**	UGU, UGC: Cys **UGA Stop** UGG Trp	U C A G
C	CUU, CUC, CUA, CUG: Leu	CCU, CCC, CCA, CCG: Pro	CAU, CAC: His CAA, CAG: Gln	CGU, CGC, CGA, CGG: Arg	U C A G
A	AUU, AUC, AUA: Ile **AUG Met**	ACU, ACC, ACA, ACG: Thr	AAU, AAC: Asn AAA, AAG: Lys	AGU, AGC: Ser AGA, AGG: Arg	U C A G
G	GUU, GUC, GUA, GUG: Val	GCU, GCC, GCA, GCG: Ala	GAU, GAC: Asp GAA, GAG: Glu	GGU, GGC, GGA, GGG: Gly	U C A G

Deletions/insertions

“Frameshifts”



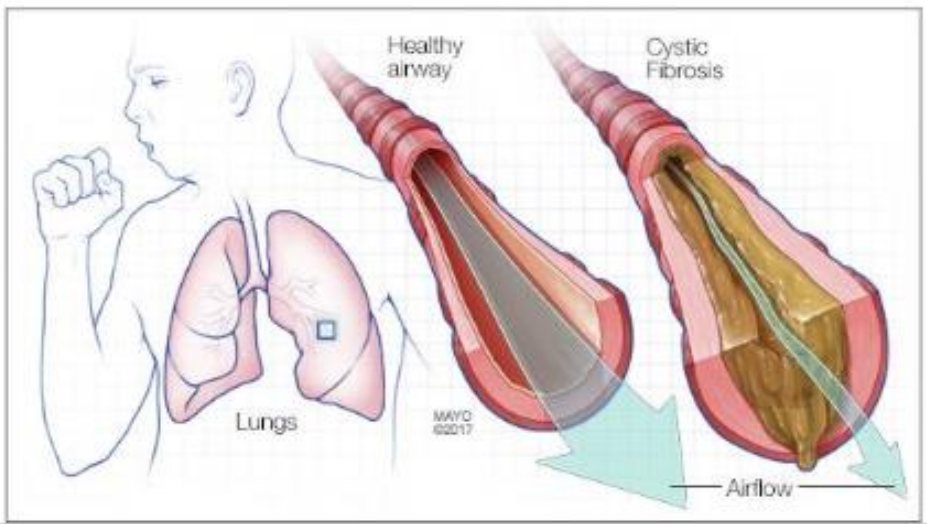
Deletion – cystic fibrosis F508del

Functioning CFTR sequence:

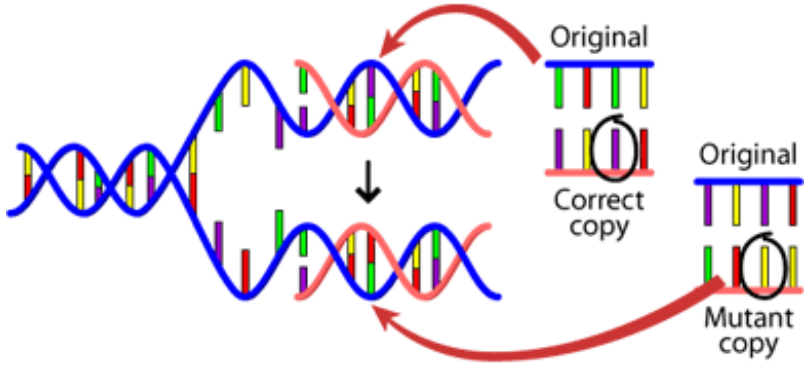
Nucleotide	ATC	ATC	TTT	GGT	GTT
Amino acid	Ile	Ile	Phe	Gly	Val

F508Del variant inactivating chloride channel:

Nucleotide	ATC	ATT	GGT	GTT
Amino acid	Ile	Ile	Gly	Val



Mutations happen all the time, with every replication



Human genome mutation rate is ~1.1×10^{-8} per site per generation.

Human genome is over 3 billion base pairs.

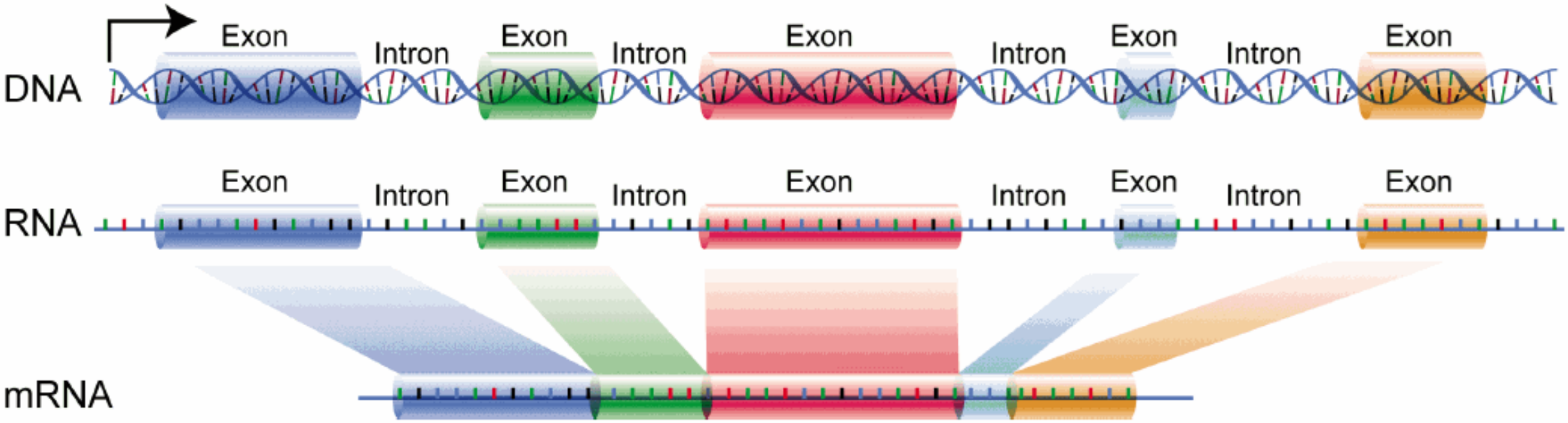
Each genome: 3,000,000,000 sites
Mutation rate: 0.000000011 errors/site

How many new mutations do you expect in expect in each cell replication?

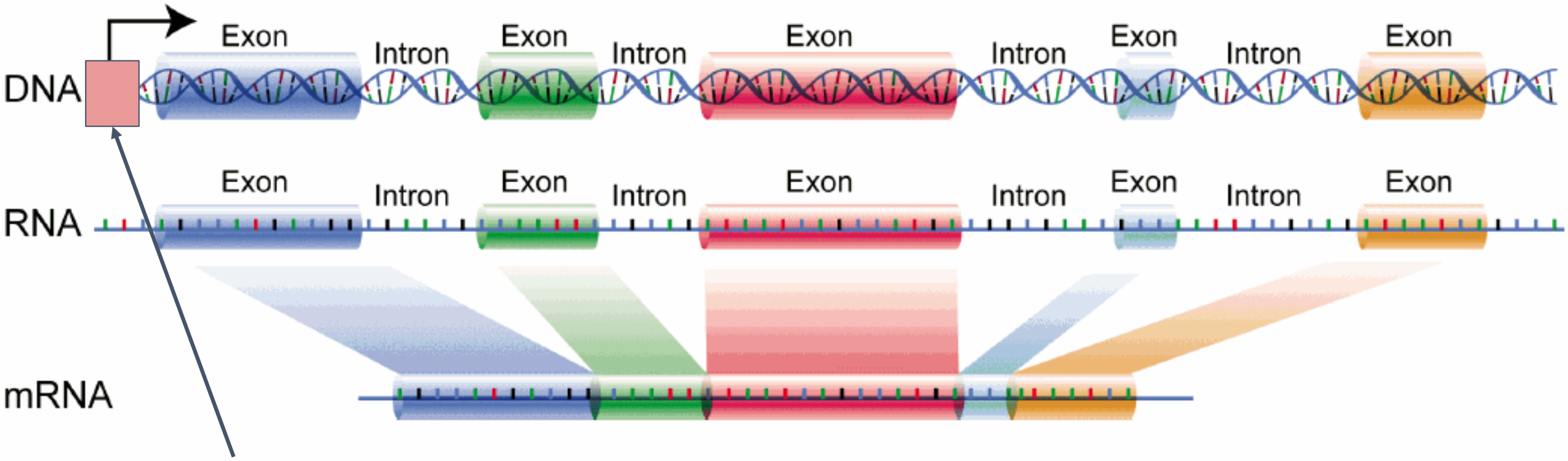
Zoom Poll

Sometimes we don’t change the protein itself…

A gene includes a lot of DNA that doesn't become protein

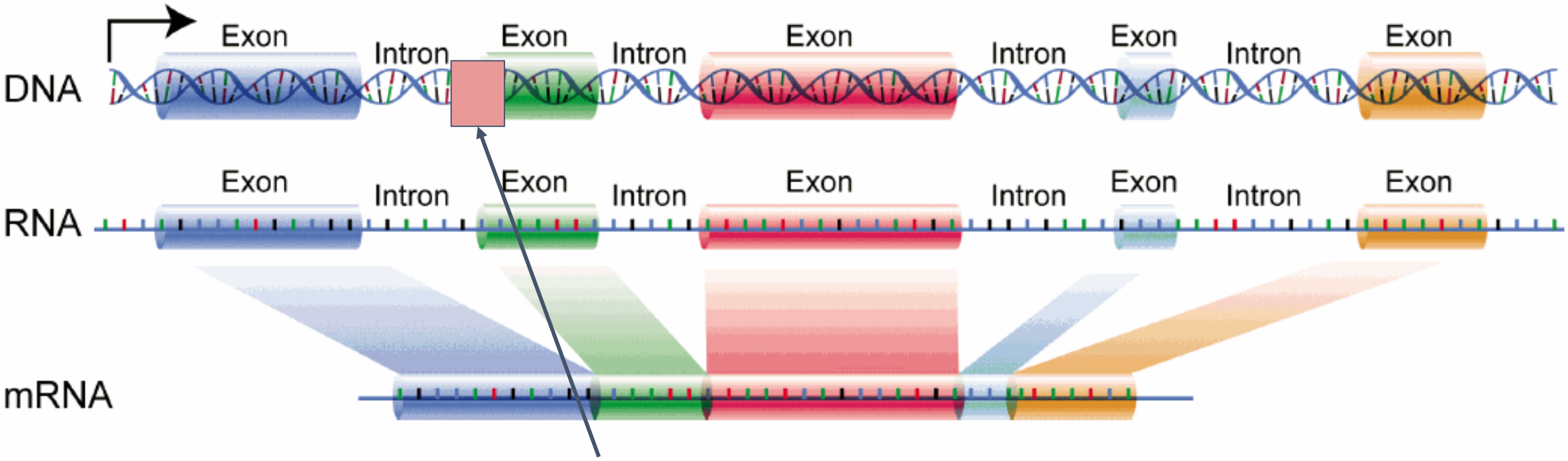


A gene includes a lot of DNA that doesn't become protein



A variant here can change gene "expression"

A gene includes a lot of DNA that doesn't become protein

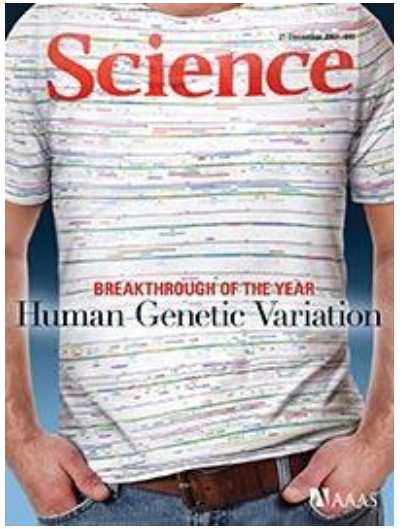


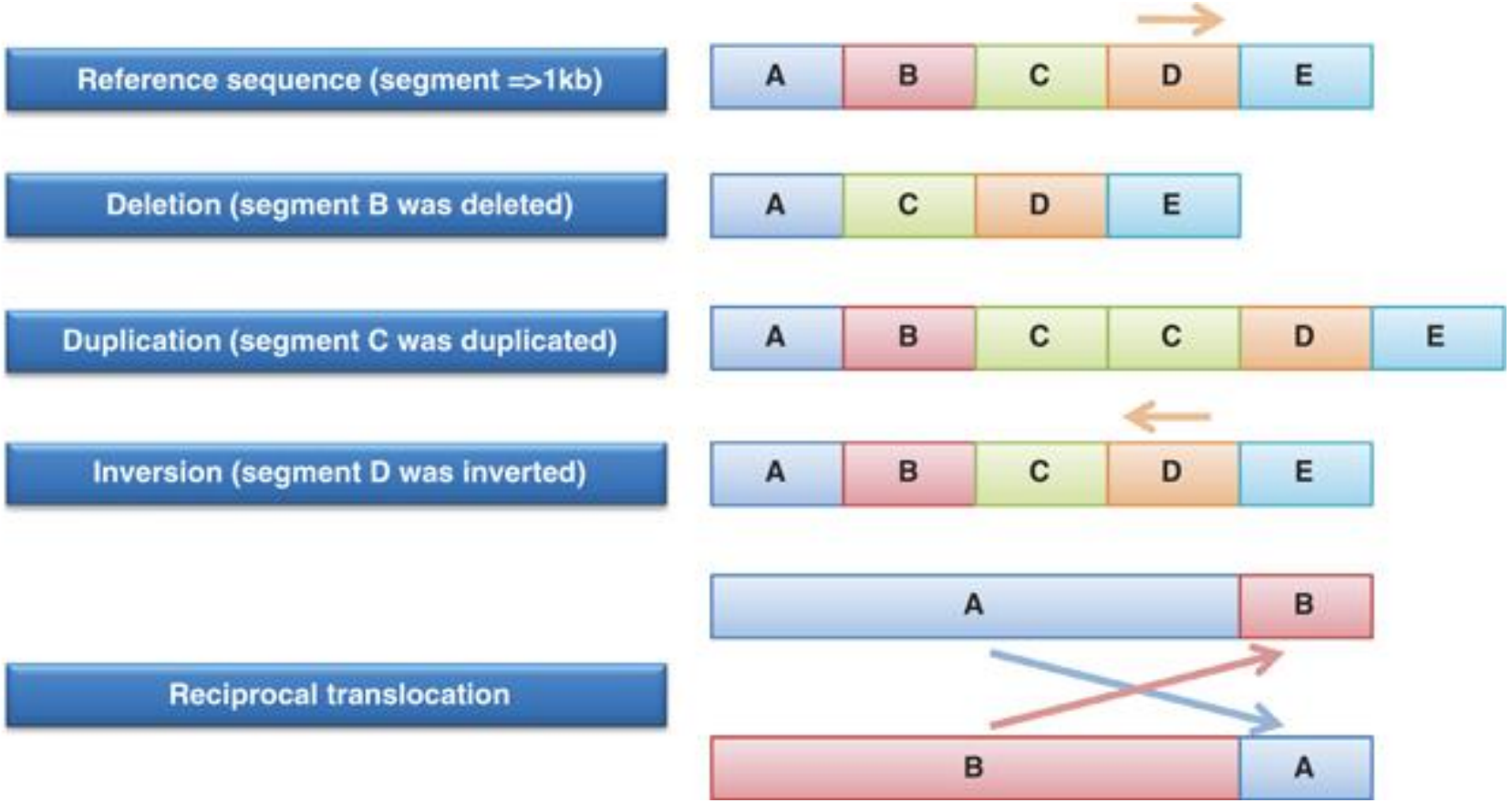
Or here can change the "splice site" to make a different protein

Zoom breakout – discuss Q1

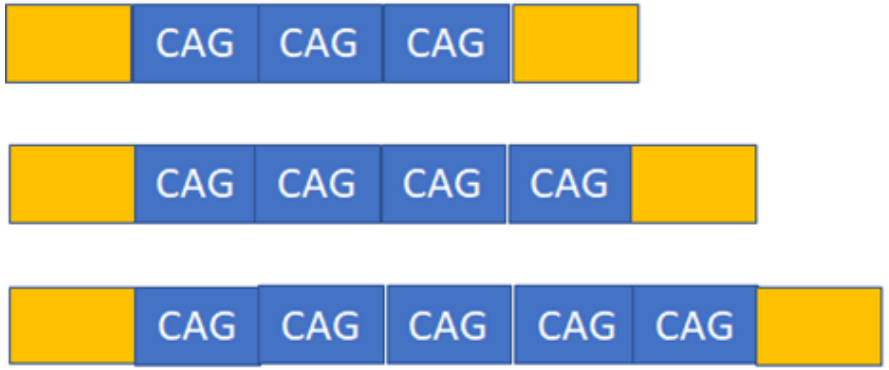
Besides single base changes, what types of changes can we have?

Genetic Variation – structural variation





tandem repeats (Huntington's disease)



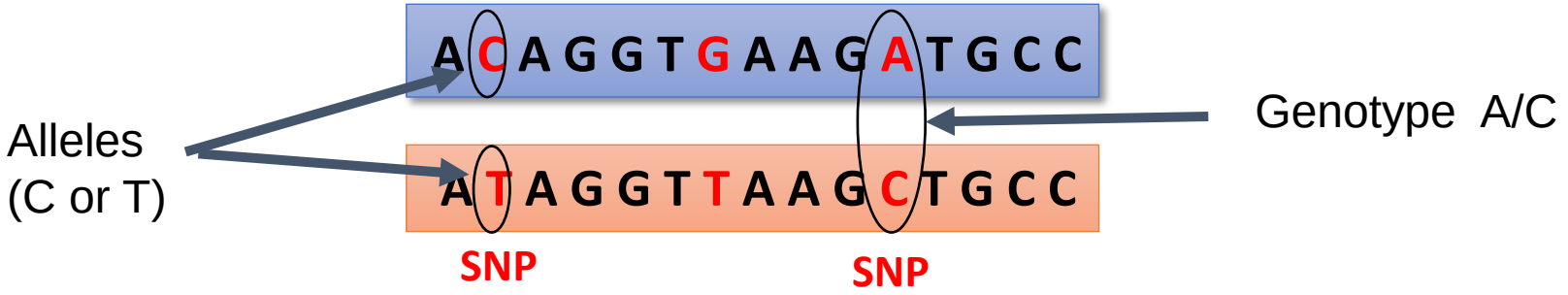
The normal function of huntingtin is unknown. The CAG repeats (polyglutamine) is in some way neurotoxic.

Repeat count	Classification	Disease status
<28	Normal	Unaffected
28–35	Intermediate	Unaffected
36–40	Reduced-penetrance	May be affected
>40	Full-penetrance	Affected

Alleles to genotypes and phenotypes

Allele vs. genotype

We inherit two copies of each chromosome



Genotypes
(A/A) – homozygous
(A/C) – heterozygous
(C/C) - homozygous

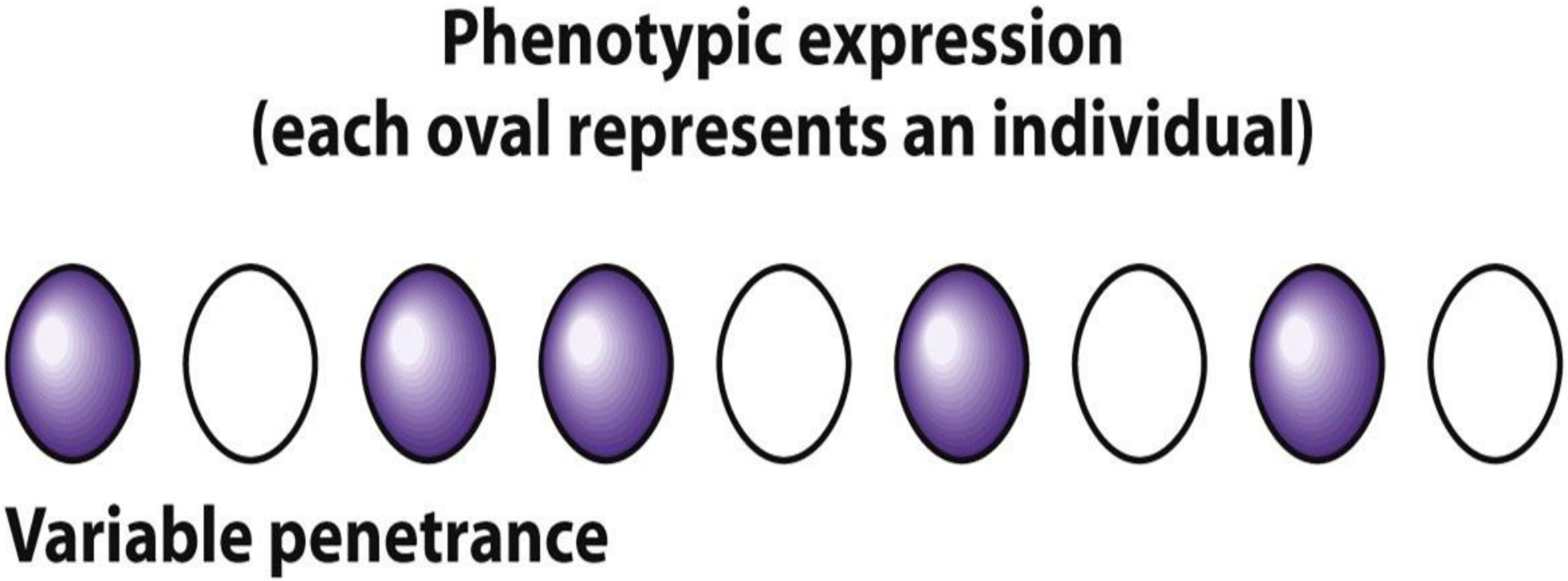
Inheritance patterns: genotype->phenotype

- 2 copies of every gene/chromosome (normally)
- Dominant (only need one copy of a variant to see the effect)
- Recessive (need two copies of the variant to see the effect)
- Additive (the effect of one variant is ½ that of two variants)

Genotypes and Phenotypes

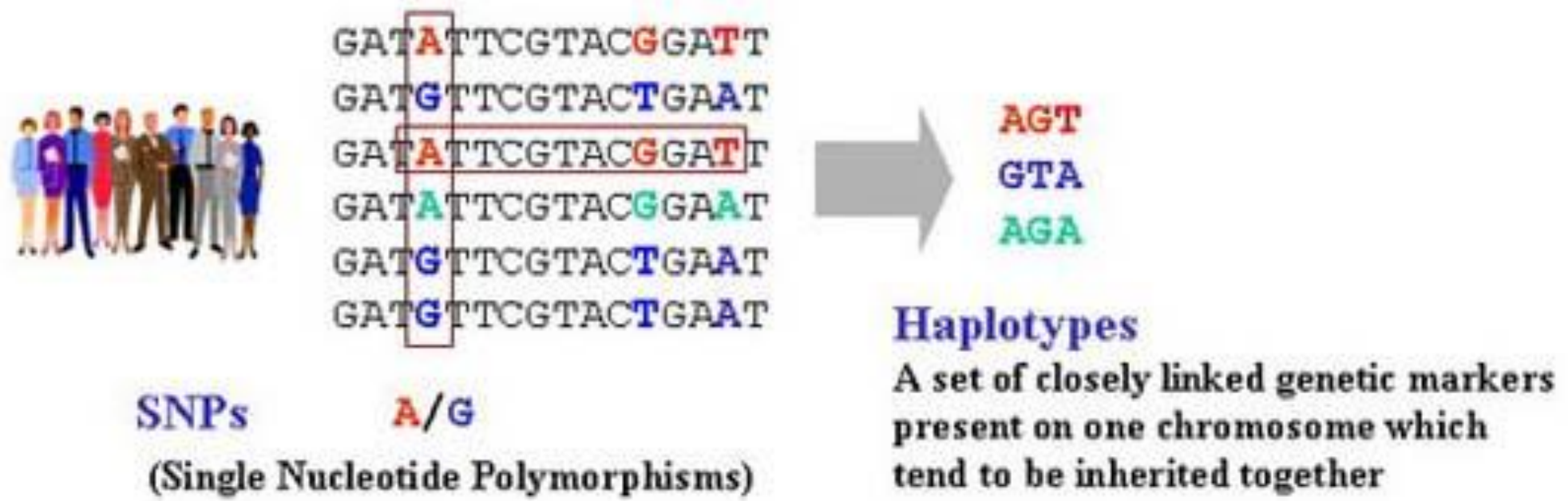
- **Mendelian phenotype** is one driven by variation at a single genetic locus.
- **Complex phenotype** does not show such simple patterns of inheritance.
  - oligogenic (a few genetic loci)
  - polygenic (many genetic loci)
- Binary outcomes (yes/no, i.e. disease status)
- Quantitative outcomes (continuous, i.e. height)

Same genetic pattern, different phenotype

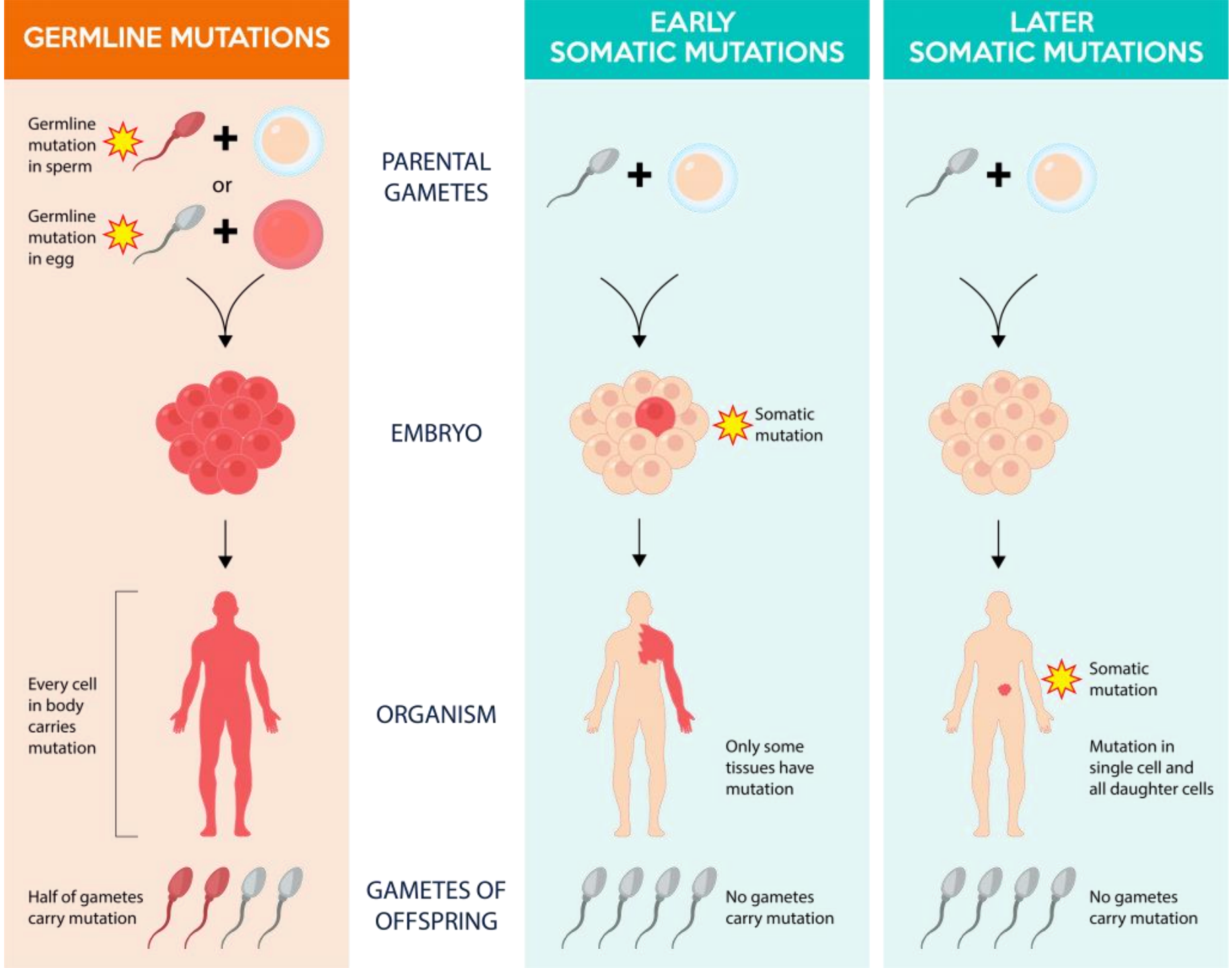


Haplotypes

Specific combination of SNPs occurring on the same segment of chromosome.



Zoom breakout Q2


GERMLINE MUTATIONS
EARLY SOMATIC MUTATIONS
LATER SOMATIC MUTATIONS
Germline mutation in sperm
or
Germline mutation in egg
PARENTAL GAMETES
EMBRYO
Somatic mutation
Every cell in body carries mutation
ORGANISM
Only some tissues have mutation
Somatic mutation
Mutation in single cell and all daughter cells
Half of gametes carry mutation
GAMETES OF OFFSPRING
No gametes carry mutation
No gametes carry mutation

Genetic data collection

- TaqMan Polymerase chain reaction (PCR)
  - Targeted, low throughput.
  - Detect deletions and structural variations.
- Genotyping chip
  - Targeted locations, high throughput.
  - Detects single, *a priori* locations.
- Sequencing
  - Collects all bases, increasingly high throughput.
  - Identify novel variants.
  - Analyzing data more intensive

Reference SNP (rs) Report

Download

Switch to classic site

rs776746

Current Build 152
Released October 2, 2018

Organism	*Homo sapiens*
Position	chr7:99672916 (GRCh38.p12)
Alleles	T>C
Variation Type	SNV Single Nucleotide Variation
Frequency	T=0.28922 (36317/125568, TOPMED) T=0.2653 (8204/30920, GnomAD) T=0.379 (1896/5008, 1000G) (+ 3 more)
Clinical Significance	Reported in ClinVar
Gene : Consequence	CYP3A5 : Splice Acceptor Variant ZSCAN25 : Intron Variant
Publications	386 citations
Genomic View	See rs on genome

FEEDBACK

Variant Details

Clinical Significance

Frequency

Genomic Placements

Sequence name	Change
GRCh37.p13 chr 7	NC_000007.13:g.99270539C>T
GRCh38.p12 chr 7	NC_000007.14:g.99672916T>C

Zoom breakout exercises #3

Summary

- Genetic variation can affect single nucleotides or longer segments through structural changes.
- Changes in DNA affect what we see (phenotypes) depending on where they are in the genome and their role in protein production.