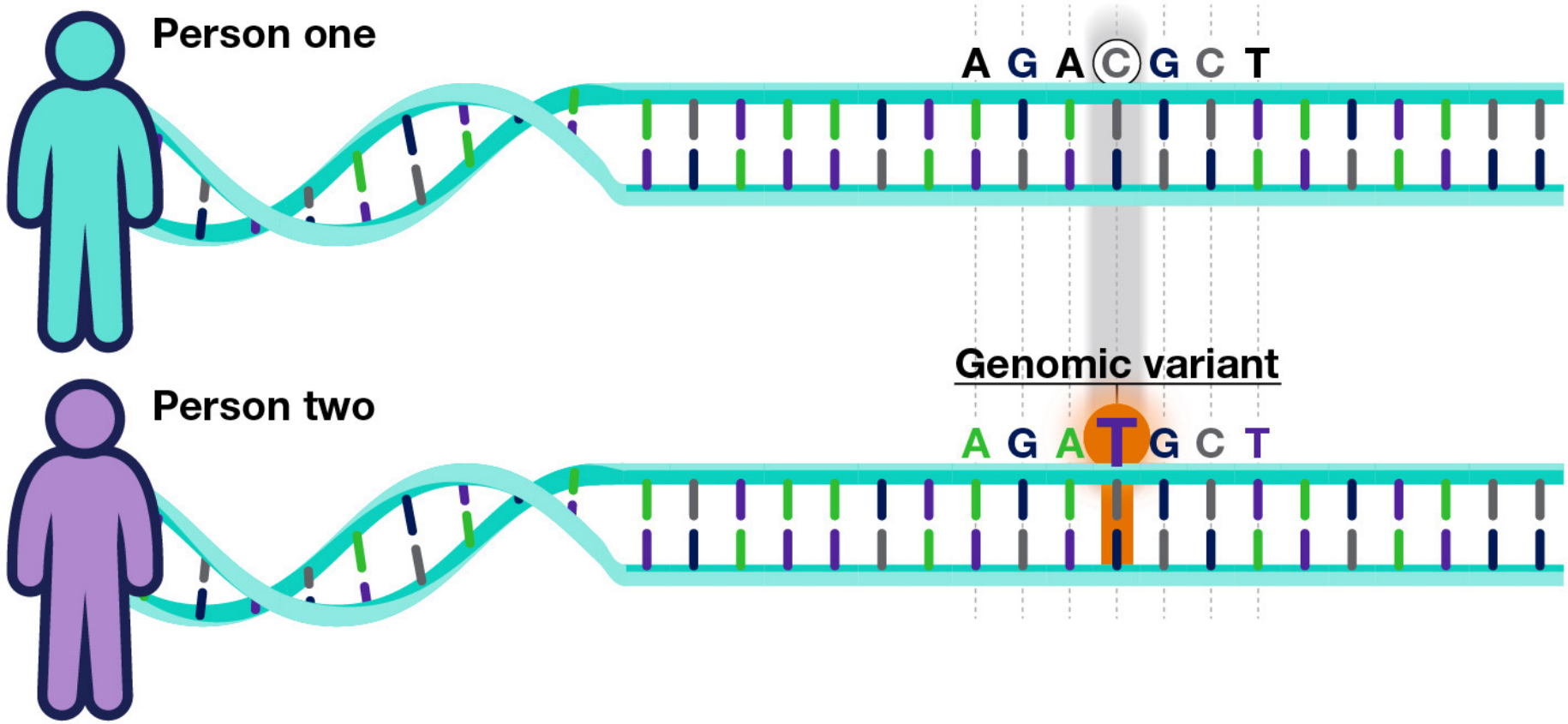
Genetic variation

- Types and classifications.
- Abundance.
- Effect.


Person one
A G A C G C T
Genomic variant
Person two
A G A T G C T

1. HOW MANY VARIANTS DOES A HUMAN GENOME HAVE?

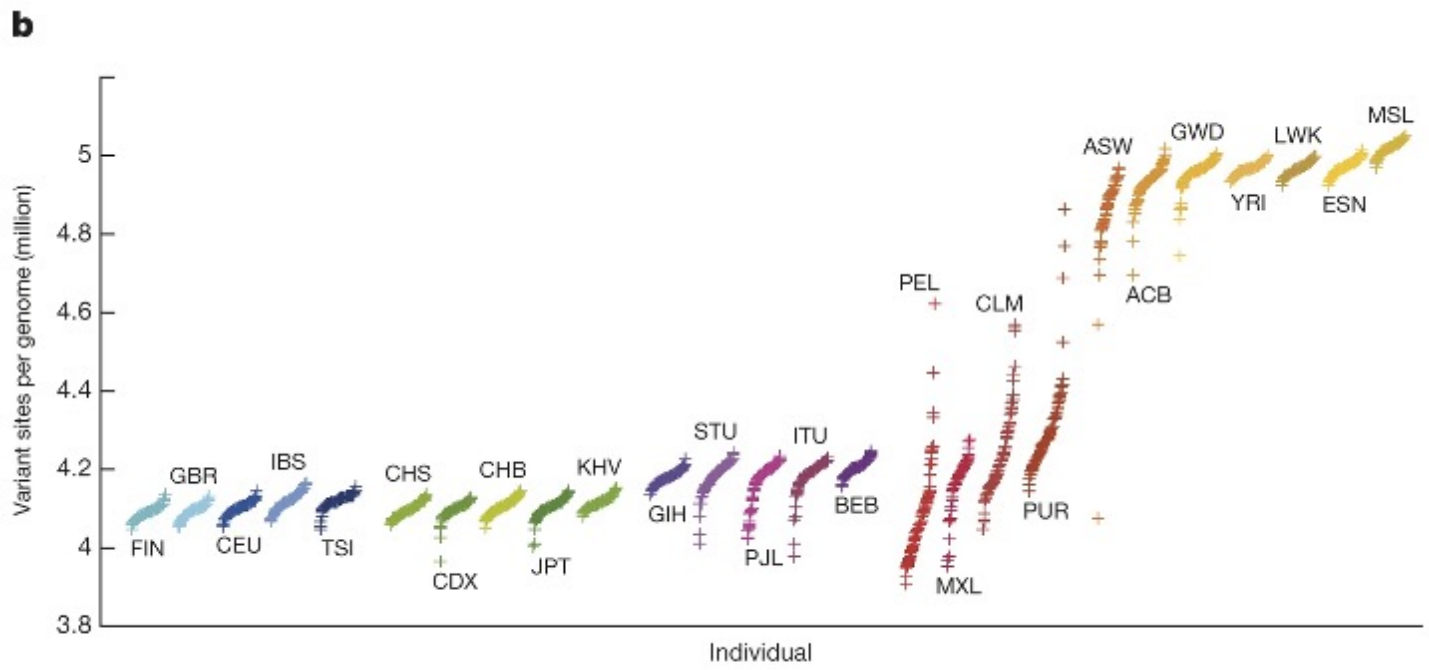
2. WHICH TYPE?

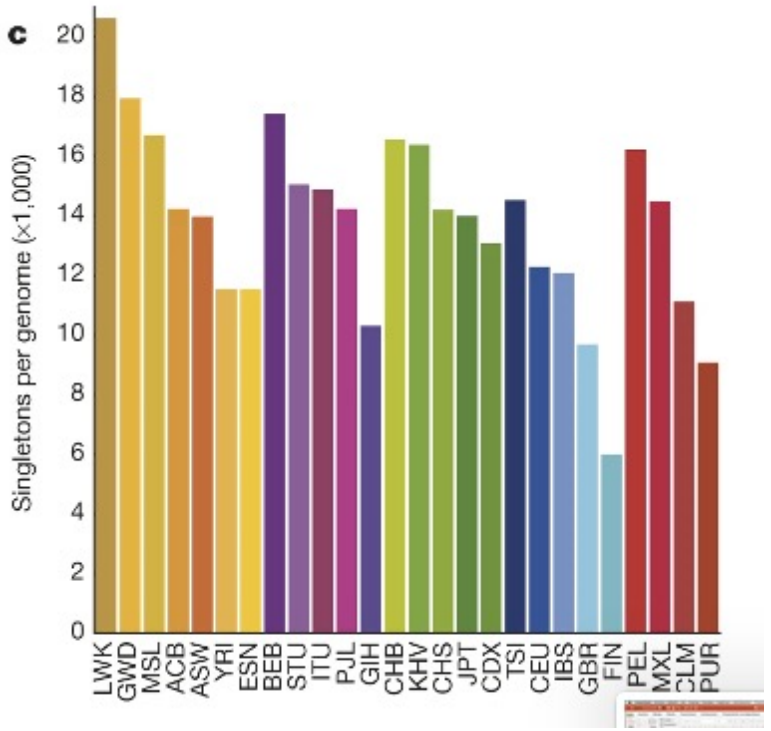
3. WHAT EFFECT DO THEY HAVE?

4. WHAT IS THE BIOLOGICAL SIGNIFICANCE?

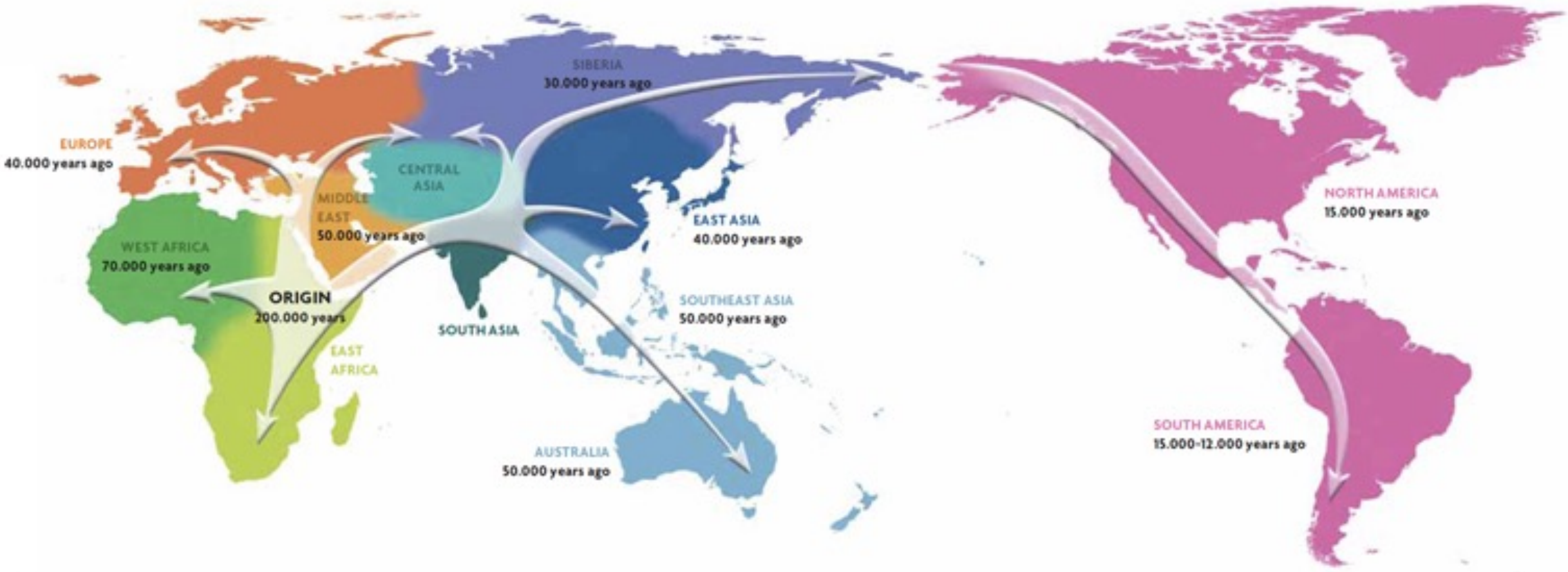
1. HOW MANY VARIANTS DOES OUR GENOME HAVE?

1. HOW MANY VARIANTS DOES OUR GENOME HAVE?





1. HOW MANY VARIANTS DOES THE GENOME HAVE?

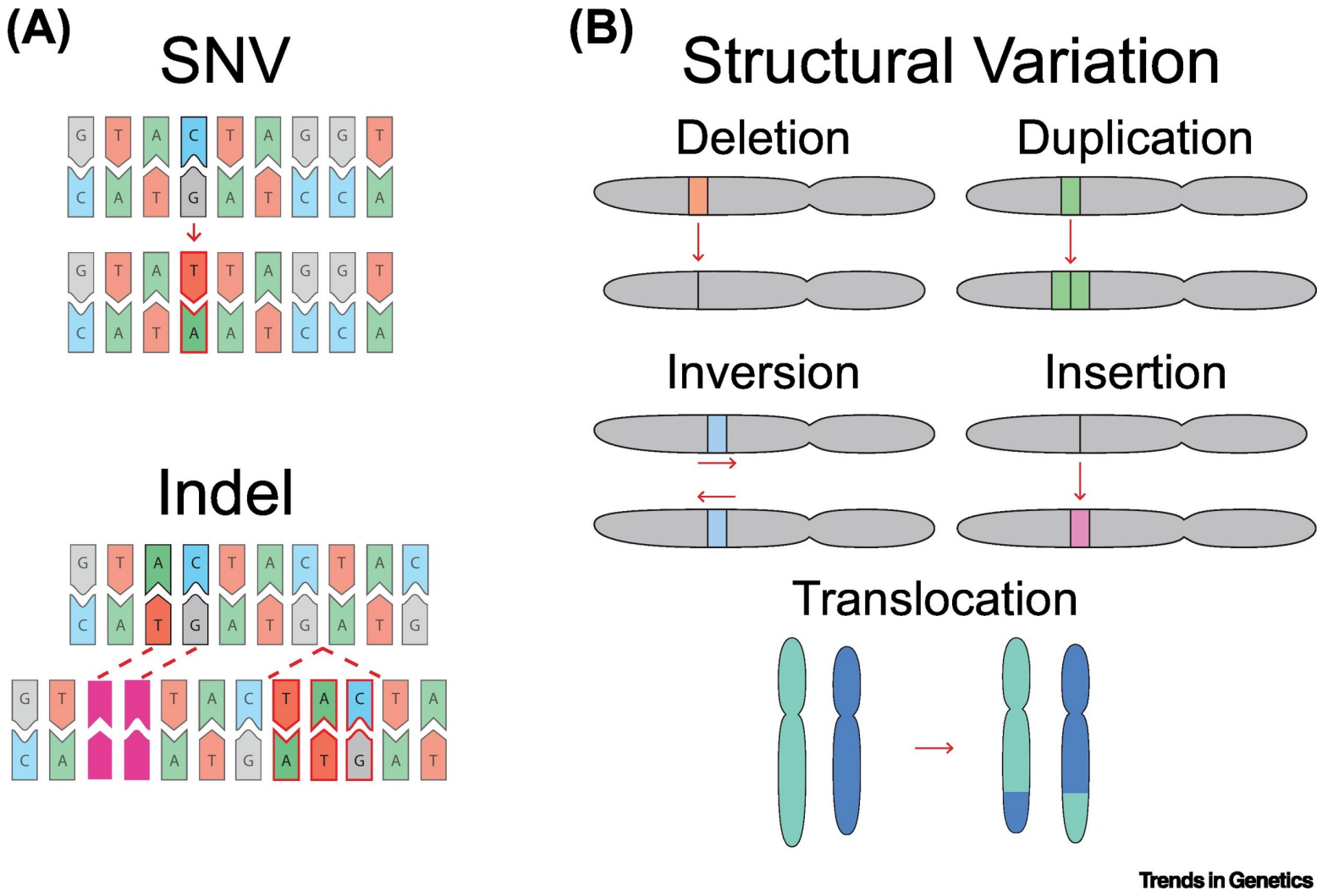


2. TYPES OF GENETIC VARIANTS

- Size (SNV, indels, CNVs)
- Frequency (rare, common)
- Nature (somatic, germinal)
- Localization (intergenic, UTRs, exons)

2. TYPES OF GENETIC VARIANTS

classification by SIZE



De *Hotspots of Human Mutation*, Nesta et al. 2020

2. TYPES OF GENETIC VARIANTS

classification by SIZE

Chromosome mutations

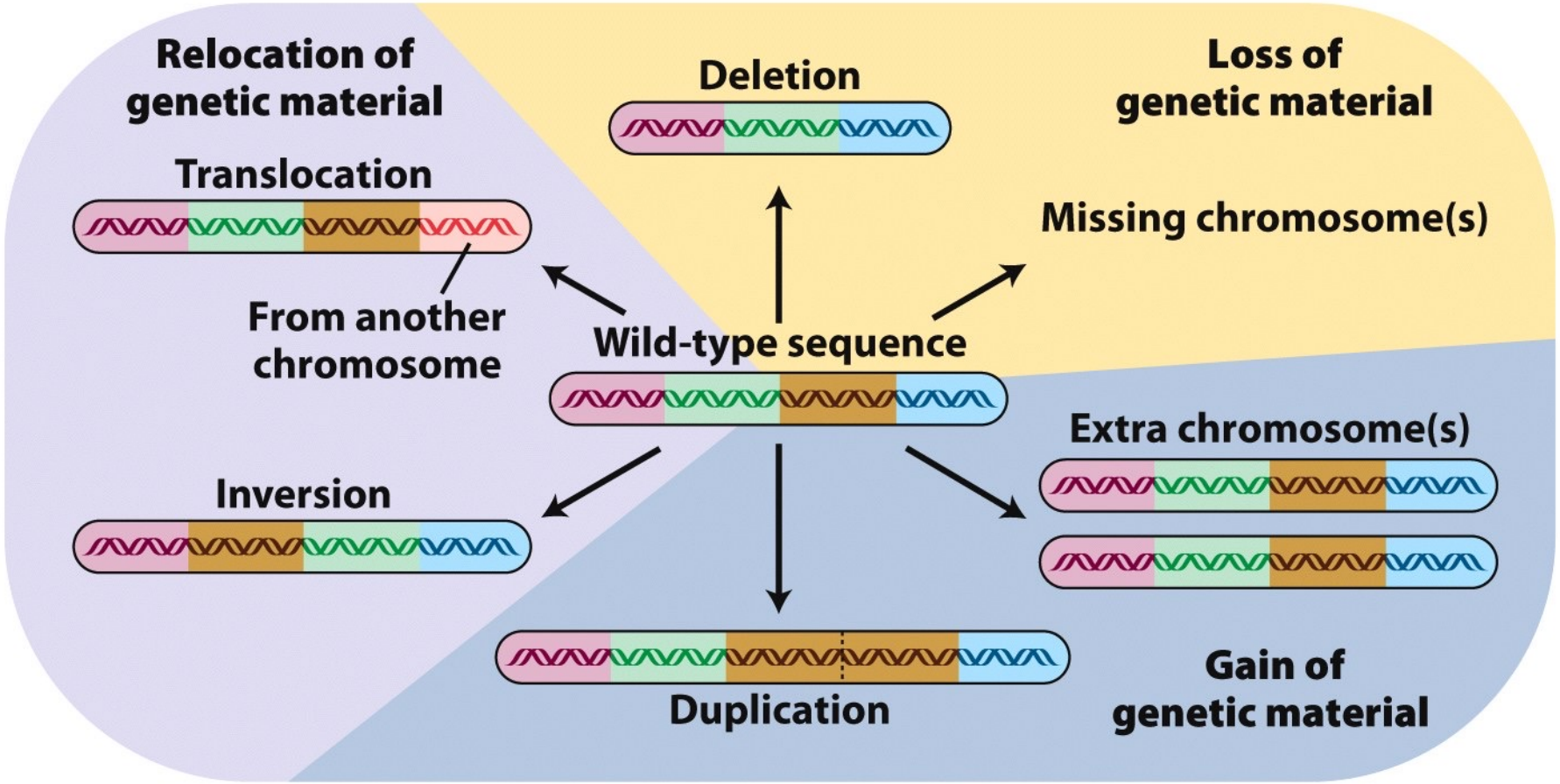
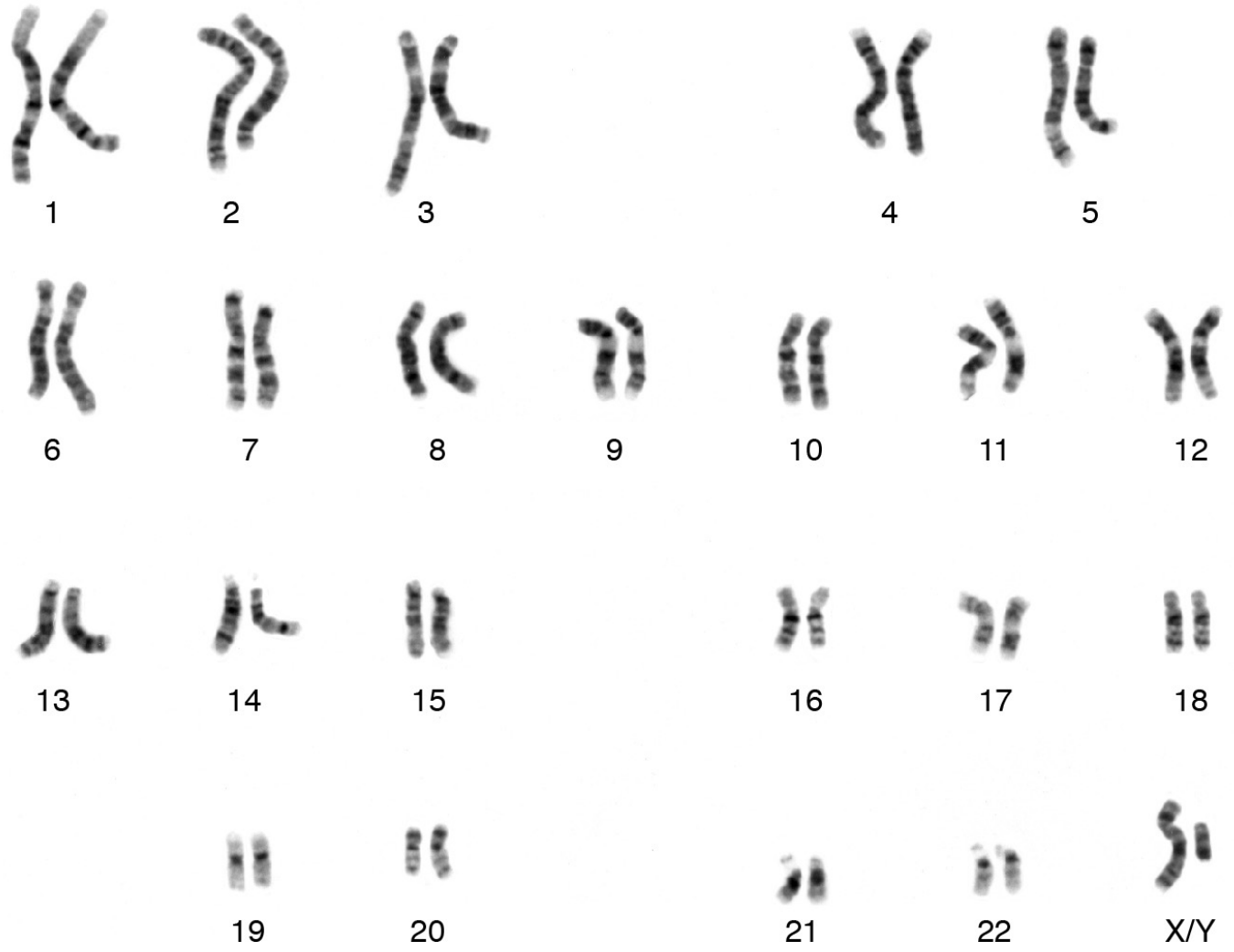


Figure 17-2
Introduction to Genetic Analysis, Tenth Edition

2. TYPES OF GENETIC VARIANTS

classification by SIZE



2. TYPES OF GENETIC VARIANTS

classification by SIZE

Turner syndrome (XO)

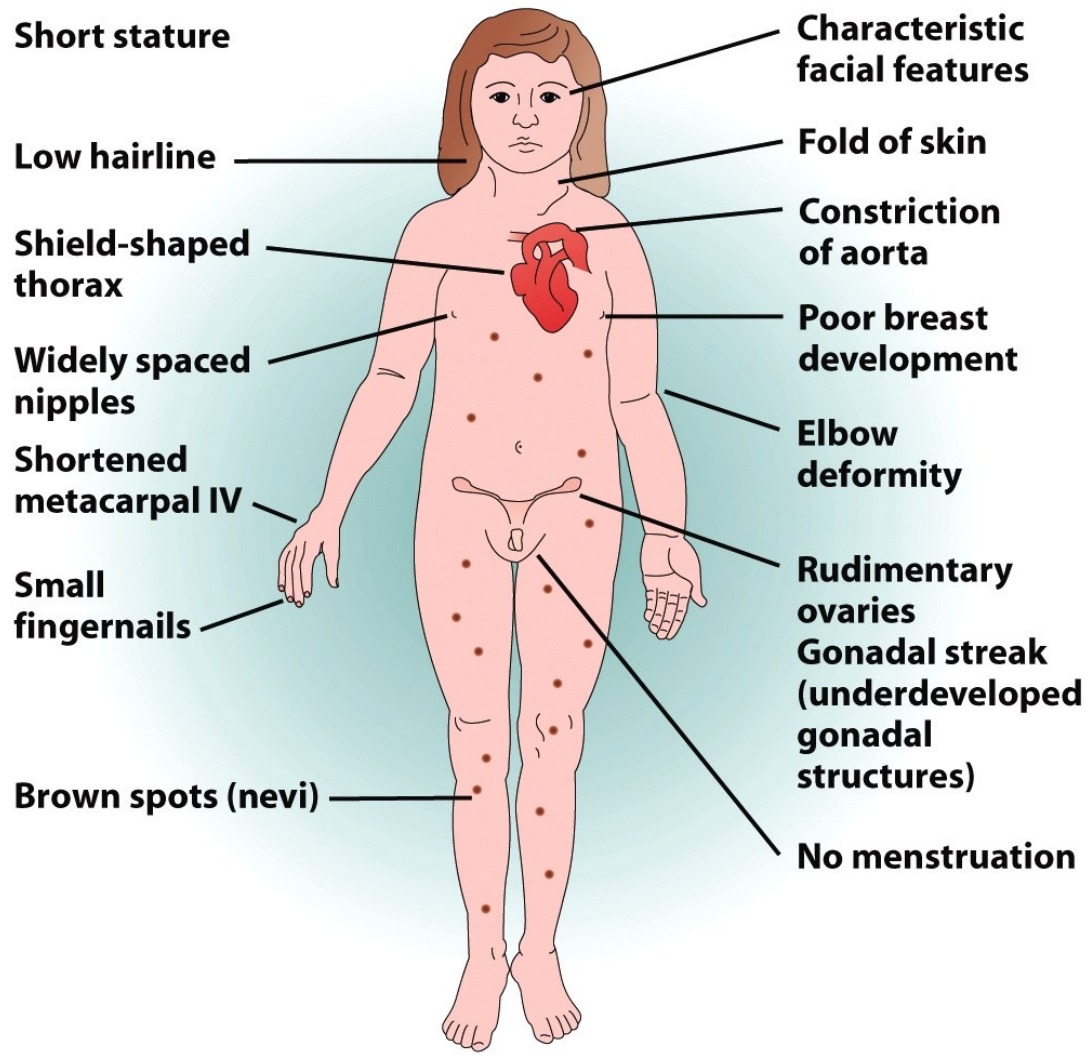


Figure 17-13
Introduction to Genetic Analysis, Tenth Edition

2. TYPES OF GENETIC VARIANTS

classification by SIZE

Klinefelter syndrome (XXY)

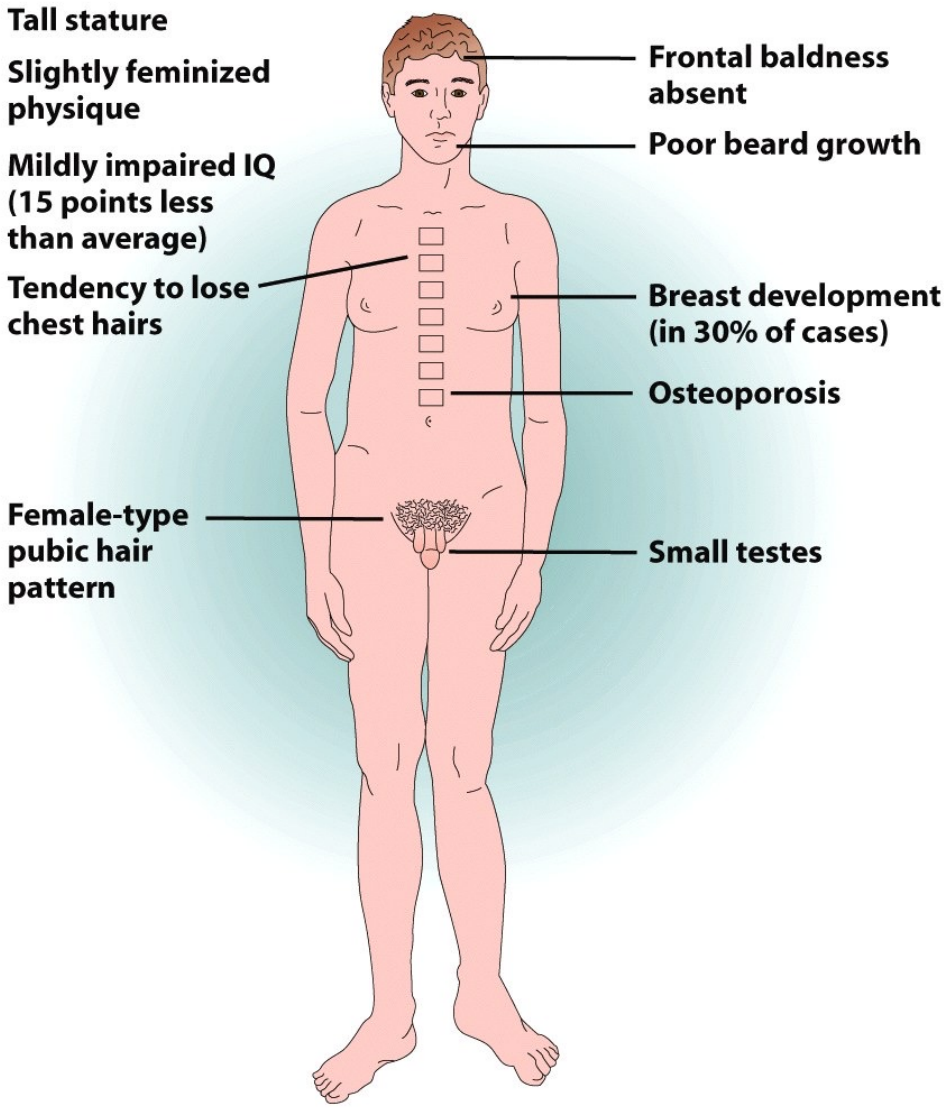


Figure 17-15
Introduction to Genetic Analysis, Tenth Edition

2. TYPES OF GENETIC VARIANTS

classification by SIZE

Down syndrome (trisomy 21)

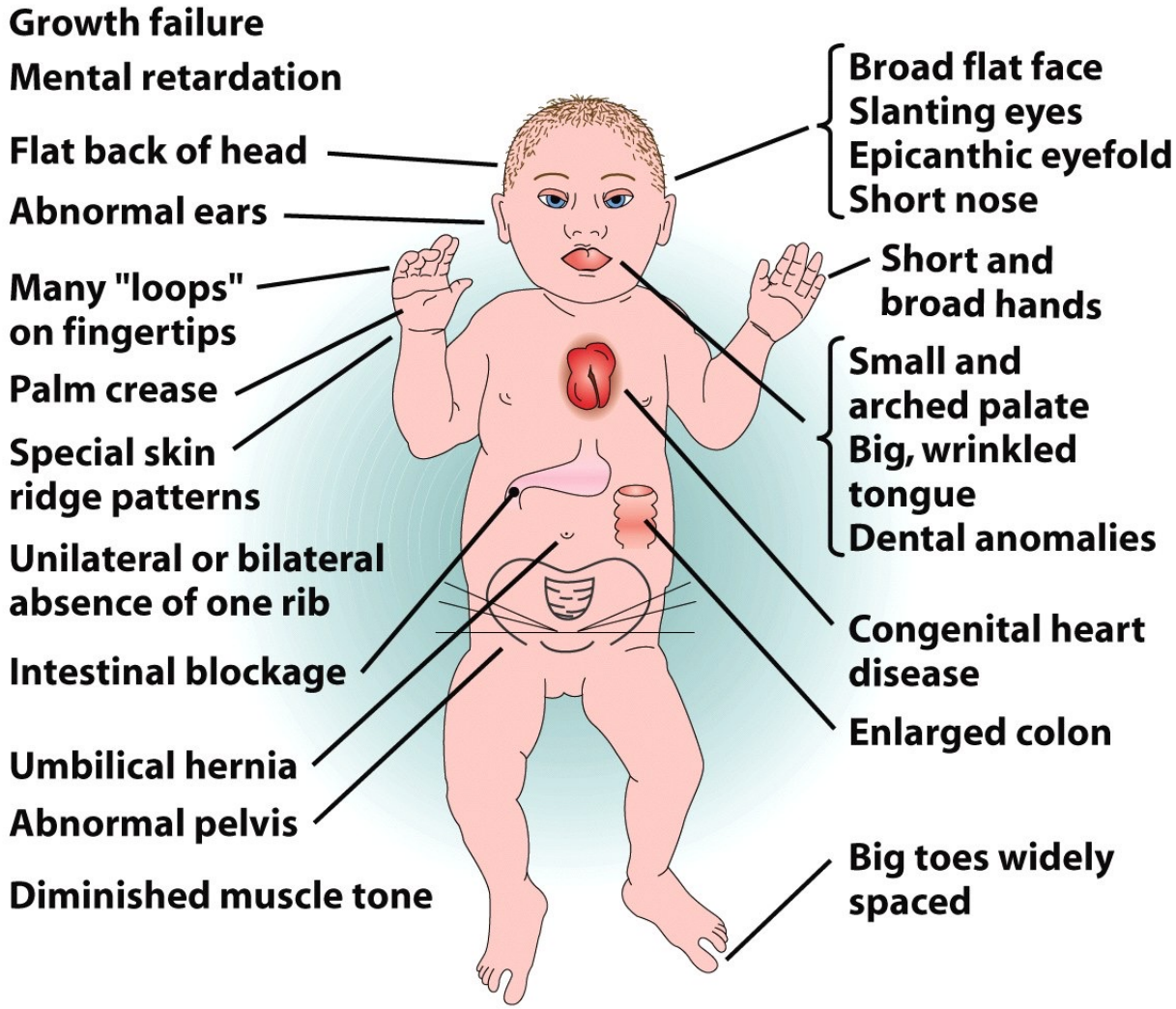


Figure 17-16
Introduction to Genetic Analysis, Tenth Edition

2. TYPES OF GENETIC VARIANTS

classification by SIZE

The maternal-age effect in Down syndrome

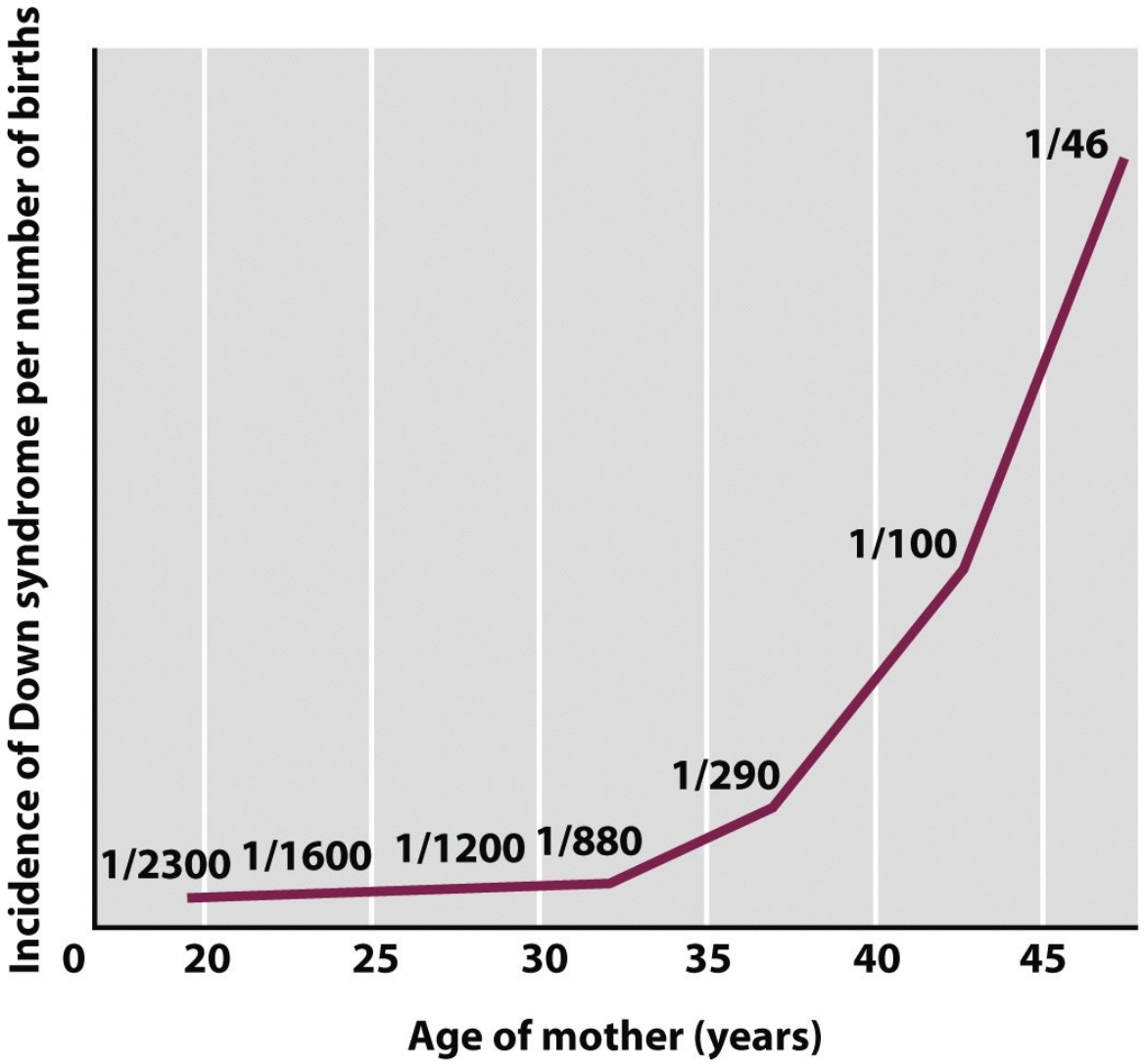


Figure 17-17
Introduction to Genetic Analysis, Tenth Edition

2. TYPES OF GENETIC VARIANTS

classification by SIZE

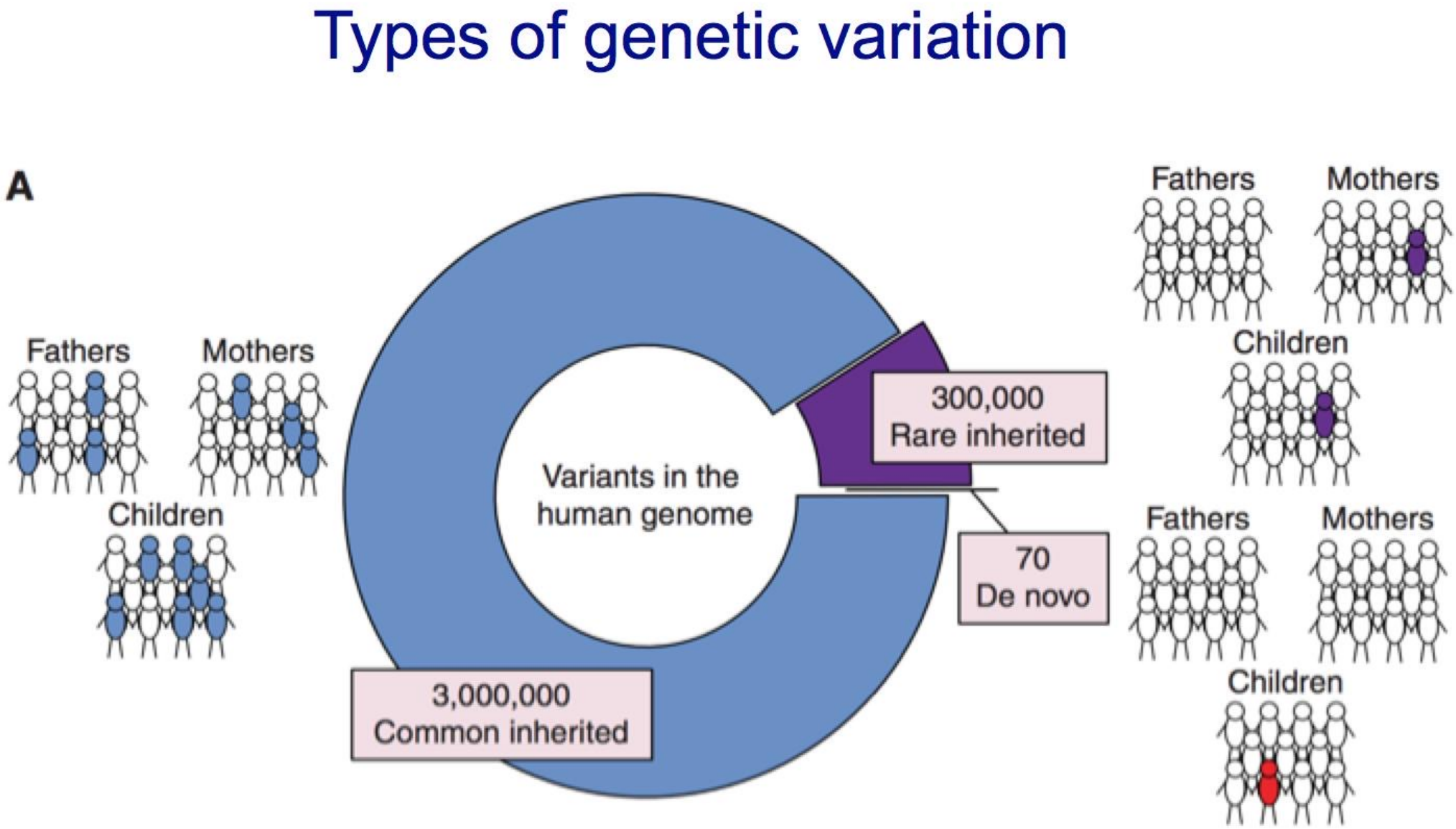
Table 1 | Median autosomal variant sites per genome

	AFR		AMR		EAS		EUR		SAS	
Samples	661		347		504		503		489	
Mean coverage	8.2		7.6		7.7		7.4		8.0	
	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons
SNPs	4.31M	14.5k	3.64M	12.0k	3.55M	14.8k	3.53M	11.4k	3.60M	14.4k
Indels	625k	-	557k	-	546k	-	546k	-	556k	-
Large deletions	1.1k	5	949	5	940	7	939	5	947	5
CNVs	170	1	153	1	158	1	157	1	165	1
MEI (Alu)	1.03k	0	845	0	899	1	919	0	889	0
MEI (L1)	138	0	118	0	130	0	123	0	123	0
MEI (SVA)	52	0	44	0	56	0	53	0	44	0
MEI (MT)	5	0	5	0	4	0	4	0	4	0
Inversions	12	0	9	0	10	0	9	0	11	0
Nonsynon	12.2k	139	10.4k	121	10.2k	144	10.2k	116	10.3k	144
Synon	13.8k	78	11.4k	67	11.2k	79	11.2k	59	11.4k	78
Intron	2.06M	7.33k	1.72M	6.12k	1.68M	7.39k	1.68M	5.68k	1.72M	7.20k
UTR	37.2k	168	30.8k	136	30.0k	169	30.0k	129	30.7k	168
Promoter	102k	430	84.3k	332	81.6k	425	82.2k	336	84.0k	430
Insulator	70.9k	248	59.0k	199	57.7k	252	57.7k	189	59.1k	243
Enhancer	354k	1.32k	295k	1.05k	289k	1.34k	288k	1.02k	295k	1.31k
TFBSs	927	4	759	3	748	4	749	3	765	3
Filtered LoF	182	4	152	3	153	4	149	3	151	3
HGMD-DM	20	0	18	0	16	1	18	2	16	0
GWAS	2.00k	0	2.07k	0	1.99k	0	2.08k	0	2.06k	0
ClinVar	28	0	30	1	24	0	29	1	27	1

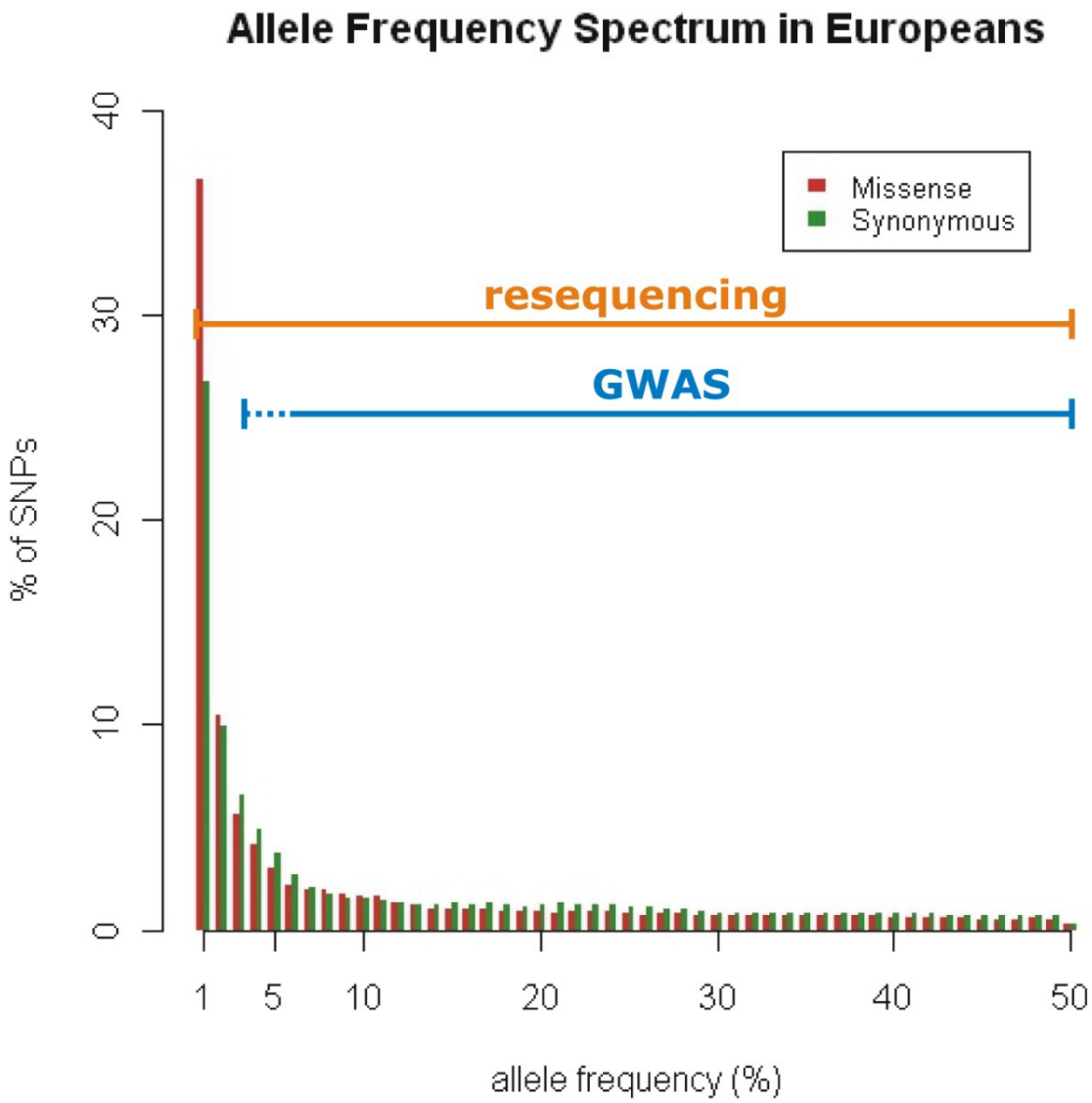
See Supplementary Table 1 for continental population groupings. CNVs, copy-number variants; HGMD-DM, Human Gene Mutation Database disease mutations; k, thousand; LoF, loss-of-function; M, million; MEI, mobile element insertions.

2. TYPES OF GENETIC VARIANTS

classification by FREQUENCY



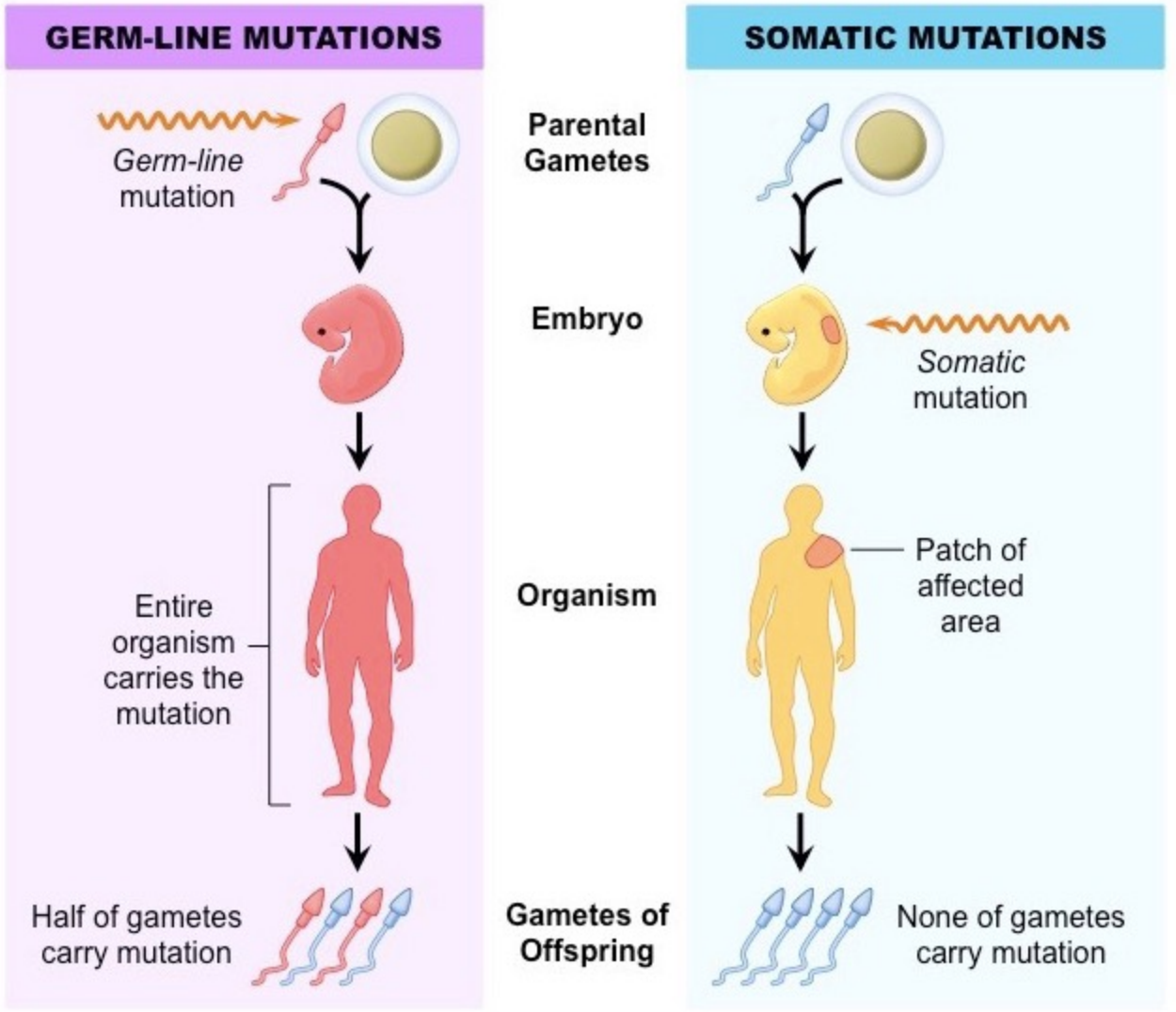
classification by FREQUENCY



From: Casals *et al.* J Neuroimmunol. 2012 *Next-generation sequencing approaches for genetic mapping of complex diseases*.

2. TYPES OF GENETIC VARIANTS

classification by ORIGEN



classification by ORIGIN

Genetic mosaicism

“The term 'mosaic' has been used since antiquity to refer to an artistic patchwork of ornamental stones, glass, gems or other precious material. At a distance, the collective image appears as it would in a painting; only on close inspection do the individual components become recognizable. In biological systems, mosaicism implies the presence of more than one genetically distinct cell line in a single organism. As in the artistic sense, mosaicism in an organism might be inapparent unless closely analysed. At the level of the whole organism, appreciation of the mosaic phenotype depends on tissue-to-tissue genetic variations that might not clearly follow Mendelian rules of inheritance”

Mechanisms and consequences of somatic mosaicism in humans, Youssoufian and Pyeritz, Nature Reviews 2002

2. TYPES OF GENETIC VARIANTS

classification by ORIGEN

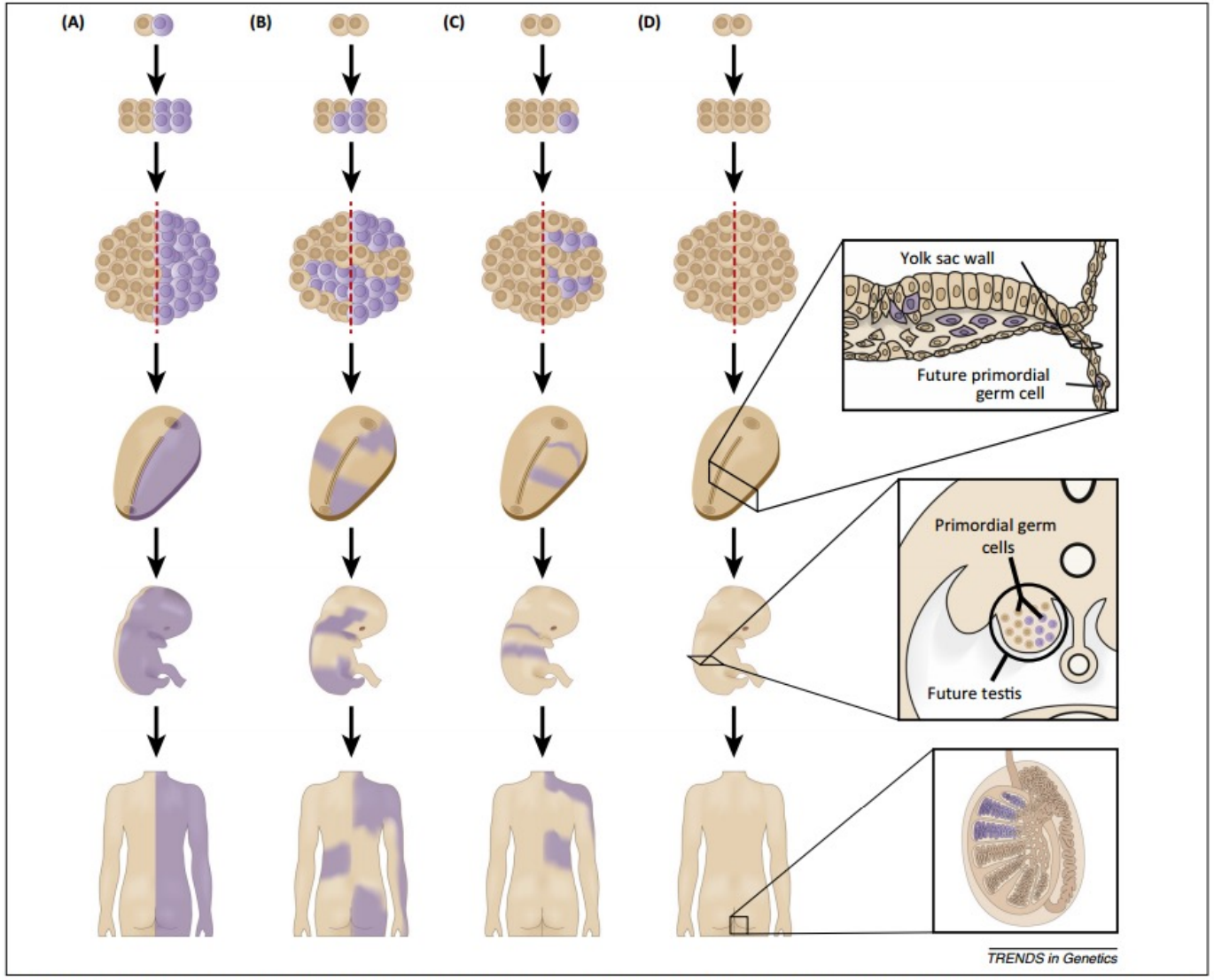


Figure 1. The timing of postzygotic mutation influences the distribution of mutant cells in the individual. **(A)** Mutations that occur during the first mitosis result in approximately half of the individual being affected. Individuals with CHILD syndrome (congenital hemidysplasia with ichthyosiform erythroderma and limb defects) have been observed with this striking pattern (see Figure 2A). **(B)** Mutations that occur before left–right determination can affect both sides of the individual, including one or both gonads. **(C)** Mutations that arise after the determination of the two sides of the embryo can be confined to only one side of the individual. Only one gonad is likely to be affected. **(D)** Mutations that occur after differentiation of primordial germ cells (PGCs) will be absent from somatic tissues. Thus, molecular investigations to detect such gonadal mosaicism must involve direct observation of germ cells. For males, this process is relatively straight forward, but for females it involves invasive biopsy of potentially both ovaries.

2. TYPES OF GENETIC VARIANTS

classification by LOCALIZATION

SNVs (3.000.000)

INTERGENIC ~ 62 %

INTRAGENIC ~ 38 %

Coding ~ 6 %

SNPs		
INTERGENIC		**2.054.900**
INTRAGENIC		**1.252.778**
Intron	1.043.427	
UTR	181.267	
Spl acc	1.898	
Spl don	398	
Coding	18.796	
Synonymous	9.612	
Missense	9.082	
Nonsense	87	
StopGain	15	
TOTAL		**3.307.678**

Data from *Comprehensive Characterization of Human Genome Variation by High Coverage Whole-Genome Sequencing of Forty Four Caucasians,* Shen *et al.* 2013

2. TYPES OF GENETIC VARIANTS

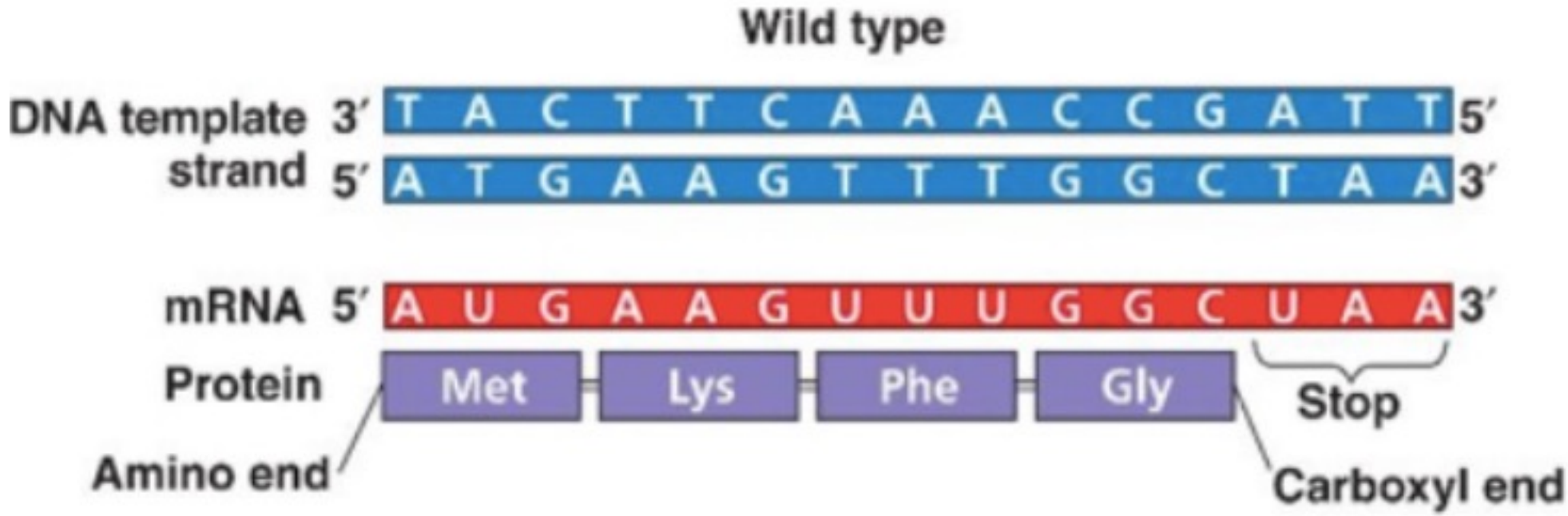
classification by LOCALIZATION

Table 1 | Median autosomal variant sites per genome

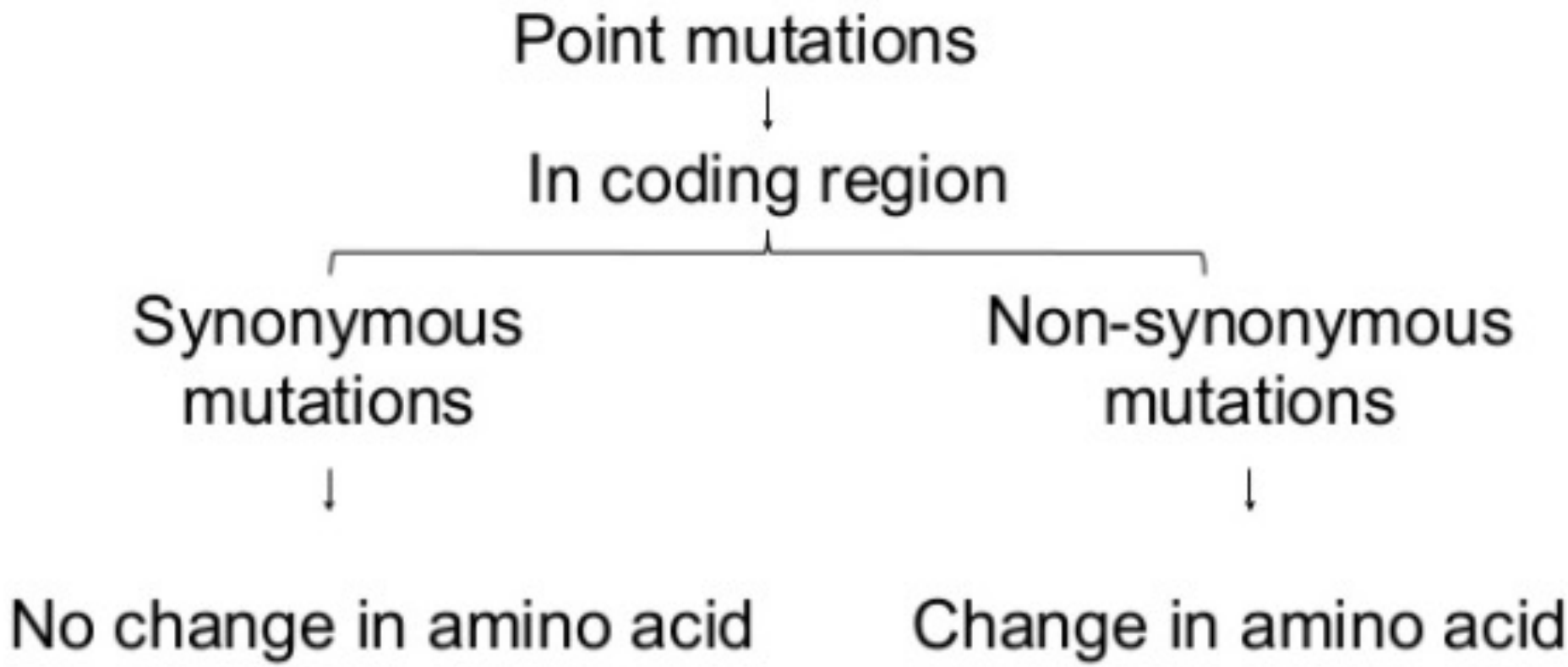
	AFR		AMR		EAS		EUR		SAS	
Samples	661		347		504		503		489	
Mean coverage	8.2		7.6		7.7		7.4		8.0	
	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons
SNPs	4.31M	14.5k	3.64M	12.0k	3.55M	14.8k	3.53M	11.4k	3.60M	14.4k
Indels	625k	-	557k	-	546k	-	546k	-	556k	-
Large deletions	1.1k	5	949	5	940	7	939	5	947	5
CNVs	170	1	153	1	158	1	157	1	165	1
MEI (Alu)	1.03k	0	845	0	899	1	919	0	889	0
MEI (L1)	138	0	118	0	130	0	123	0	123	0
MEI (SVA)	52	0	44	0	56	0	53	0	44	0
MEI (MT)	5	0	5	0	4	0	4	0	4	0
Inversions	12	0	9	0	10	0	9	0	11	0
Nonsynon	12.2k	139	10.4k	121	10.2k	144	10.2k	116	10.3k	144
Synon	13.8k	78	11.4k	67	11.2k	79	11.2k	59	11.4k	78
Intron	2.06M	7.33k	1.72M	6.12k	1.68M	7.39k	1.68M	5.68k	1.72M	7.20k
UTR	37.2k	168	30.8k	136	30.0k	169	30.0k	129	30.7k	168
Promoter	102k	430	84.3k	332	81.6k	425	82.2k	336	84.0k	430
Insulator	70.9k	248	59.0k	199	57.7k	252	57.7k	189	59.1k	243
Enhancer	354k	1.32k	295k	1.05k	289k	1.34k	288k	1.02k	295k	1.31k
TFBSs	927	4	759	3	748	4	749	3	765	3
Filtered LoF	182	4	152	3	153	4	149	3	151	3
HGMD-DM	20	0	18	0	16	1	18	2	16	0
GWAS	2.00k	0	2.07k	0	1.99k	0	2.08k	0	2.06k	0
ClinVar	28	0	30	1	24	0	29	1	27	1

See Supplementary Table 1 for continental population groupings. CNVs, copy-number variants; HGMD-DM, Human Gene Mutation Database disease mutations; k, thousand; LoF, loss-of-function; M, million; MEI, mobile element insertions.

SNVs in the coding region

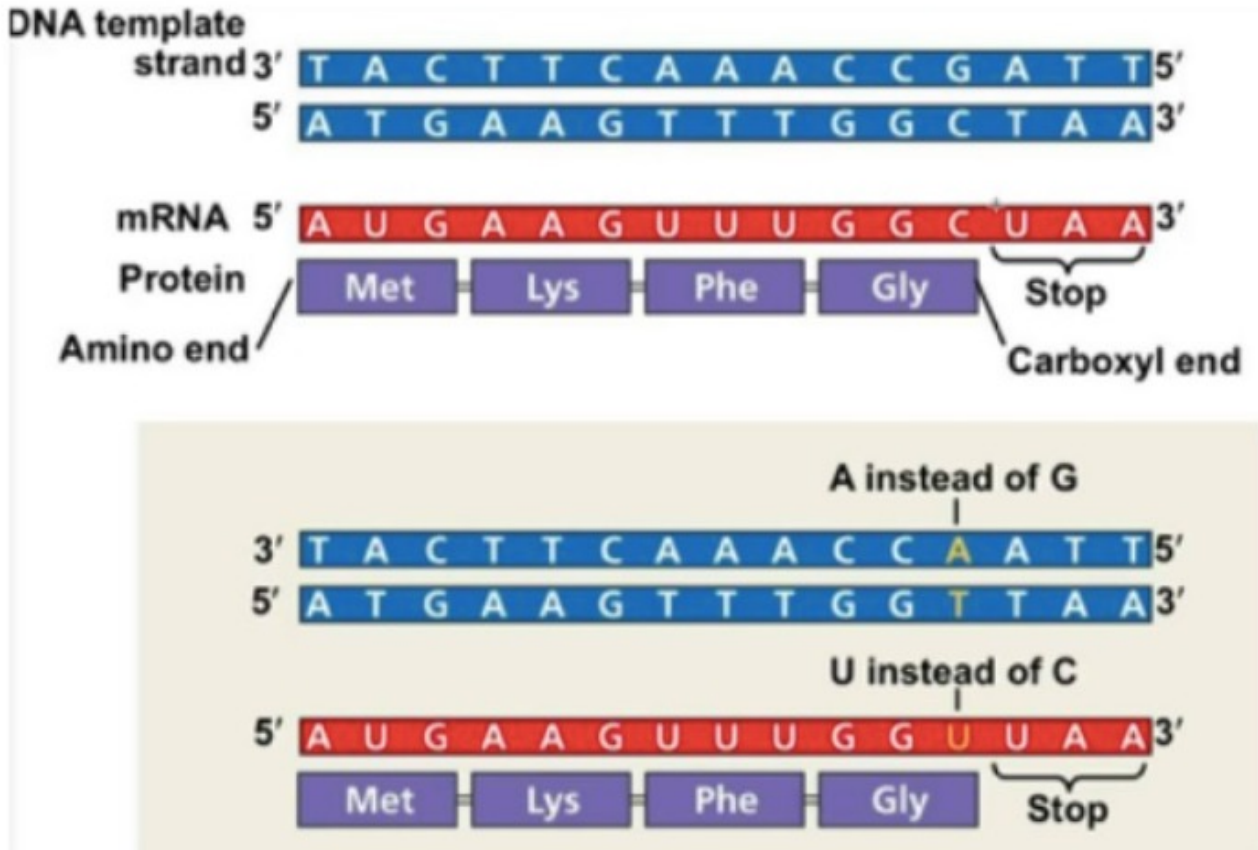


SNVs in the coding region



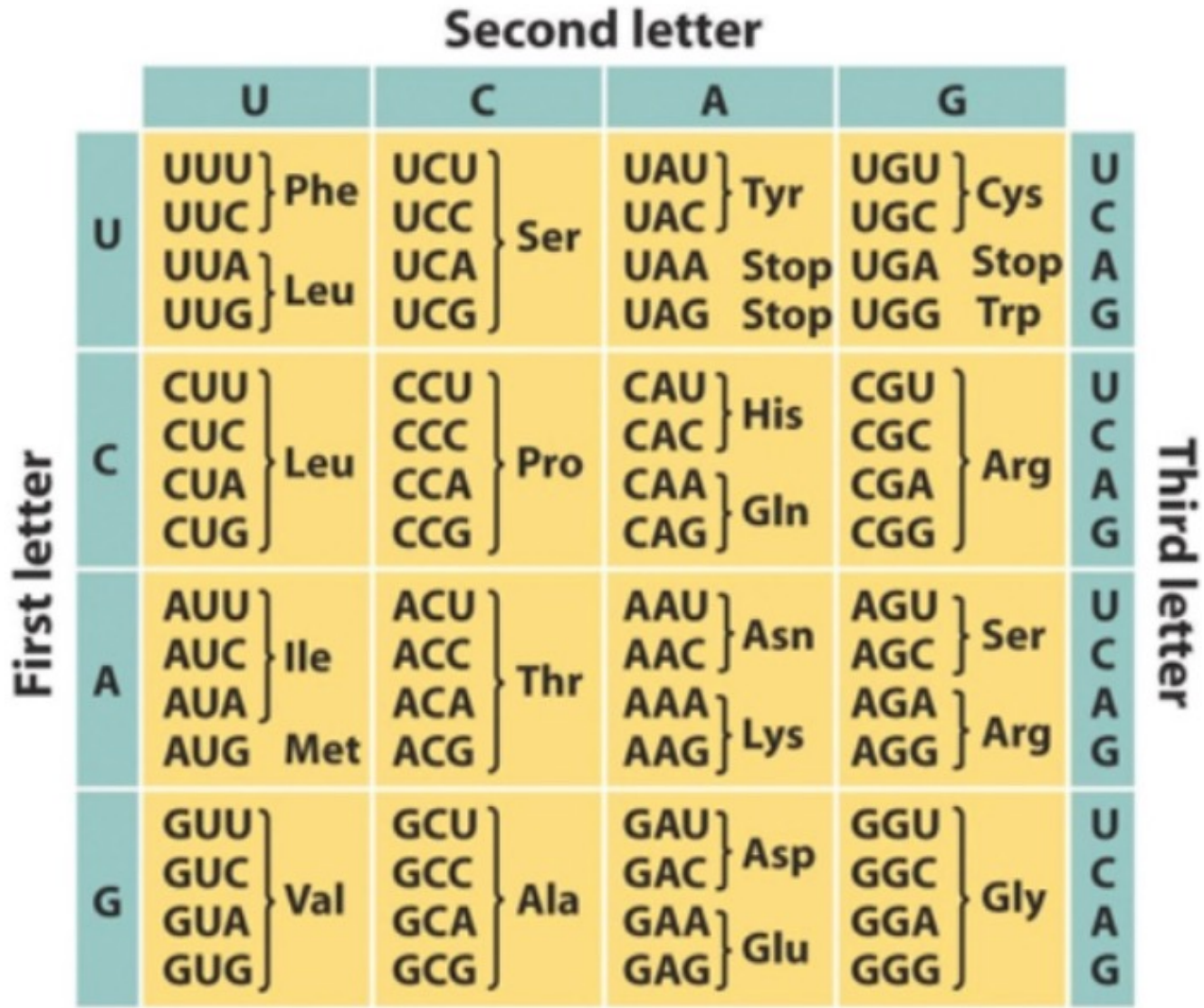
SNVs in the coding region

A change in the third nucleotide of codon GGC to GGT does not change the amino acid Glycine to another one.



[SNVs in the coding region]

The genetic code is degenerate



SNVs in the coding region

Point mutations

↓

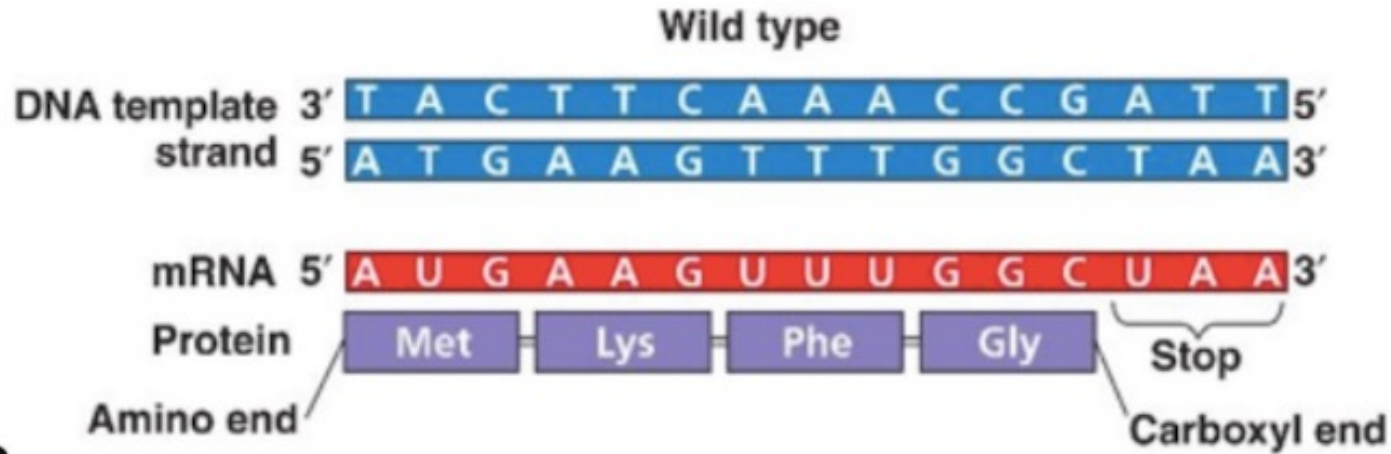
In coding region

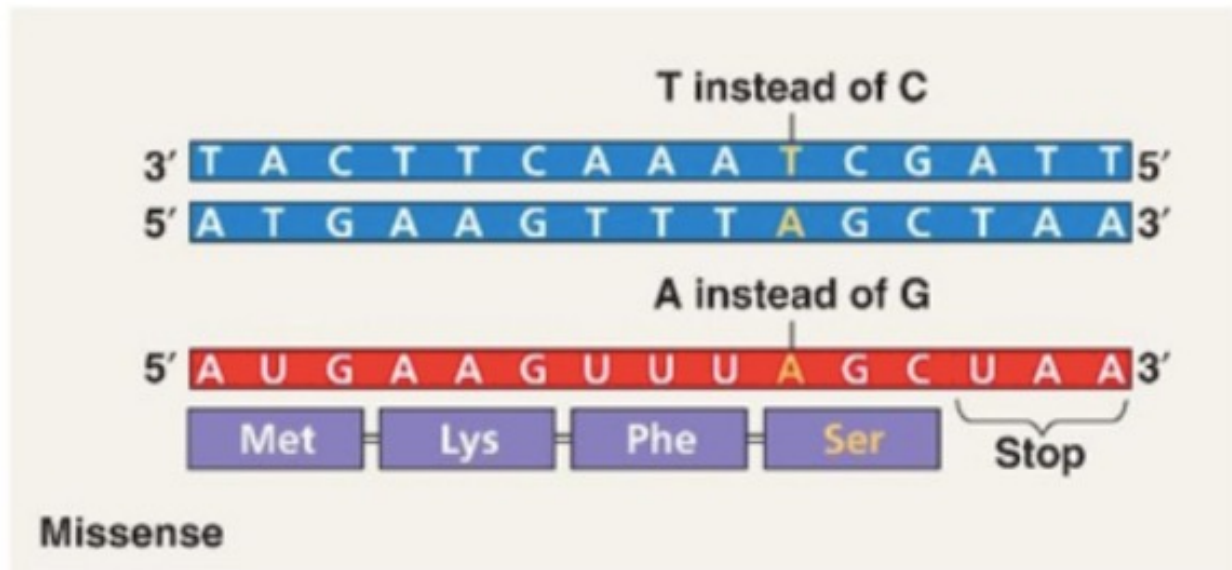
↓

Non-synonymous mutations

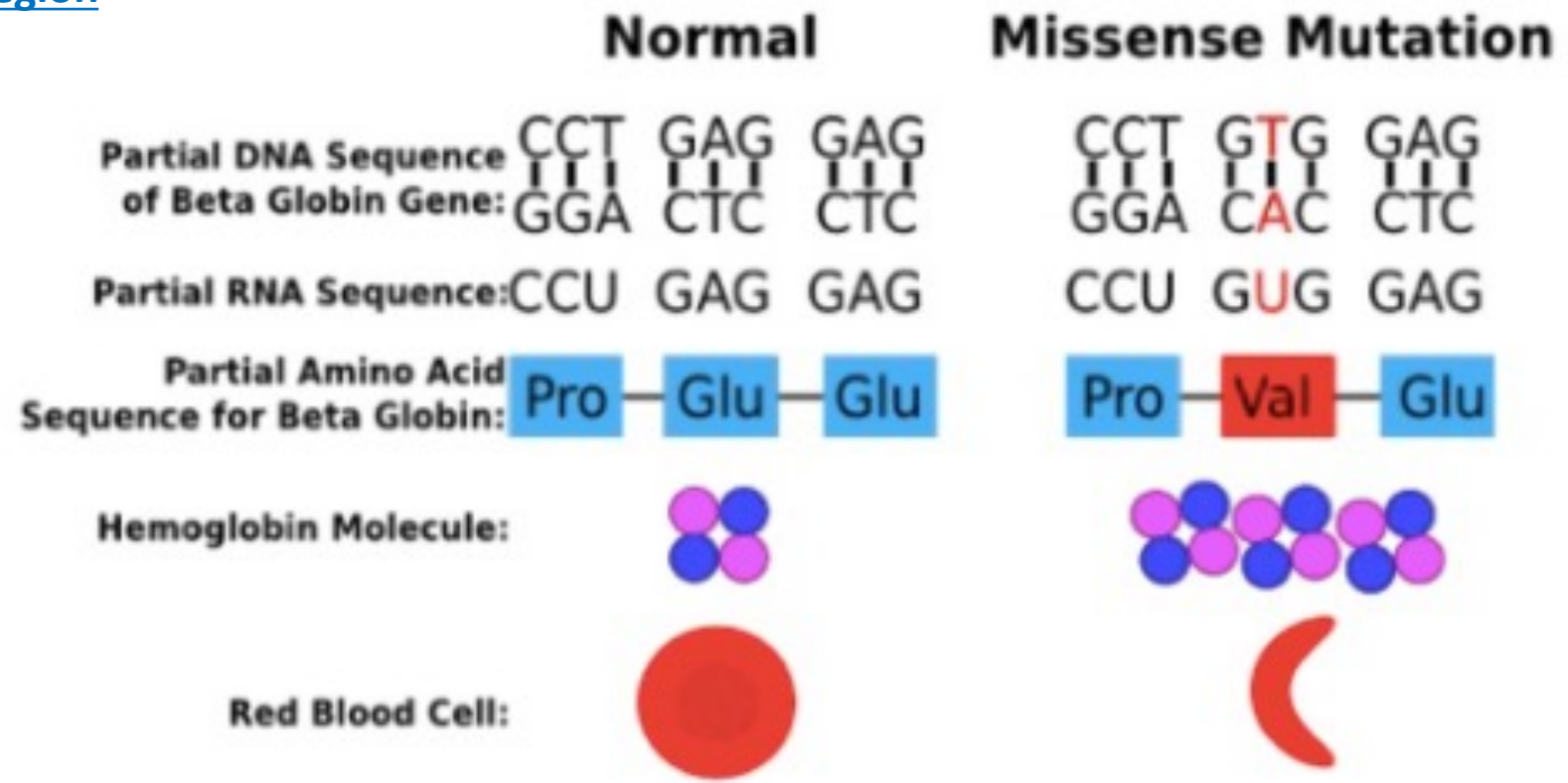
↓

Missense mutation





SNVs in the coding region



- Missense mutations may alter the function of the protein by substituting an amino acid with an unfavorable one.

- Example: hemoglobin and sickle cell anemia.

SNVs in the coding region

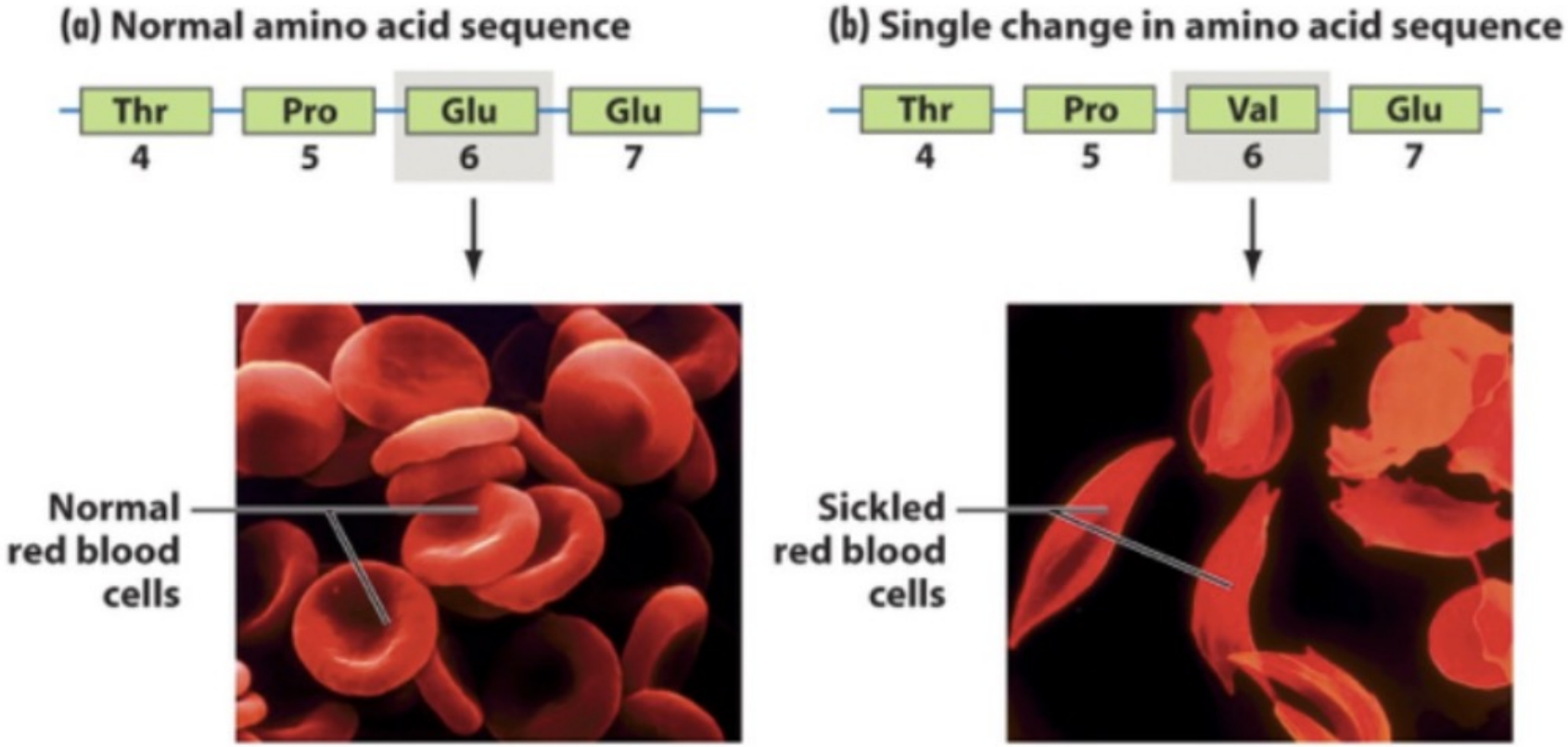


Figure 3-13 Biological Science, 2/e

A change from glutamic acid (negatively charged) to a valine (non-polar) causes a severe change in the protein function and

SNVs in the coding region

Point mutations

↓

In coding region

↓

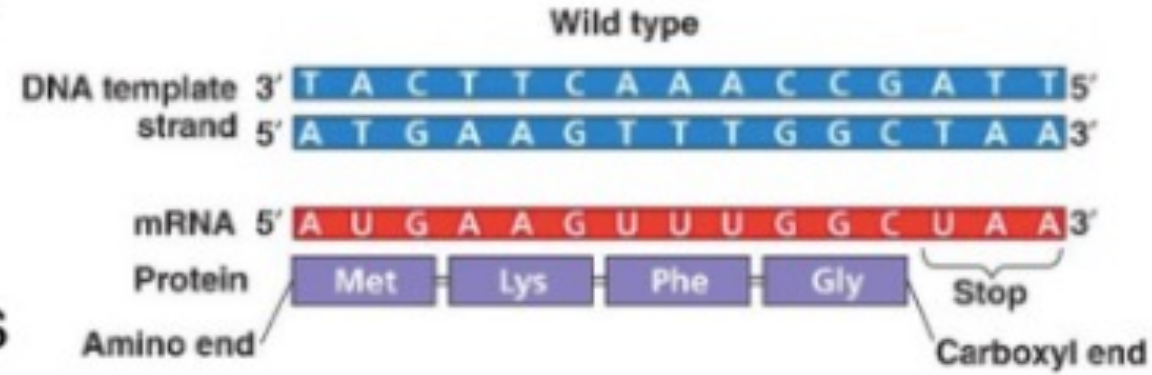
Non-synonymous mutations

↓

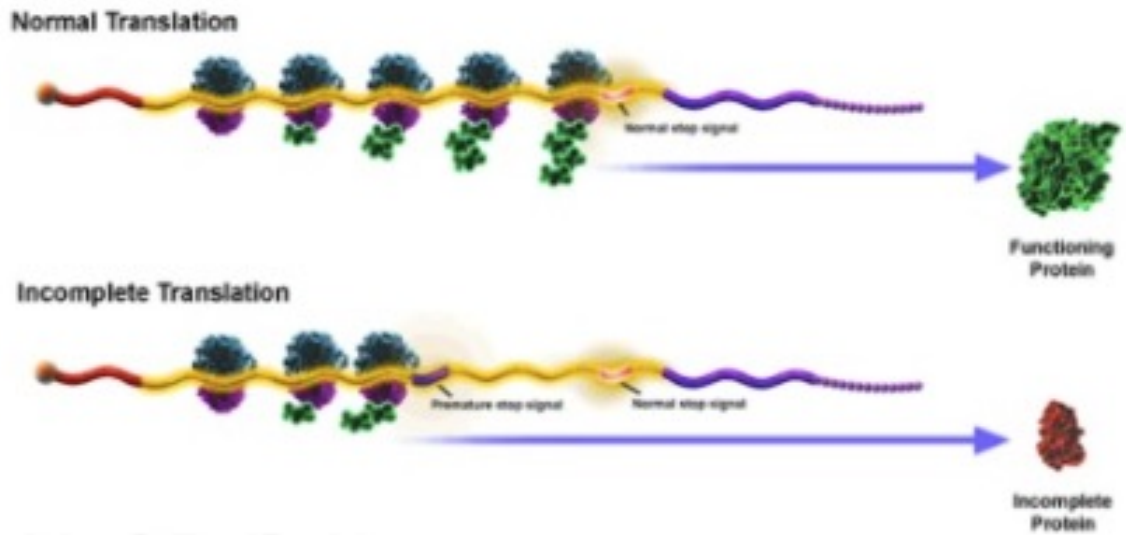
Nonsense mutation

↓

Change an amino acid into a stop codon



SNVs in the coding region



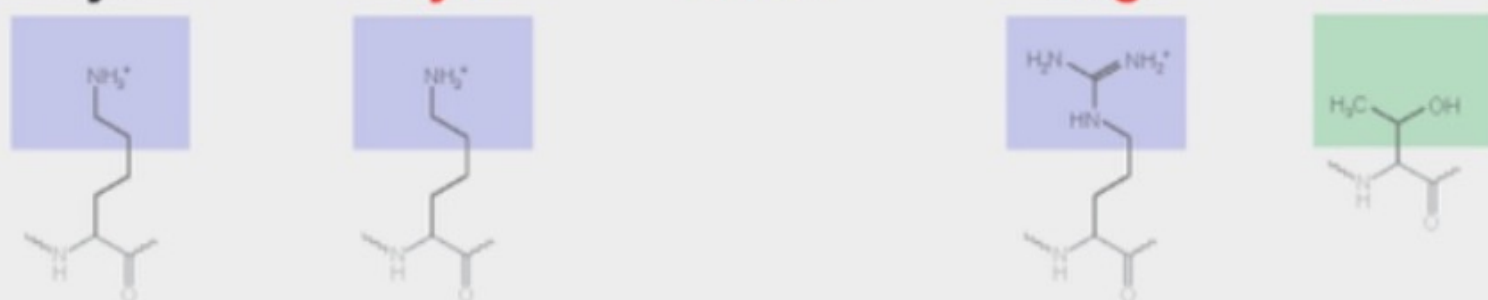
- Nonsense mutation occurs early the gene results in the truncation (shorter) of protein and loss of function (**truncated protein**).

- Nonsense mutation occurs later in the sequence of the gene results in the protein loss of function or reduction in function.

SNVs in the coding region

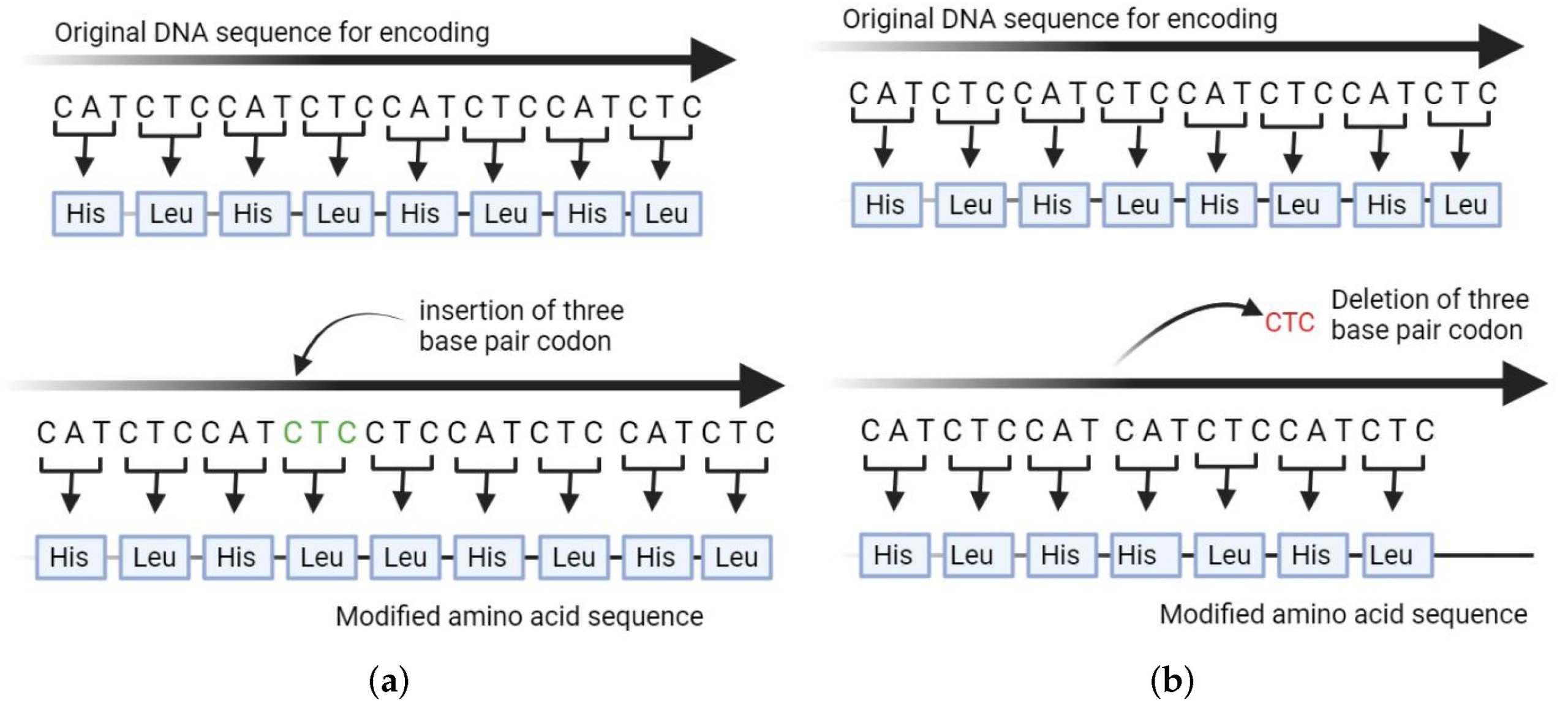
Summary of substation mutations in coding region

	No mutation	Point mutations			
		Silent	Nonsense	Missense	
				conservative	non-conservative
DNA level	TTC	TTT	ATC	TCC	TGC
mRNA level	AAG	AAA	UAG	AGG	ACG
protein level	Lys	Lys	STOP	Arg	Thr



basic
polar

In-frame indels



Frameshift indels