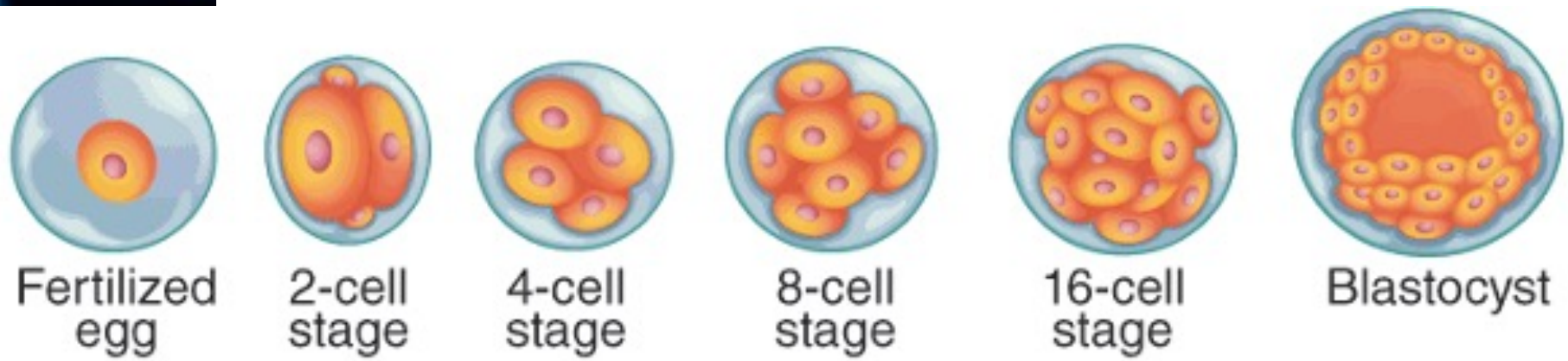
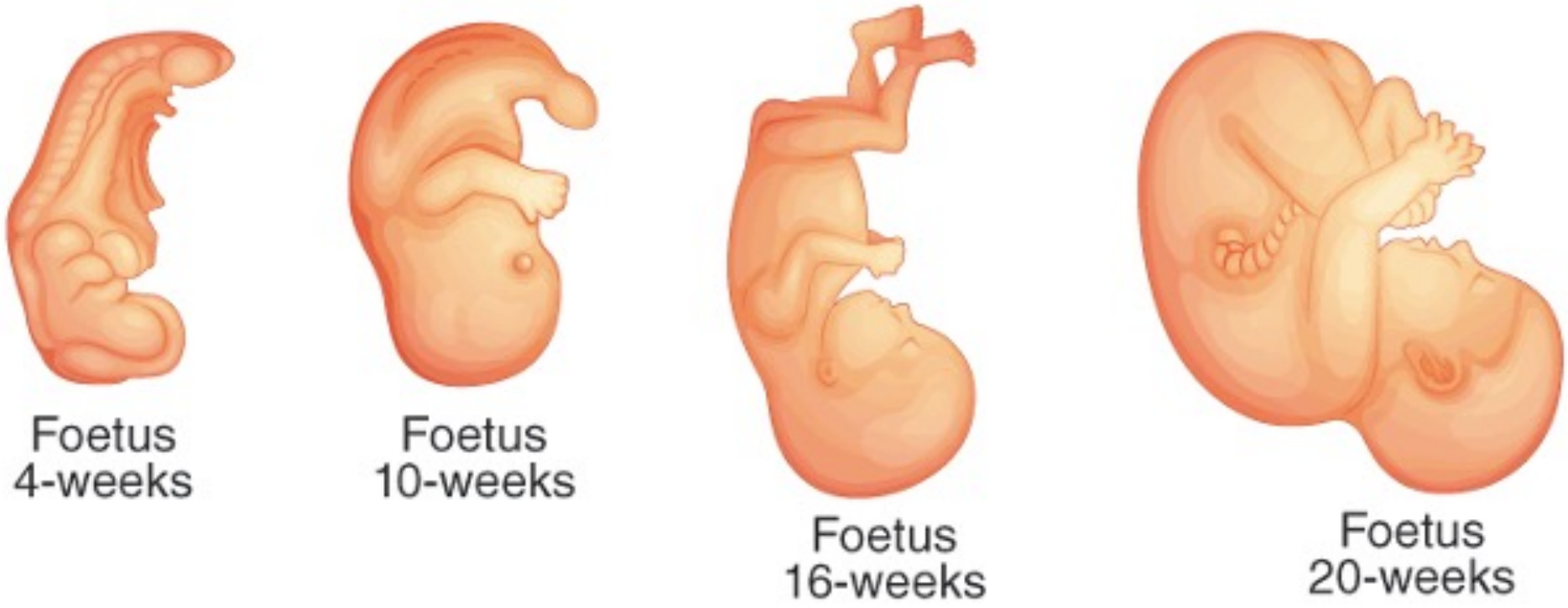
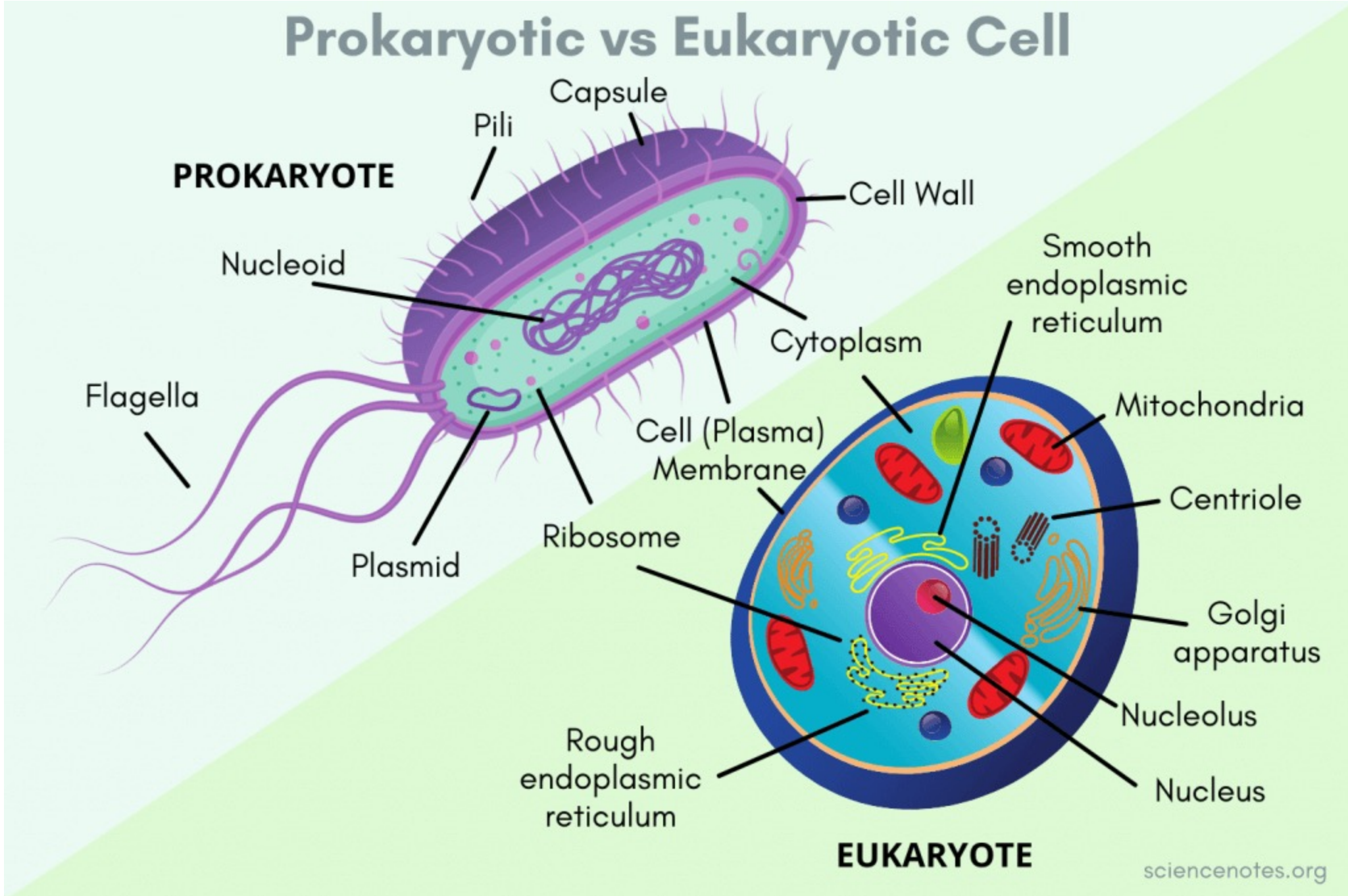
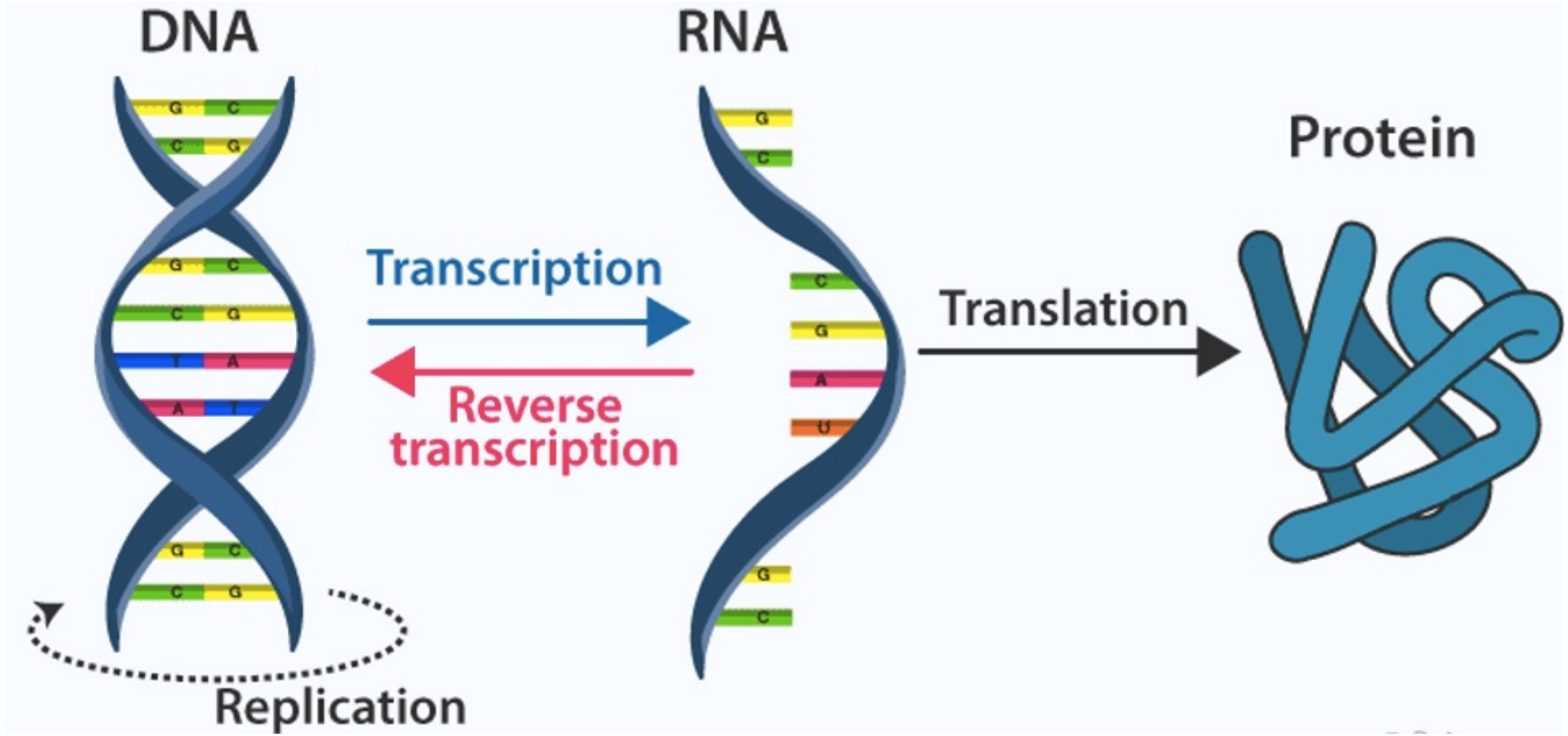

Fertilized egg
2-cell stage
4-cell stage
8-cell stage
16-cell stage
Blastocyst


Foetus 4-weeks
Foetus 10-weeks
Foetus 16-weeks
Foetus 20-weeks


Prokaryotic vs Eukaryotic Cell
PROKARYOTE
Capsule
Pili
Cell Wall
Nucleoid
Cytoplasm
Flagella
Cell (Plasma) Membrane
Ribosome
Plasmid
Smooth endoplasmic reticulum
Mitochondria
Centriole
Golgi apparatus
Nucleolus
Nucleus
Rough endoplasmic reticulum
EUKARYOTE
sciencenotes.org


DNA
RNA
Protein
Transcription
Reverse transcription
Translation
Replication

DNA replication is the basis for the perpetuation of life through time

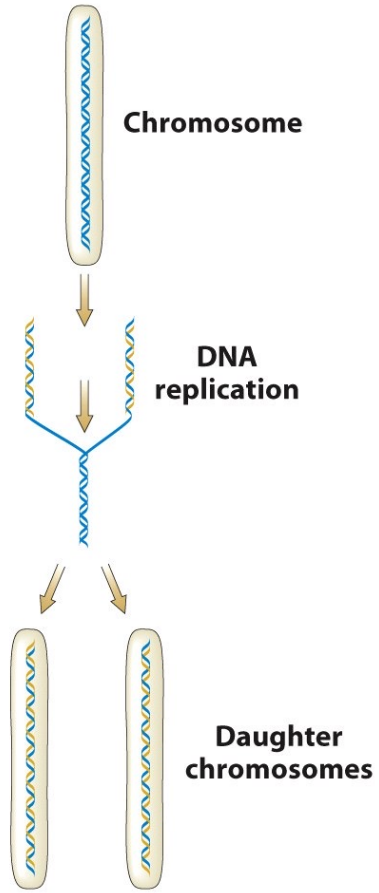


Figure 1-14
Introduction to Genetic Analysis, Tenth Edition

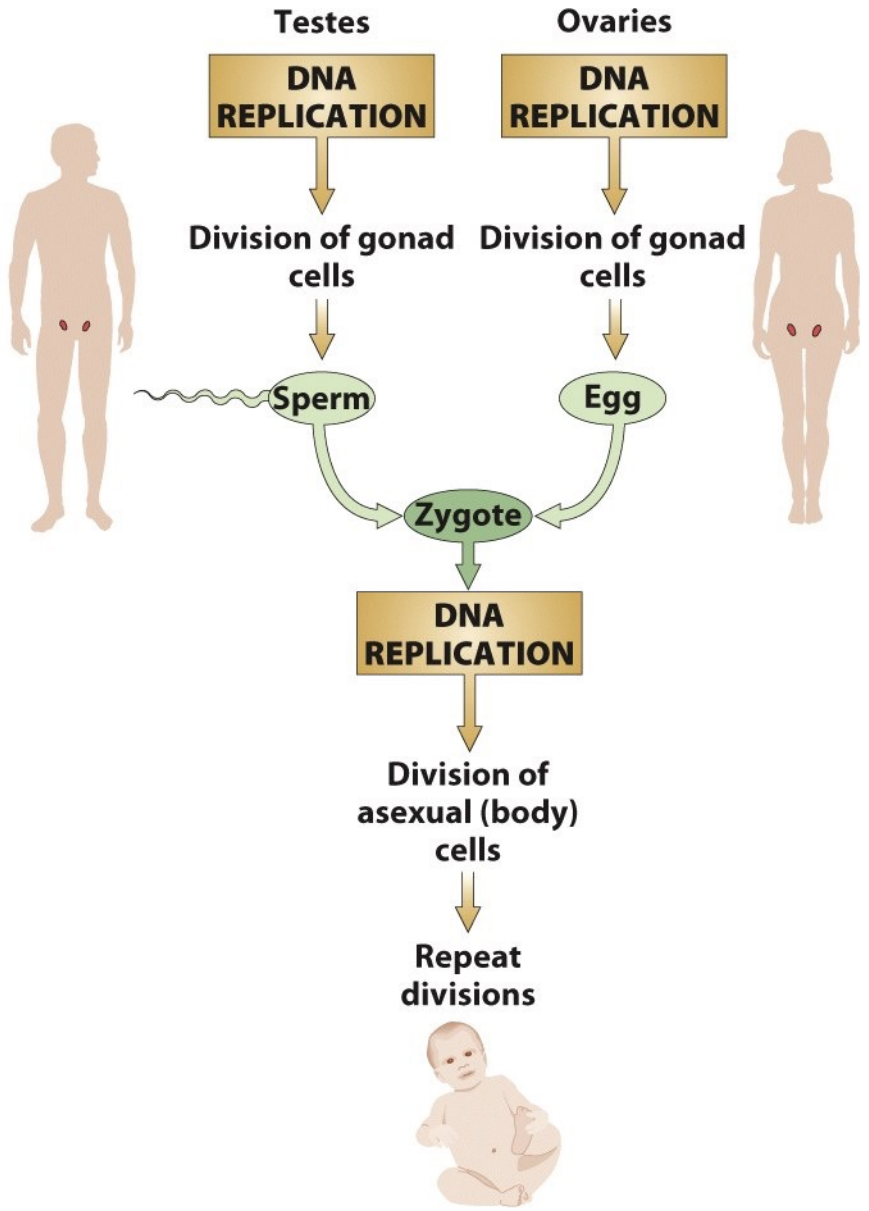


Figure 1-15
Introduction to Genetic Analysis, Tenth Edition

Key stages of meiosis and mitosis

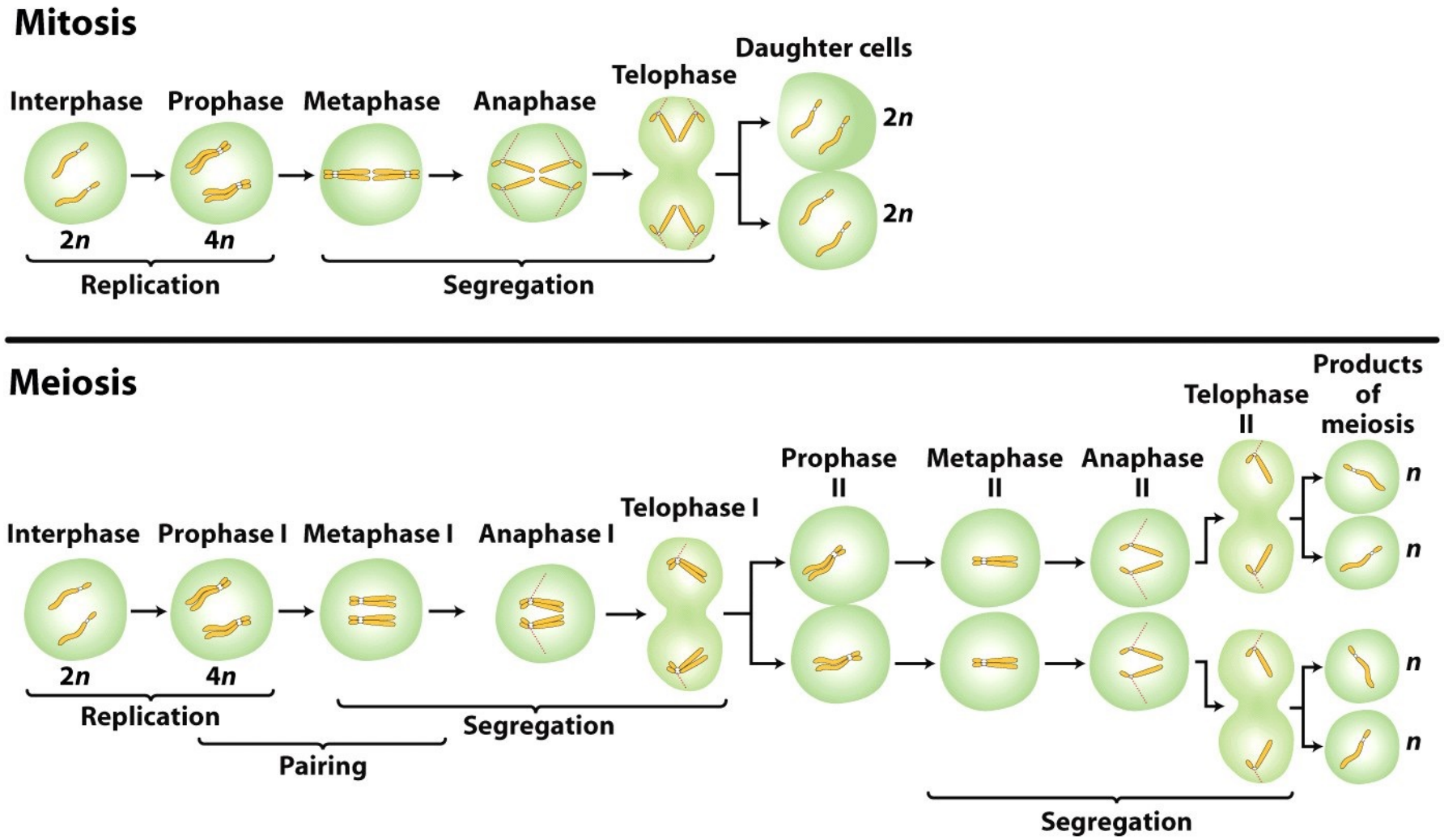
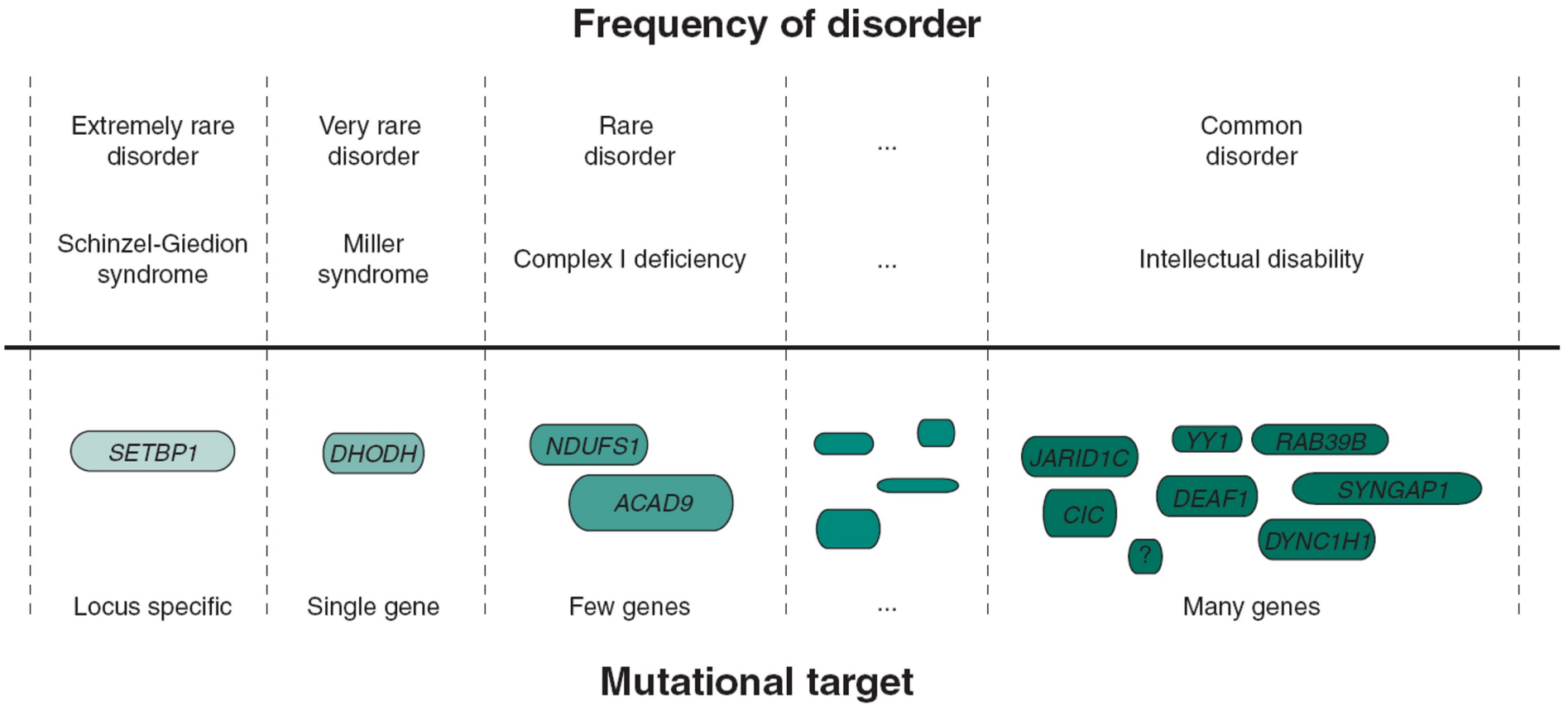
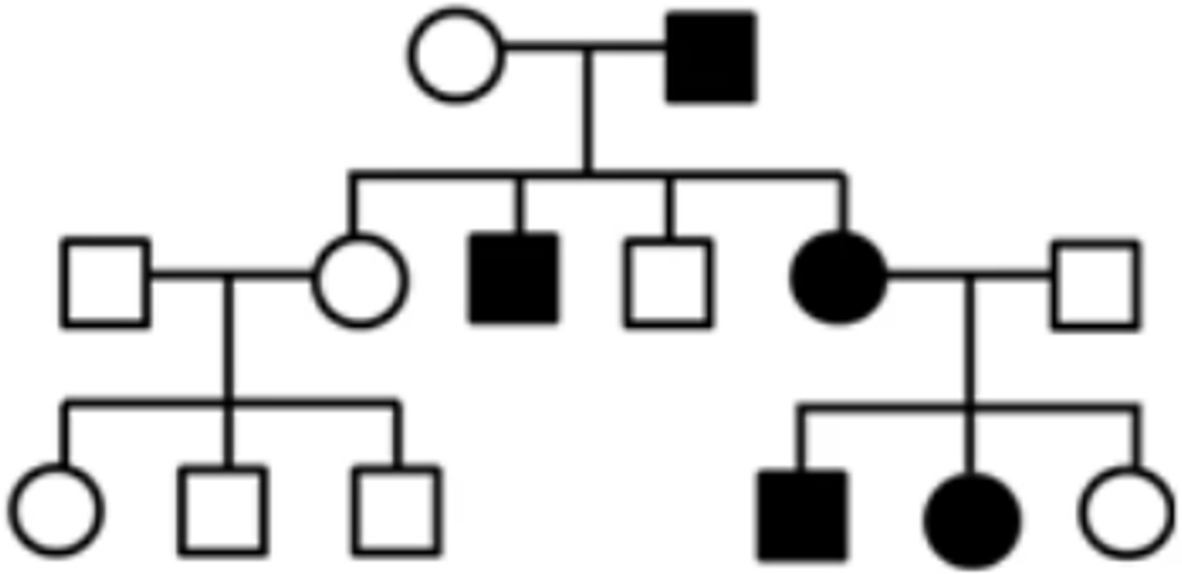


Figure 2-8
Introduction to Genetic Analysis, Tenth Edition

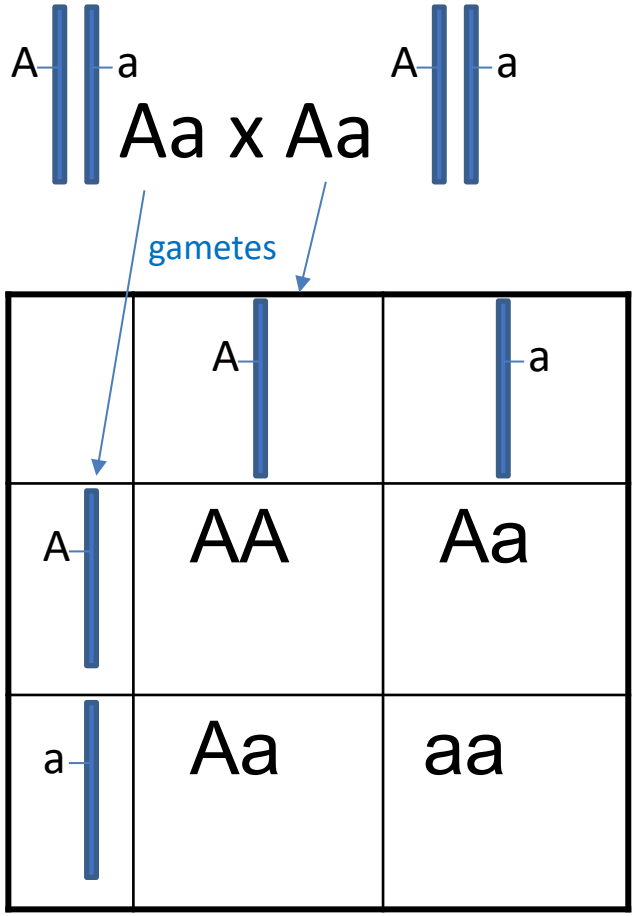
Disease prevalence and genetic model



From: Gilissen *et al. Genome Biology 2011,* ***12:228***



?



Autosomal recessive

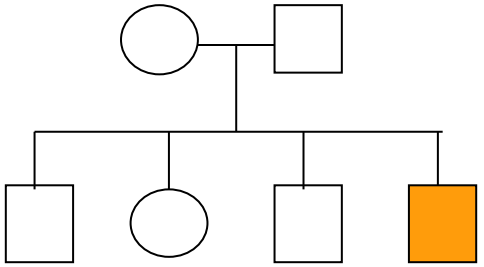
P (healthy) = 3/4
P (affected) = 1/4

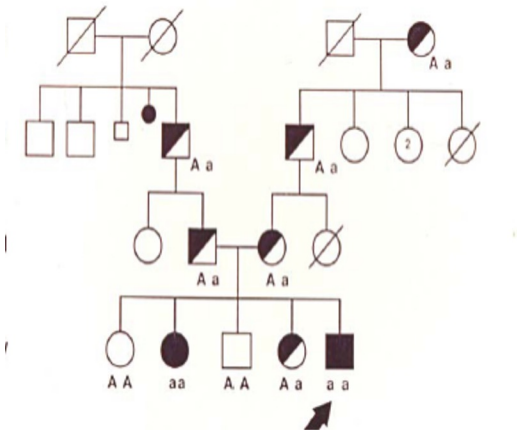
Two possible phenotypes
Three possible genotypes

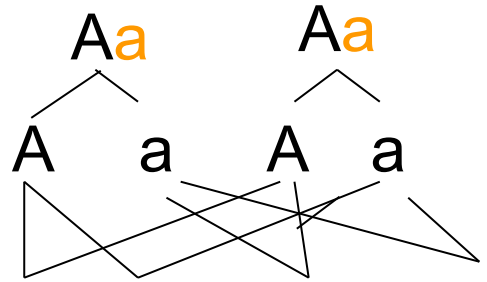
Autosomal recessive inheritance

Two copies of the mutant allele are necessary to produce an increase in risk, or equivalently, one copy of the normal allele is sufficient to provide protection.

Cystic Fibrosis (Chromosome 17) 1/ 2.000

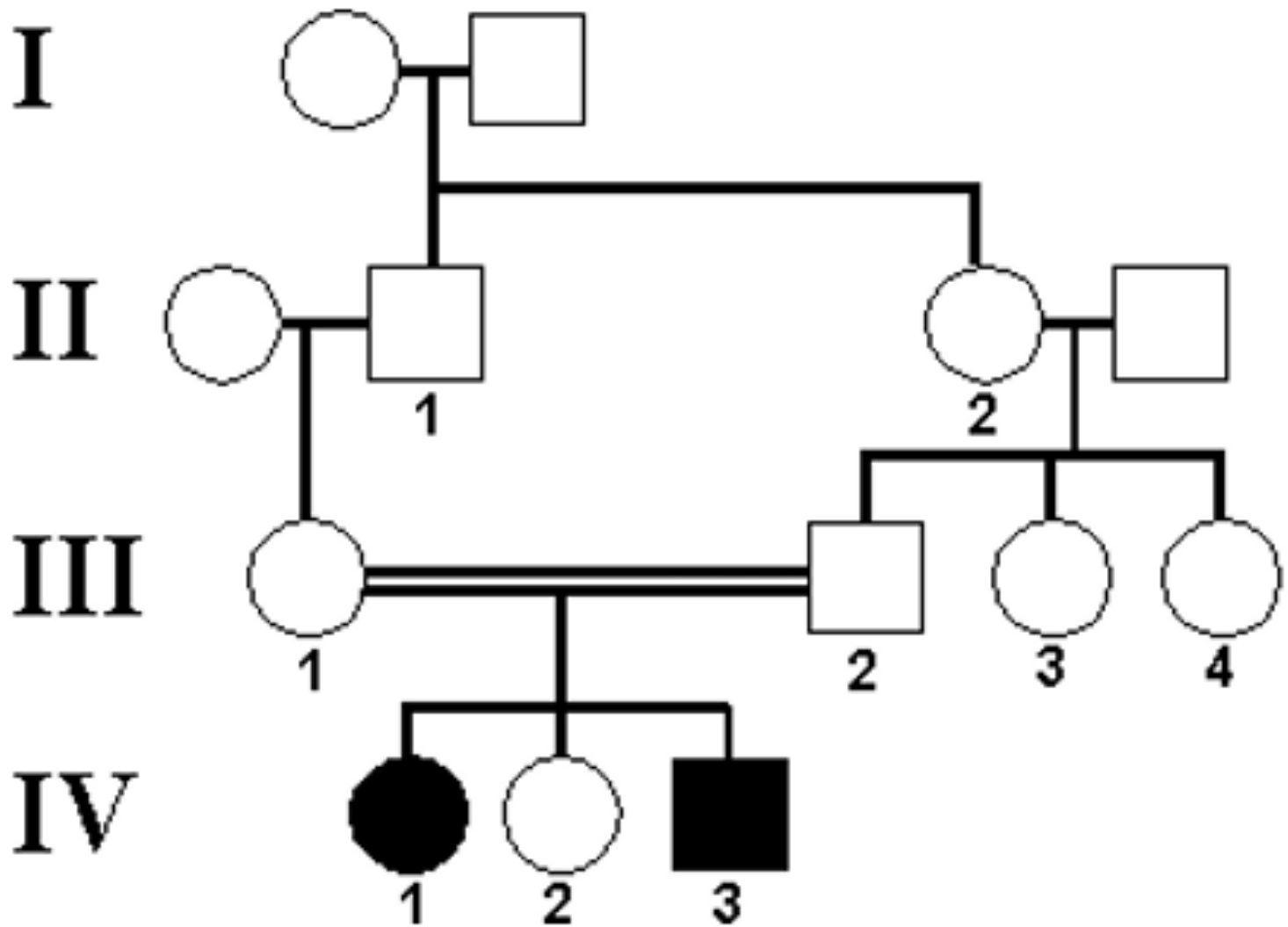






AA Aa Aa aa

¼ healthy ½ carriers ¼ affected


I
II
1
2
III
1
2
3
4
IV
1
2
3

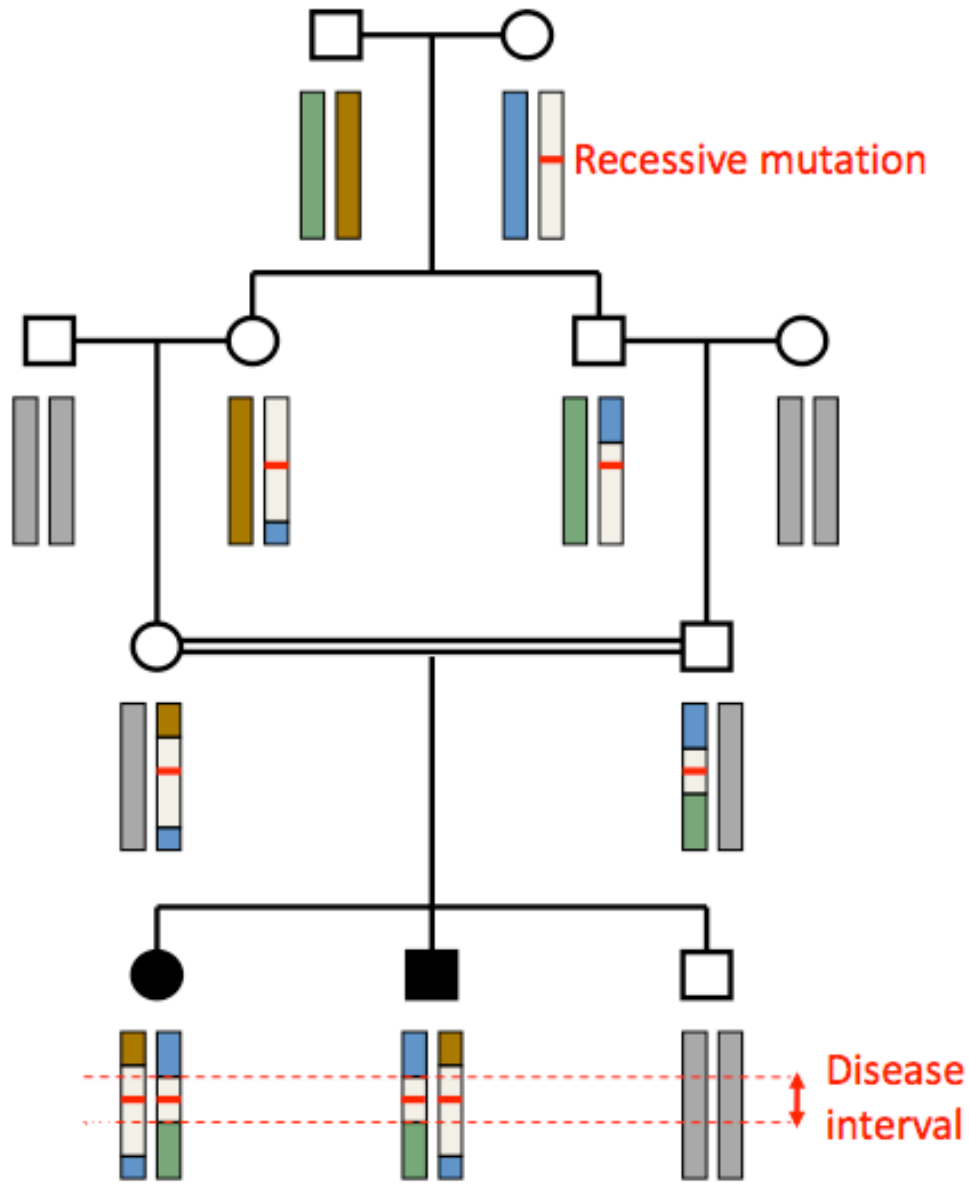

Recessive mutation
Disease
interval

Table 17.1 Human genetic relationships

Biological relationship	Genetic relationships	Coefficient of relationship	Coefficient of inbreeding
Incest[a]	First degree	0.5	0.25
Uncle-niece Double first cousin	Second degree	0.25	0.125
First cousin	Third degree	0.125	0.0625
First cousin once removed Double second cousin	Fourth degree	0.0625	0.0313
Second cousin	Fifth degree	0.0313	0.0156
Second cousin once removed Double third cousin	Sixth degree	0.0156	0.0078
Third cousin	Seventh degree	0.0078	0.0039

[a]Incest is defined as a sexual relationship between father–daughter, mother–son or brother–sister

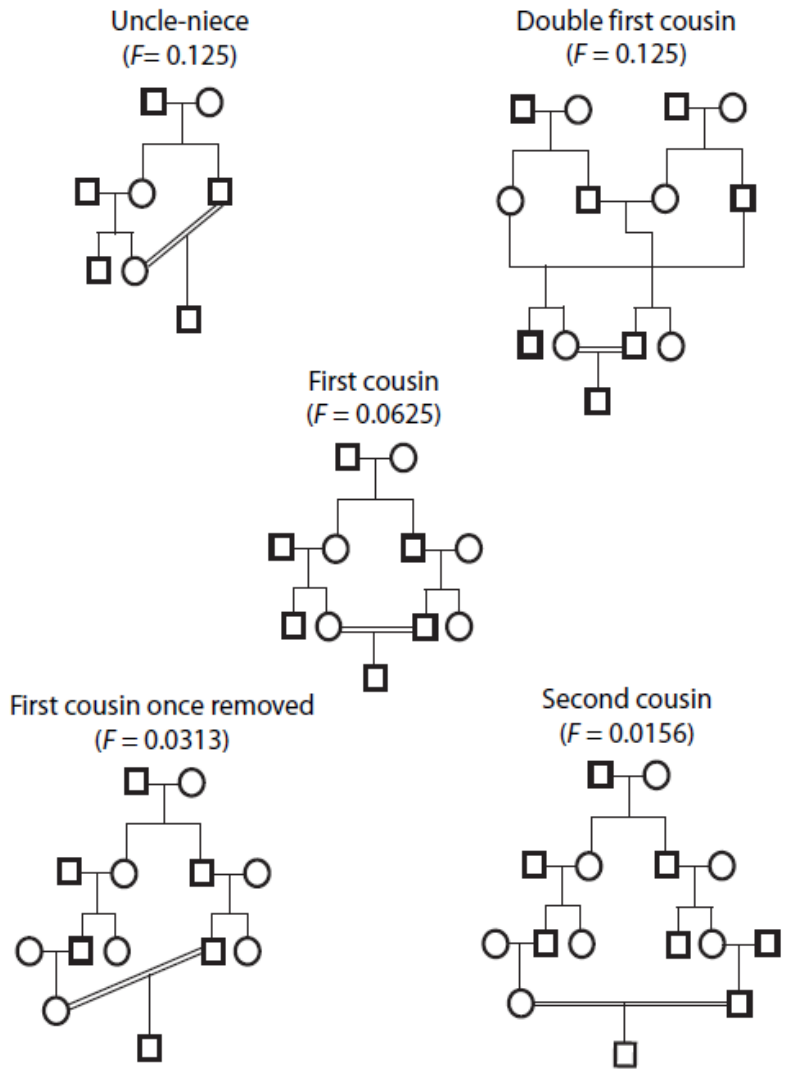
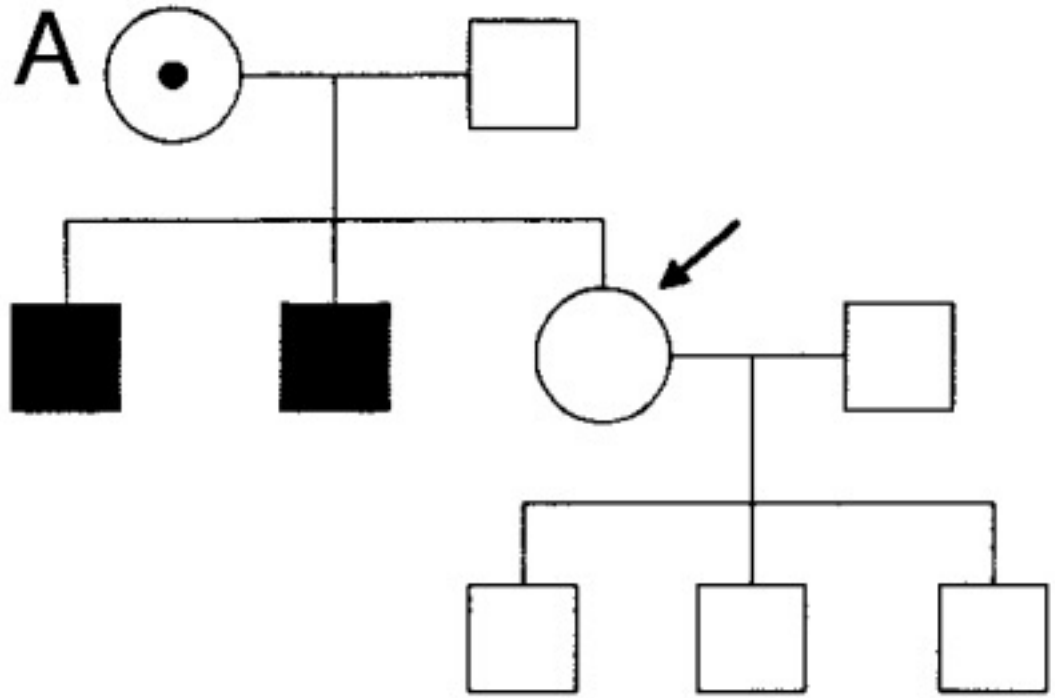


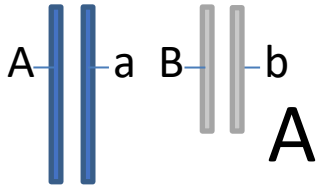
Fig. 17.1 Consanguineous pedigrees

A

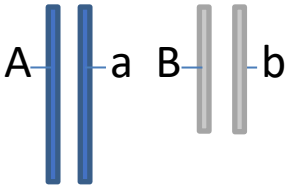


B

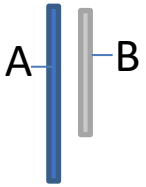
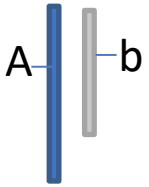
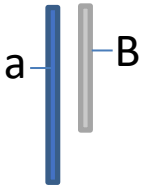
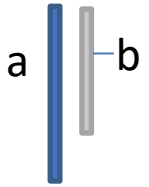
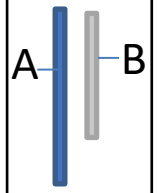
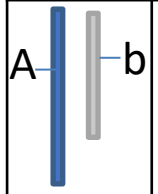
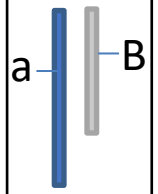
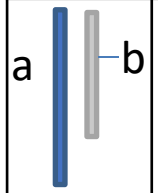
Hypothesis	Carrier	Non-carrier
Prior Probability	1/2	1/2
Conditional Probability (of three normal sons)	1/8	~1
Joint Probability	1/16	1/2
Posterior Probability	(1/16) / (1/16 + 1/2) = 1/9	(1/2) / (1/16 + 1/2) = 8/9



Aa Bb x Aa Bb



gametes

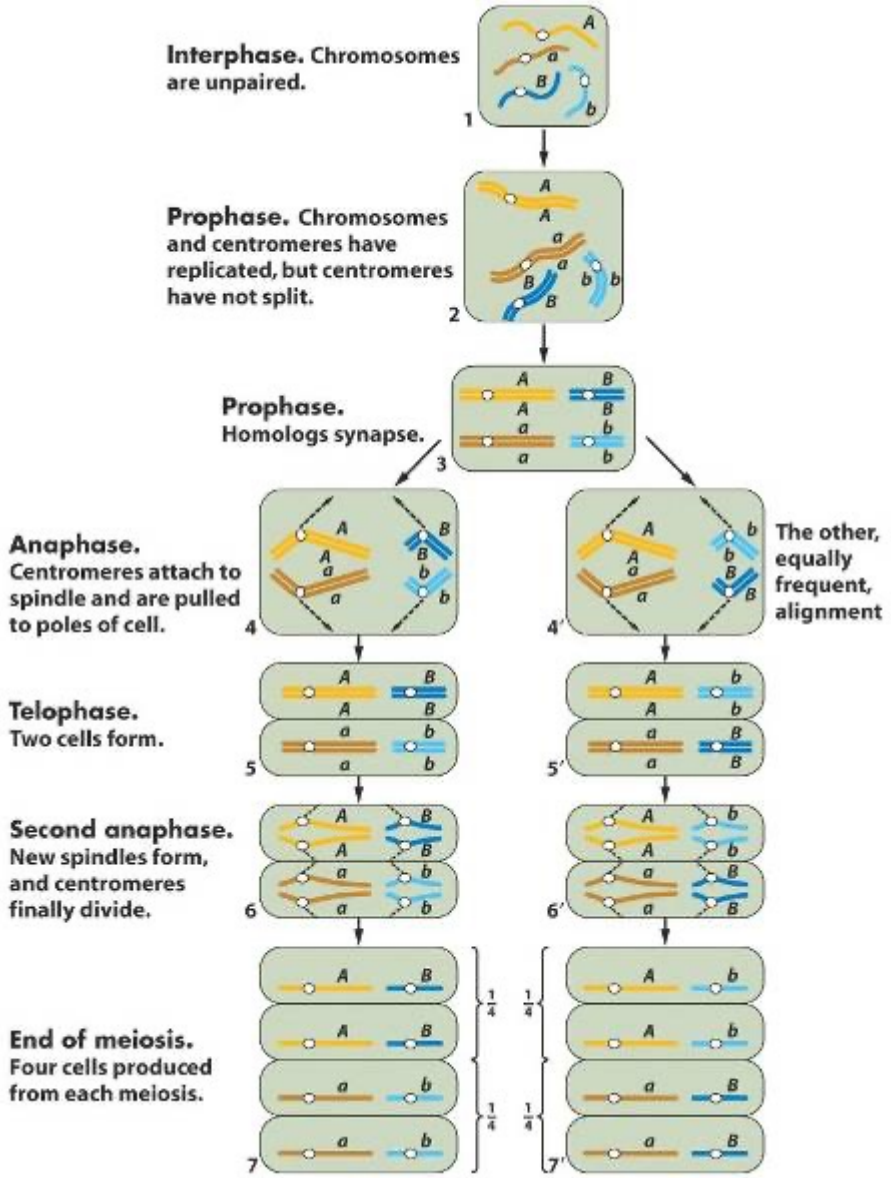
				
				
				
				
				

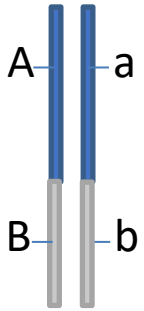
The Mendel's law of independent transmission is explained by the independent segregation of non homologous chromosomes during meiosis

Meosis in a diploid cell with genotype (A/a; B/b)

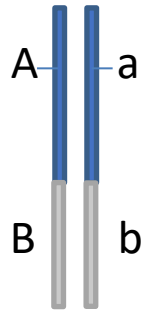
Genotipe AaBb

Gametes: AB ¼
Ab ¼
aB ¼
ab ¼

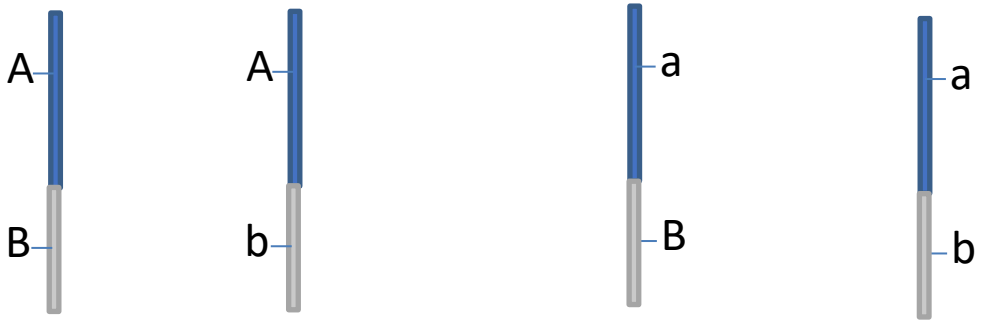




Aa Bb x Aa Bb



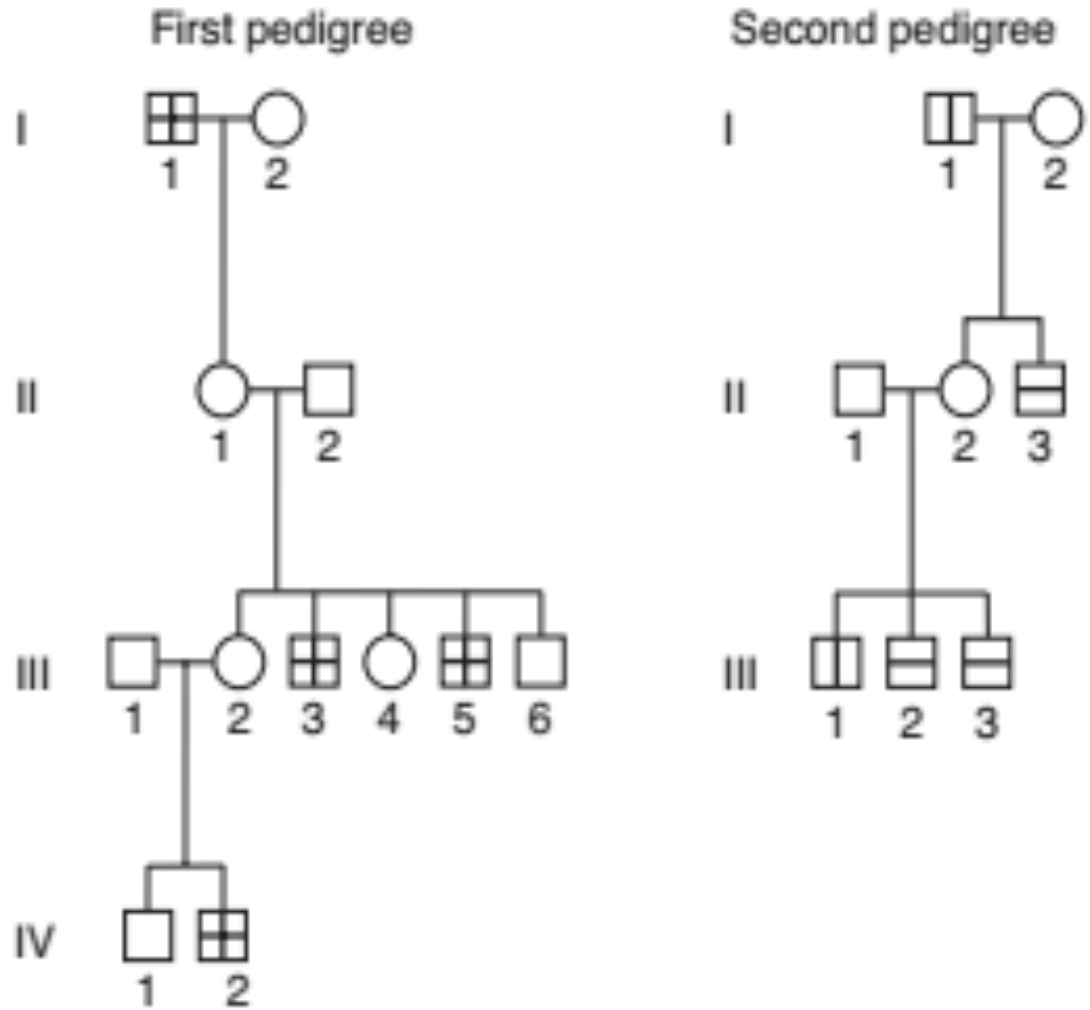
gametes

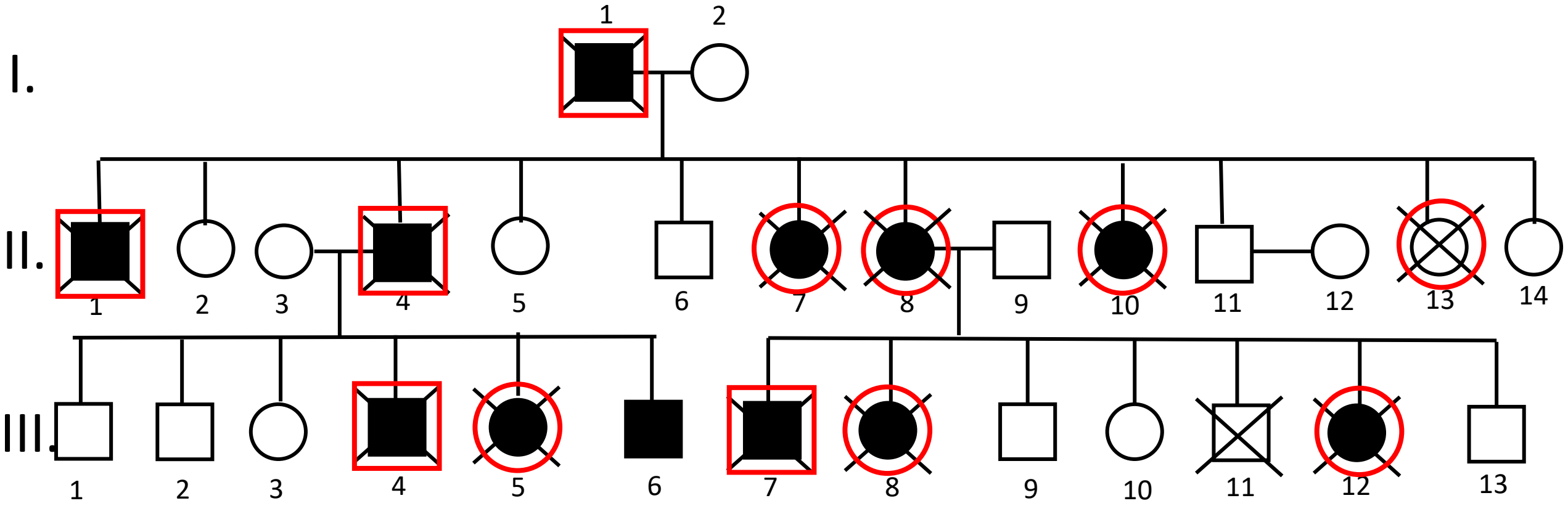


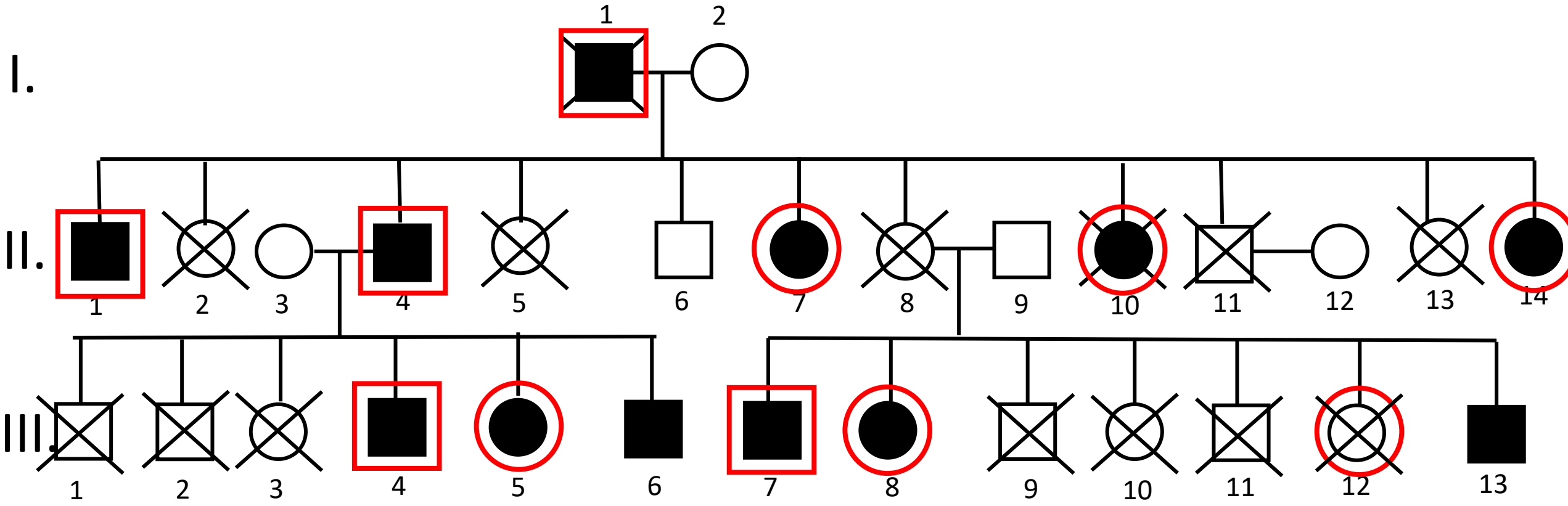
Do the observed proportions fit to the expected ones?

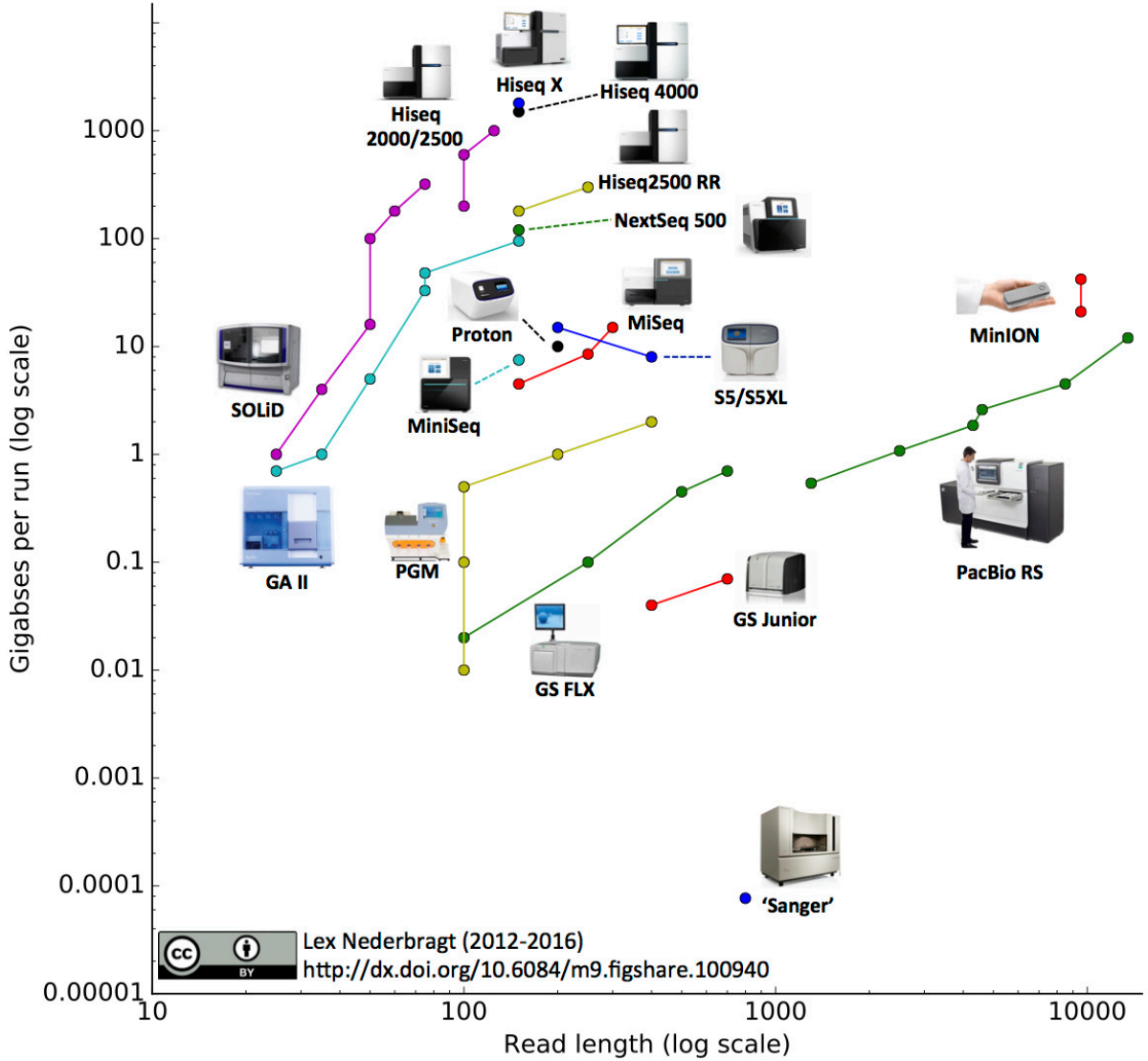
Do the observed proportions fit to the expected ones?

$$X^2 = \sum \frac{(\text{Observed value} - \text{Expected value})^2}{\text{Expected value}}$$


First pedigree
I
1
2
II
1
2
III
1
2
3
4
5
6
IV
1
2
Second pedigree
I
1
2
II
1
2
3
III
1
2
3


I.
1
2
II.
1
2
3
4
5
6
7
8
9
10
11
12
13
14
III.
1
2
3
4
5
6
7
8
9
10
11
12
13


I.
1
2
II.
1
2
3
4
5
6
7
8
9
10
11
12
13
14
III.
1
2
3
4
5
6
7
8
9
10
11
12
13



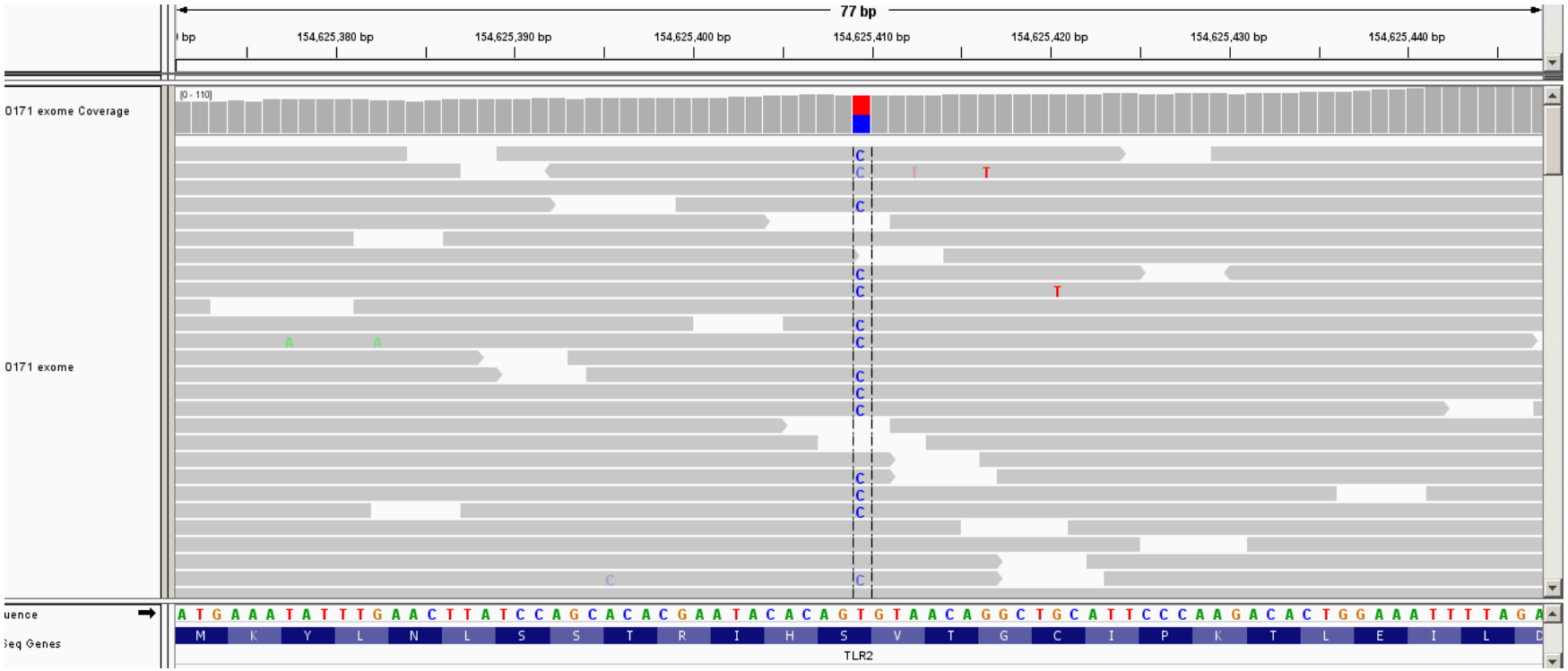
https://flxlexblog.wordpress.com/2016/07/08/developments-in-high-throughput-sequencing-july-2016-edition/

FASTQ files

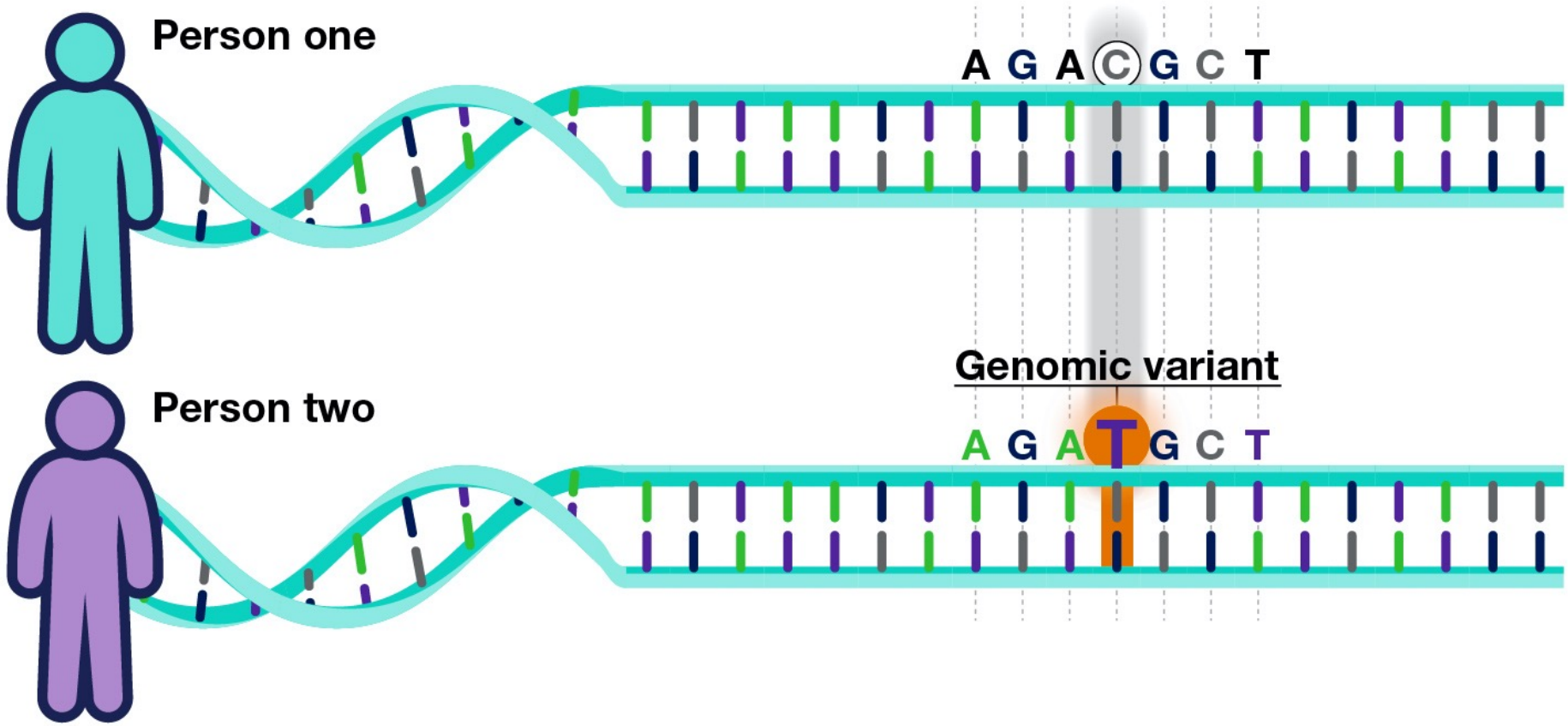
Line1: Sequence identifier
Line2: Raw sequence
Line3: meaningless
Line4: quality values for the sequence

```
@HWI-ST508:210:C0EDTACXX:1:1101:1872:1227 1:N:0:
AATTGTGAAAACCCAAAAGGTGGAGCAGCCATTNTTATACATTGCAGAAGGGNGANNNANCNTTATGAAATTTAGCACCTGCCTTCCTGAATGATAAATGG
+
@CCFFEFFHHHHHJJJJIJJCGHEIIIJIJJJJ#1BFHIJJJJJIJJIJJI#-;###-#-#-5?BFFFFEEEEEECCDDDDDDDDDDCCDDDDDDCCEED
@HWI-ST508:210:C0EDTACXX:1:1101:1895:1233 1:N:0:
TGACATAAGCTTGCATTTGAAAAGCACCTCCGAAAGCTTCCCAGCCTCAAAGNCANNATCGNCTTCTGATGCAGTTAGGCACCACAAGAGCTTCCCCACAA
+
CCCFFFFFHHHHGJJJJJIIJJIIIHJJIJJJJJJJJJJJJJJJJJJJGHJJ#.;##--;C#-5CEFFFFFEEECCEEDDDDDDDDDBDDDDDDDDDDDB
@HWI-ST508:210:C0EDTACXX:1:1101:1761:1235 1:N:0:
GCTCTACTAAAAATATAAAAATTGGCCAGGCGCAGTGACACATGCCTGTAGTCCCNGCTATTCGGGAGGCTGACACACAAGAATCAATCACTTGAACCCAG
+
CCCFFFFFHGHHHJJJJJJJIJJJJJJJJIEIIIJFHGIIIIJJIJJJJHIJJIJ#-;FGGIJIJHHFFDDEEDDCCDDDDCCDDDDDDCDDDDDDDDDDD
@HWI-ST508:210:C0EDTACXX:1:1101:1971:1236 1:N:0:
CAGGATGAAAGAGGTCTGGCCAGGTGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAAGTCAGGAGTT
+
CCCFFFFFHGHHGJHIJIIJJJJI3CFGIJJ9DFHJDEHGIJIJJJJJIIJJJGGIJJJJJJJJFIJHFFFFDDDB/?BB@BD<39?CD@B8+:@CDCB##
@HWI-ST508:210:C0EDTACXX:1:1101:1830:1239 1:N:0:
TATTGATTCTTTTTTTTTTTTTTTTTTTTTTTTTTCCTGGGGTCCTGCTTTGGGGGTCTCGGGGTCCCAAATTGCTGGTTTTACACCTCCCCCCCCCCG
+
?@@DD?DEFBBFDHEHIGEDA@FH>C??BBBCB6B##########################################################################
@HWI-ST508:210:C0EDTACXX:1:1101:1999:1240 1:N:0:
AAAGAGTGAGAGAAGCAAGGCTTGTGTGAAGAGAGCAAAACTTAGAATCAACATTGGTTGAGCATCTCCTATGAGCTAATATTAATTAGCACTTTACATGC
+
@@@DDA2?FHBHHEGEHIHGIGGHBFCGIEHGAEGGIIEGIIIIGHIGEHEGHIGIGBFHEHIEAHGHHFHEH;B@DEBDCDEEBCDDCCCCC@@CCCDCC
@HWI-ST508:210:C0EDTACXX:1:1101:1806:1245 1:N:0:
ACATGCTAATATATGTACTGATATGGAACAATCTTTAAGATGTATTATTACATGGAATAAACCAACCAGACCACAAAACAGATGTTTTTGCTTTTGCTAAA
+
CCCFFFFFHHHHGJJHJJJJIJIJJJJJJJJJJJJJJJJJJIJJJJJJJJJJJJJJJJGIJJJJJIJJJHJIJGIHHHGFFFEFFEDDDDDDDDDDEDD
```

Visualization



#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	BGMUB-1	BGMUB-2	BGMUB-3	BGMUB-C1	BGM
13	20187061	rs9509085	C	T	100	PASS	AA=C|||;AC=	GT	1|1	1|1	1|0	1|1	1|1
13	20187624	rs7988691	A	G	100	PASS	AA=G|||;AC	GT	1|1	1|1	1|1	1|1	1|1
13	20187749	rs7623	C	T	100	PASS	AA=T|||;AC=	GT	1|1	1|0	1|1	1|1	1|1
13	20187970	rs5030700	G	A	100	PASS	AA=G|||;AC	GT	1|0	0|0	0|0	0|0	0|0
13	20188733	rs55704559	T	C	100	PASS	AA=T|||;AC=	GT	1|0	0|0	0|0	0|0	0|0
13	20188817	rs3751385	A	G	100	PASS	AA=G|||;AC	GT	1|0	1|1	1|0	1|1	1|1
13	20188940	rs530484784	T	C	100	PASS	AA=T|||;AC=	GT	0|0	1|0	0|0	0|0	0|0
13	20189214	rs111033188	G	T	100	PASS	AA=G|||;AC	GT	1|0	0|0	0|0	0|0	0|0
13	20189241	rs2274083	T	C	100	PASS	AA=T|||;AC=	GT	0|0	0|0	0|0	1|0	0|0
13	20189324	rs139362103	C	T	100	PASS	AA=C|||;AC=	GT	0|0	0|0	0|0	0|0	1|0
13	20189510	.	C	T	100	PASS	AA=T|||;AC=	GT	0|0	1|0	0|0	0|0	0|0
13	20189548	.	CC	C	100	PASS	AA=T|||;AC=	GT	0|0	1|0	0|0	0|0	0|0
13	20189596	rs72561725	G	A	100	PASS	AA=G|||;AC	GT	0|0	0|0	1|0	0|0	0|0
13	20189783	rs73156818	A	T	100	PASS	AA=A|||;AC	GT	0|0	0|0	0|0	1|0	0|0
13	20189935	rs7318163	G	T	100	PASS	AA=T|||;AC=	GT	1|1	1|1	1|0	1|1	1|1
13	20223243	.	G	C	100	PASS	AA=G|||;AC	GT	1|1	0|0	0|0	0|0	0|0
13	20223280	rs554688212	G	T	100	PASS	AA=G|||;AC	GT	0|0	0|0	1|0	0|0	0|0
13	20223451	rs377181573	G	A	100	PASS	AA=G|||;AC	GT	0|0	0|0	1|0	0|0	0|0
13	20223475	rs200415730	A	G	100	PASS	AA=A|||;AC	GT	0|0	1|0	0|0	0|0	0|0
13	20229254	rs145808643	T	C	100	PASS	AA=T|||;AC=	GT	0|0	1|0	0|0	0|0	1|0
13	20229305	rs71424089	T	C	100	PASS	AA=C|||;AC=	GT	0|0	0|0	1|0	0|0	0|0
13	20232512	rs148043721	C	CG	100	PASS	AA=?|GGGG	GT	1|1	1|1	1|1	1|1	1|1
13	20232627	rs9506446	G	A	100	PASS	AA=G|||;AC	GT	0|0	1|1	1|0	1|0	1|1


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 THE GENOME HAVE?

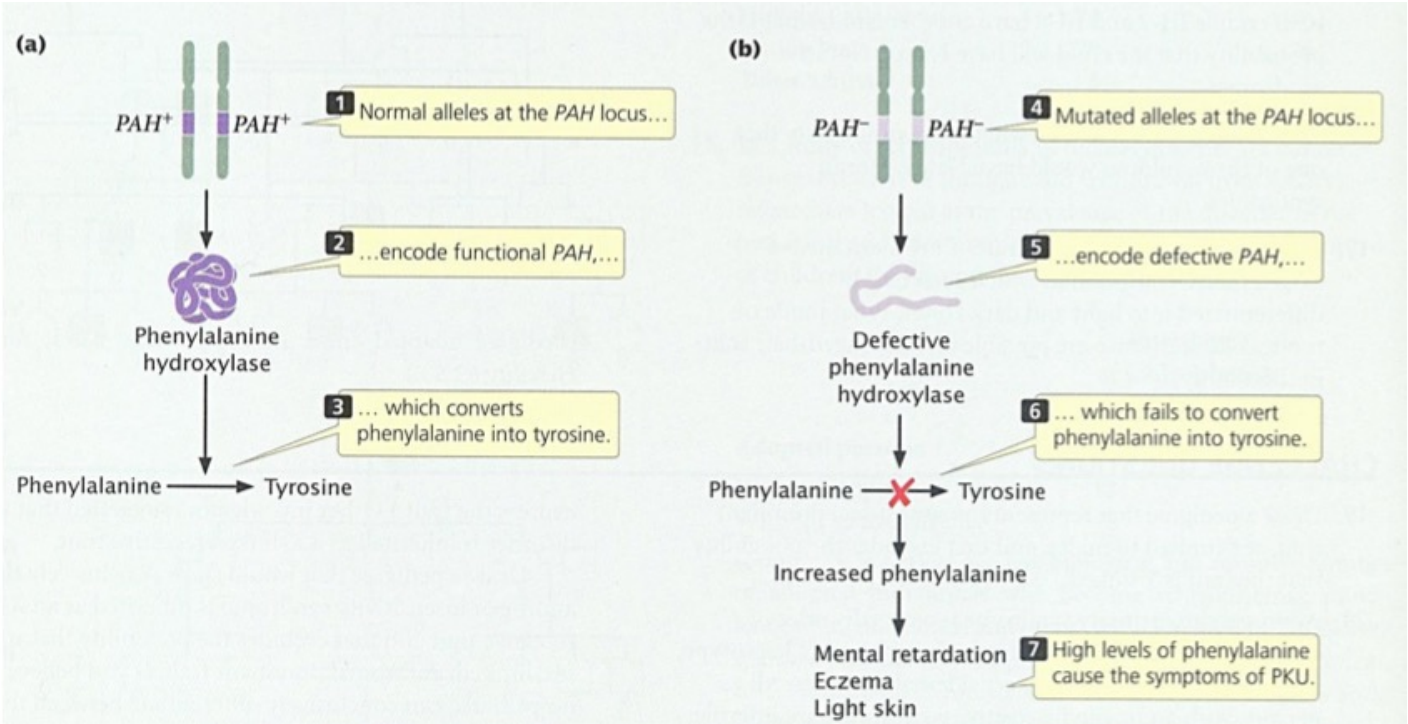
2. WHICH TYPE?

3. WHAT EFFECT DO THEY HAVE?

4. WHAT IS THE BIOLOGICAL SIGNIFICANCE?

Table 2 Frequency of different types of mutations that result in PKU

Type of PKU mutation	Percentage of all PKU mutations
Missense	67%
Deletion	14%
Splice	12%
Nonsense	6%
Insertion	1%



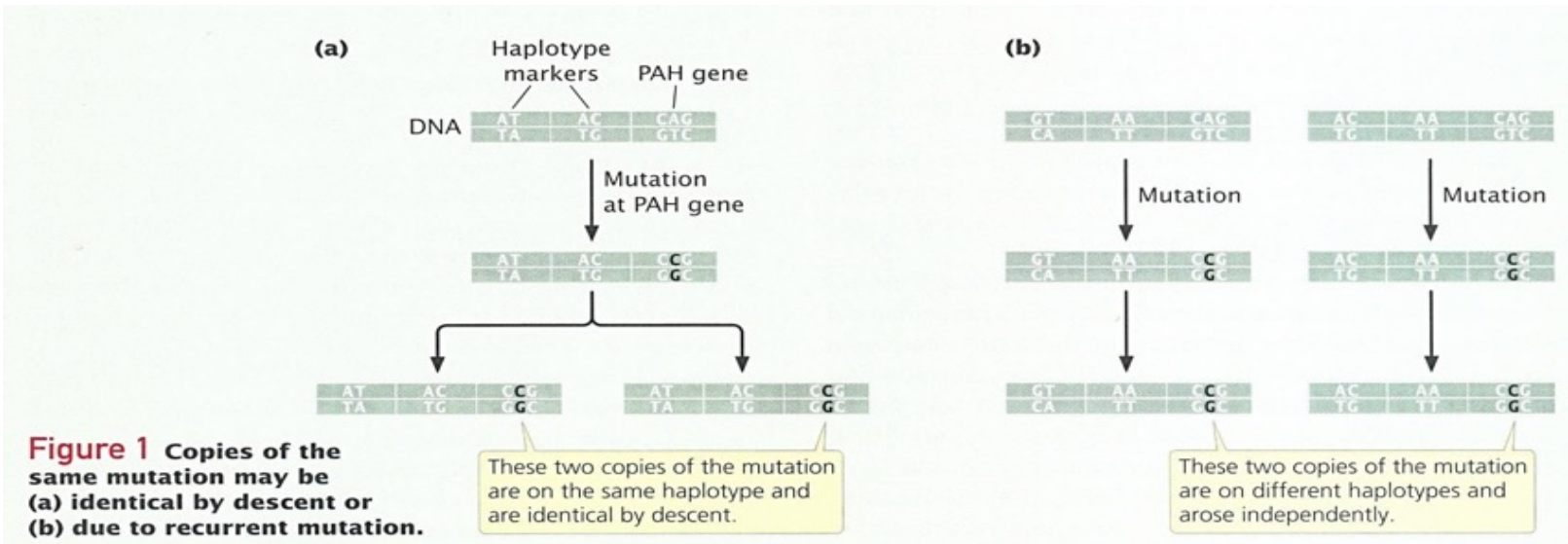
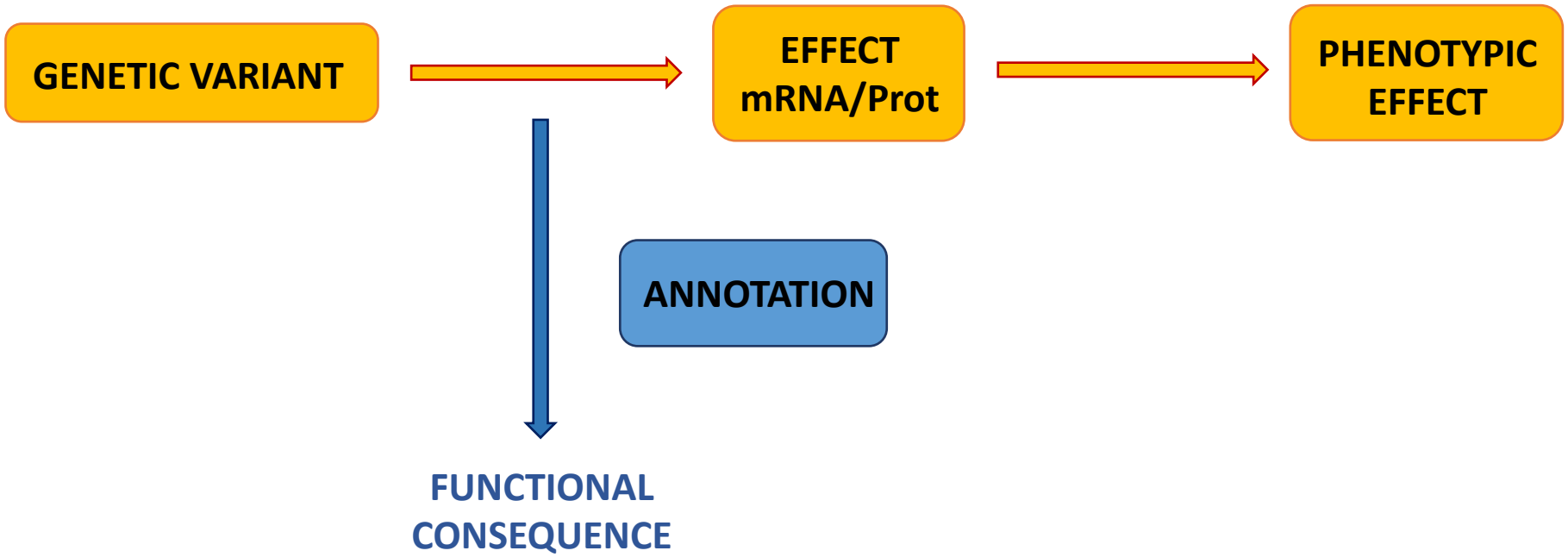


Figure 1 **Copies of the same mutation may be (a) identical by descent or (b) due to recurrent mutation.**

3. WHAT EFFECT DO THEY HAVE?

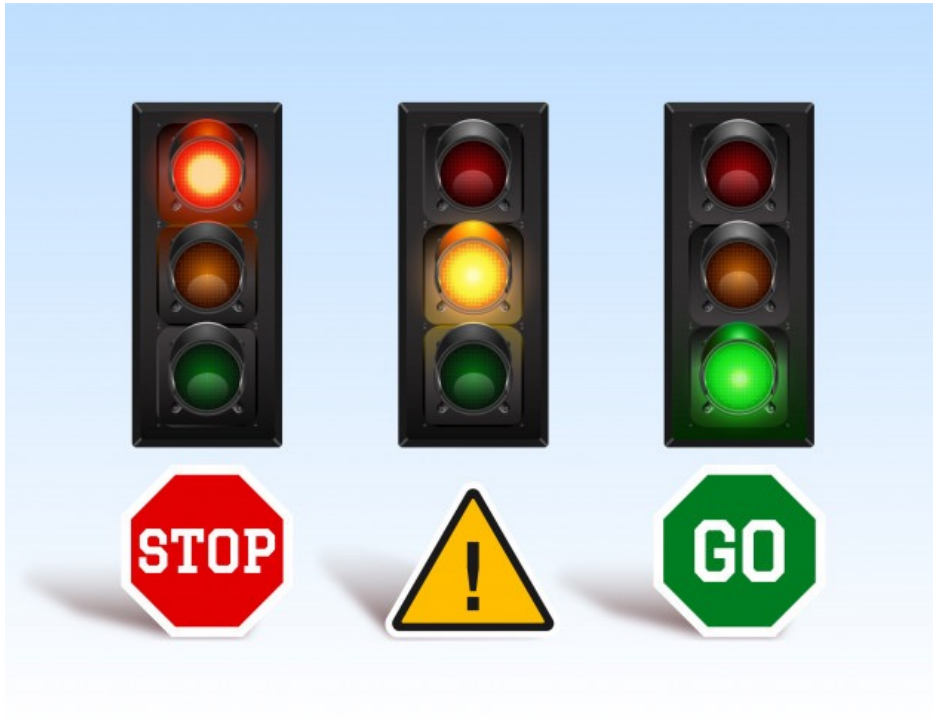
Functional annotation



3. EFFECT OF GENETIC VARIANTS

Functional annotation

- synonymous
- splice donor
- missense
- frameshift
- intron
- stop gained



Functional annotation

Computational approaches to predict the impact of SNVs

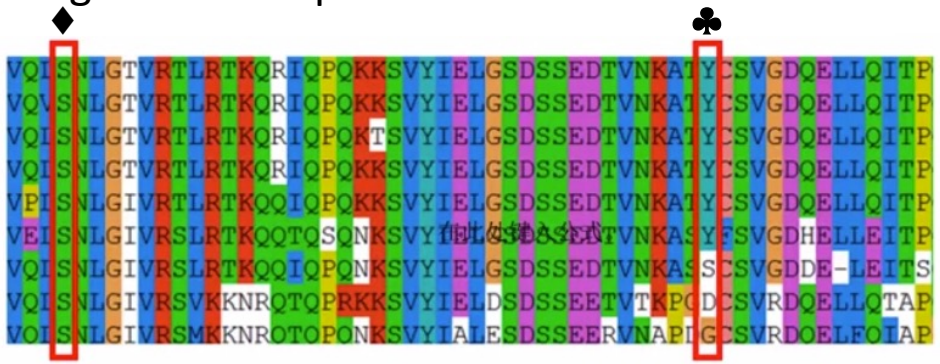
Sorting **I**ntolerant **F**rom **T**olerant substitutions (**SIFT**)

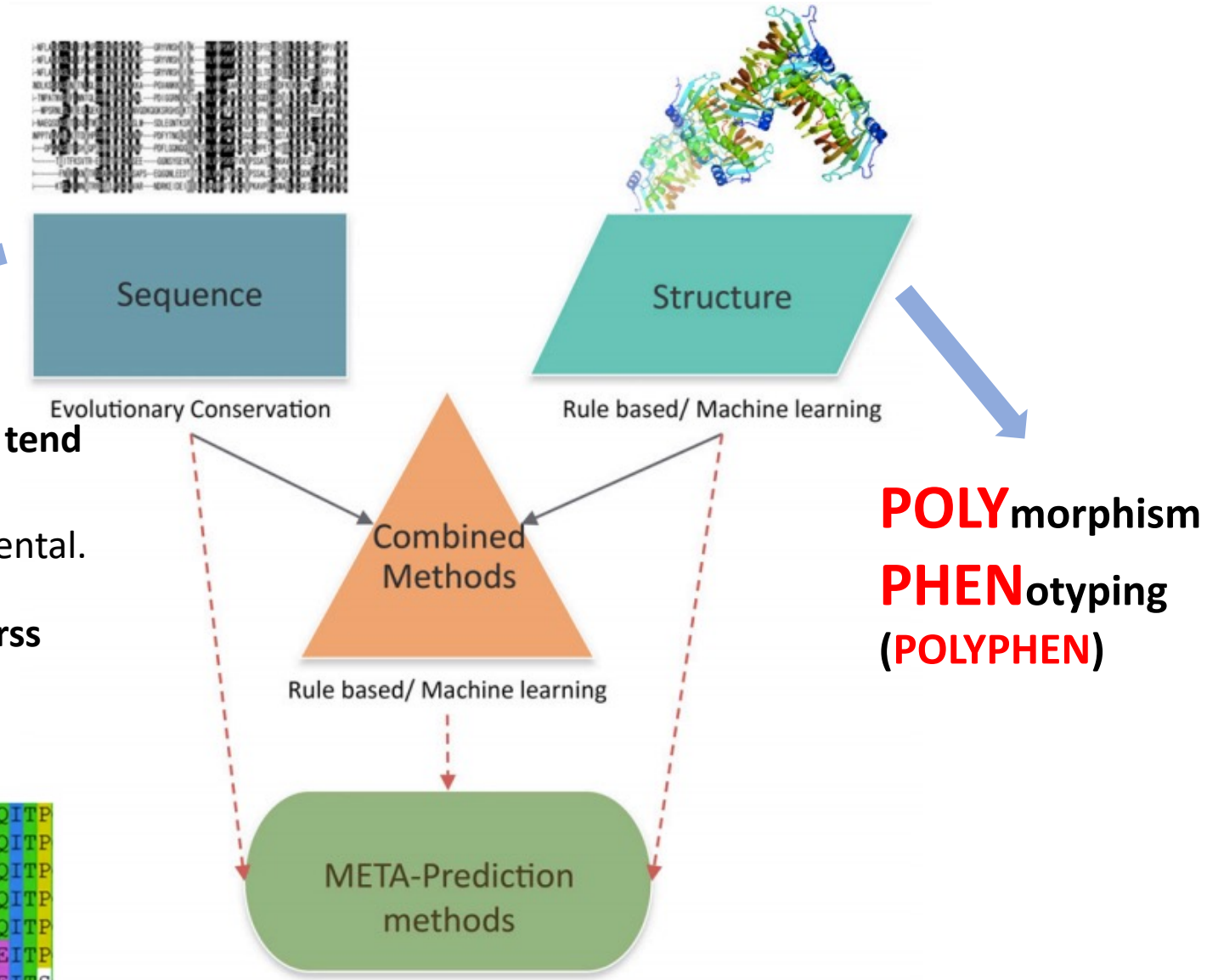
➢ **Funtionally important positions (e.g. active sites) tend to be conserved across species.**

- ♦ Changes at conserved sites tend to be detrimental.

➢ **Some positions show a high level of diversity acorss species.**

- ♣ Changes at these positions tend to be neutral.



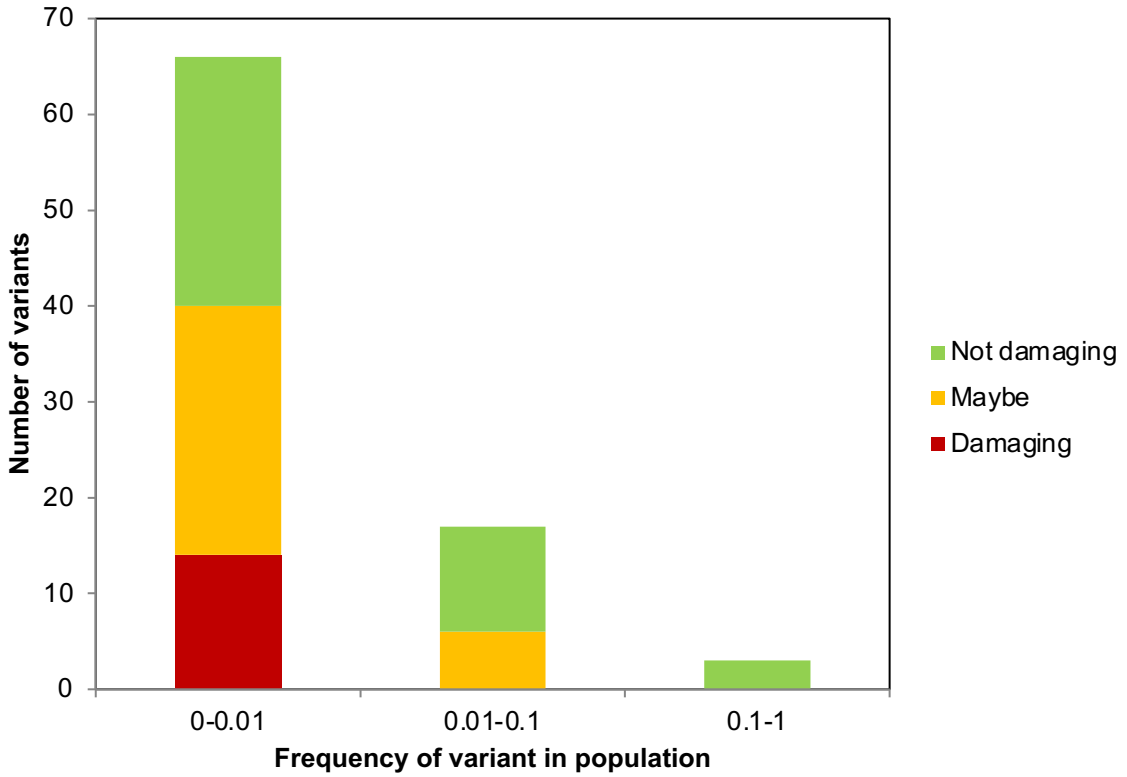


POLYmorphism
PHENotyping
(**POLYPHEN**)

3. EFFECT OF GENETIC VARIANTS

Functional annotation

Genetic variant	Position	Change	Consequence	Gene	Freq
A	3:1403401-1403401	A/C	missense	CNTN6	0.003
B	12:49022301-49022301	G/A	missense	KMT2D	-
C	3:25634002-25634002	G/A	missense	TOP2B	4×10^{-6}
D	15:72353105-72353105	C/T	missense	HEXA	5×10^{-5}
E	4:38797314-38797314	C/A	synonymous	TLR1	0.38
F	7:117639961-117639961	C/T	intron	CFTR	6×10^{-5}
G	17:31230383-31230383	G/A	splice_donor	NF1	4×10^{-6}
H	11:63290453-63290453	G/A	stop_gained	SLC22A10	0.434



Marth *et al*. (2011) *Genome Biol*. **12**, R84

3. EFFECT OF GENETIC VARIANTS

Functional annotation

Genetic variant	Position	Change	Consequence	Gene	Freq	SIFT	Polyphen	Mutation Assessor	CADD	Condel	LoF tol	RVIS
A	3:1403401-1403401	A/C	missense	CNTN6	0.003	0.47	0	Neutral	9.52	Neutral	0.974	-0.92 (9.81%)
B	12:49022301-49022301	G/A	missense	KMT2D	-	0	1	Medium	32	0.945	-	-5.29 (0.06%)
C	3:25634002-25634002	G/A	missense	TOP2B	4×10^{-6}	0.01	0.152	Medium	29.3	0.778	0.82	-0.15 (42.28%)
D	15:72353105-72353105	C/T	missense	HEXA	5×10^{-5}	0	1	Damaging	32	0.945	0.17	-0.33 (30.7%)
E	4:38797314-38797314	C/A	synonymous	TLR1	0.38	-	-	-	0.29	-	0.963	1.32 (94.07%)
F	7:117639961-117639961	C/T	intron	CFTR	6×10^{-5}	-	-	-	1.08	-	0.0235	-0.51 (21.73%)
G	17:31230383-31230383	G/A	splice_donor	NF1	4×10^{-6}	-	-	-	34	-	0.116	-3.09 (0.47%)
H	11:63290453-63290453	G/A	stop_gained	SLC22A10	0.434	-	-	-	36	-	0.7	1.85 (97.12%)

3. EFFECT OF GENETIC VARIANTS

Functional annotation

- G is the more pathogenic variant…….but is it really pathogenic?

3. EFFECT OF GENETIC VARIANTS

- ➢ Pathogenic
- ➢ Likely pathogenic
- ➢ Benign
- ➢ Likely benign
- ➢ Uncertain significance

Table 5 Rules for combining criteria to classify sequence variants

Pathogenic	(i) 1 Very strong (PVS1) *AND* (a) ≥1 Strong (PS1–PS4) *OR* (b) ≥2 Moderate (PM1–PM6) *OR* (c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) *OR* (d) ≥2 Supporting (PP1–PP5) (ii) ≥2 Strong (PS1–PS4) *OR* (iii) 1 Strong (PS1–PS4) *AND* (a)≥3 Moderate (PM1–PM6) *OR* (b)2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR* (c)1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Likely pathogenic	(i) 1 Very strong (PVS1) *AND* 1 moderate (PM1–PM6) *OR* (ii) 1 Strong (PS1–PS4) *AND* 1–2 moderate (PM1–PM6) *OR* (iii) 1 Strong (PS1–PS4) *AND* ≥2 supporting (PP1–PP5) *OR* (iv) ≥3 Moderate (PM1–PM6) *OR* (v) 2 Moderate (PM1–PM6) *AND* ≥2 supporting (PP1–PP5) *OR* (vi) 1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Benign	(i) 1 Stand-alone (BA1) *OR* (ii) ≥2 Strong (BS1–BS4)
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) *OR* (ii) ≥2 Supporting (BP1–BP7)
Uncertain significance	(i) Other criteria shown above are not met OR (ii) the criteria for benign and pathogenic are contradictory

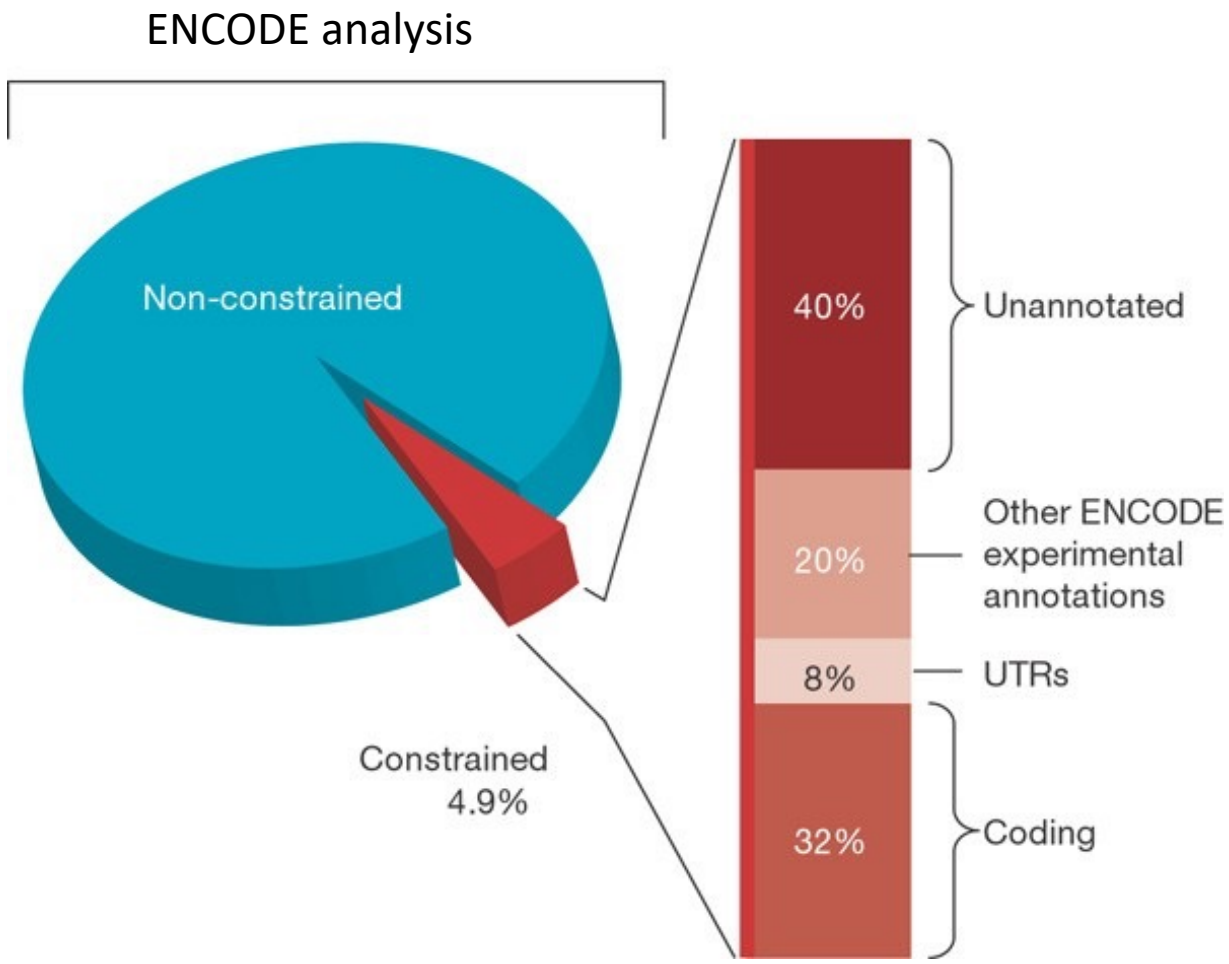
Functional interpretation

ACMG STANDARDS AND GUIDELINES

	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very strong
Population data	MAF is too high for disorder BA1/BS1 OR observation in controls inconsistent with disease penetrance BS2			Absent in population databases PM2	Prevalence in affecteds statistically increased over controls PS4	
Computational and predictive data		Multiple lines of computational evidence suggest no impact on gene /gene product BP4 Missense in gene where only truncating cause disease BP1 Silent variant with non predicted splice impact BP7 In-frame indels in repeat w/out known function BP3	Multiple lines of computational evidence support a deleterious effect on the gene /gene product PP3	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before PM5 Protein length changing variant PM4	Same amino acid change as an established pathogenic variant PS1	Predicted null variant in a gene where LOF is a known mechanism of disease PVS1
Functional data	Well-established functional studies show no deleterious effect BS3		Missense in gene with low rate of benign missense variants and path. missenses common PP2	Mutational hot spot or well-studied functional domain without benign variation PM1	Well-established functional studies show a deleterious effect PS3	
Segregation data	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and maternity confirmed) PS2	
Allelic data		Observed in *trans* with a dominant variant BP2 Observed in *cis* with a pathogenic variant BP2		For recessive disorders, detected in trans with a pathogenic variant PM3		
Other database		Reputable source w/out shared data – benign BP6	Reputable source – pathogenic PP5			
Other data		Found in case with an alternate cause BP5	Patient's phenotype or FH highly specific for gene PP4			

Figure 1 Evidence framework. This chart organizes each of the criteria by the type of evidence as well as the strength of the criteria for a benign (left side) or pathogenic (right side) assertion. Evidence code descriptions can be found in **Tables 3** and **4**. BS, benign strong; BP, benign supporting; FH, family history; LOF, loss of function; MAF, minor allele frequency; path., pathogenic; PM, pathogenic moderate; PP, pathogenic supporting; PS, pathogenic strong; PVS, pathogenic very strong.

4. BIOLOGICAL MEANING OF GENETIC VARIANTS

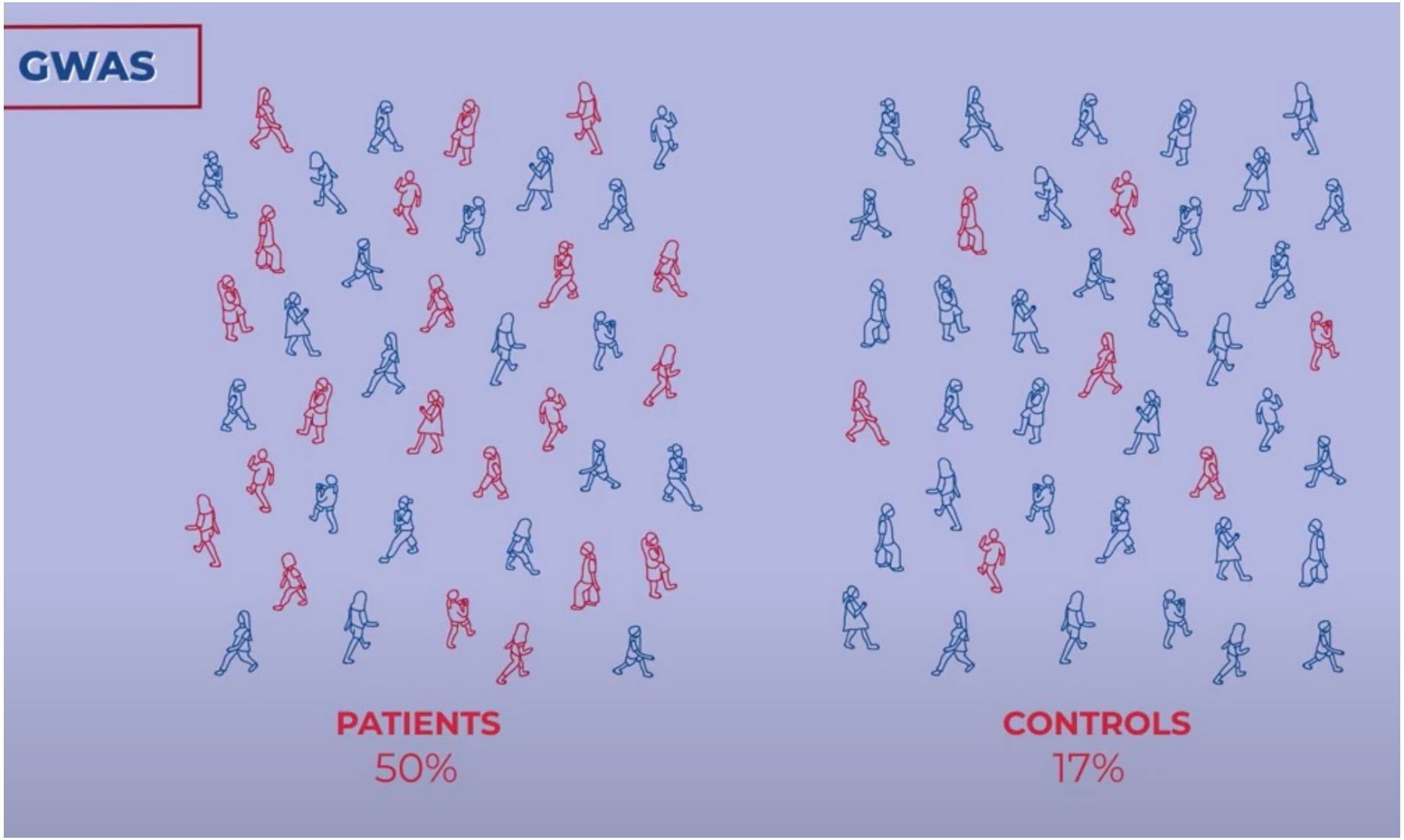


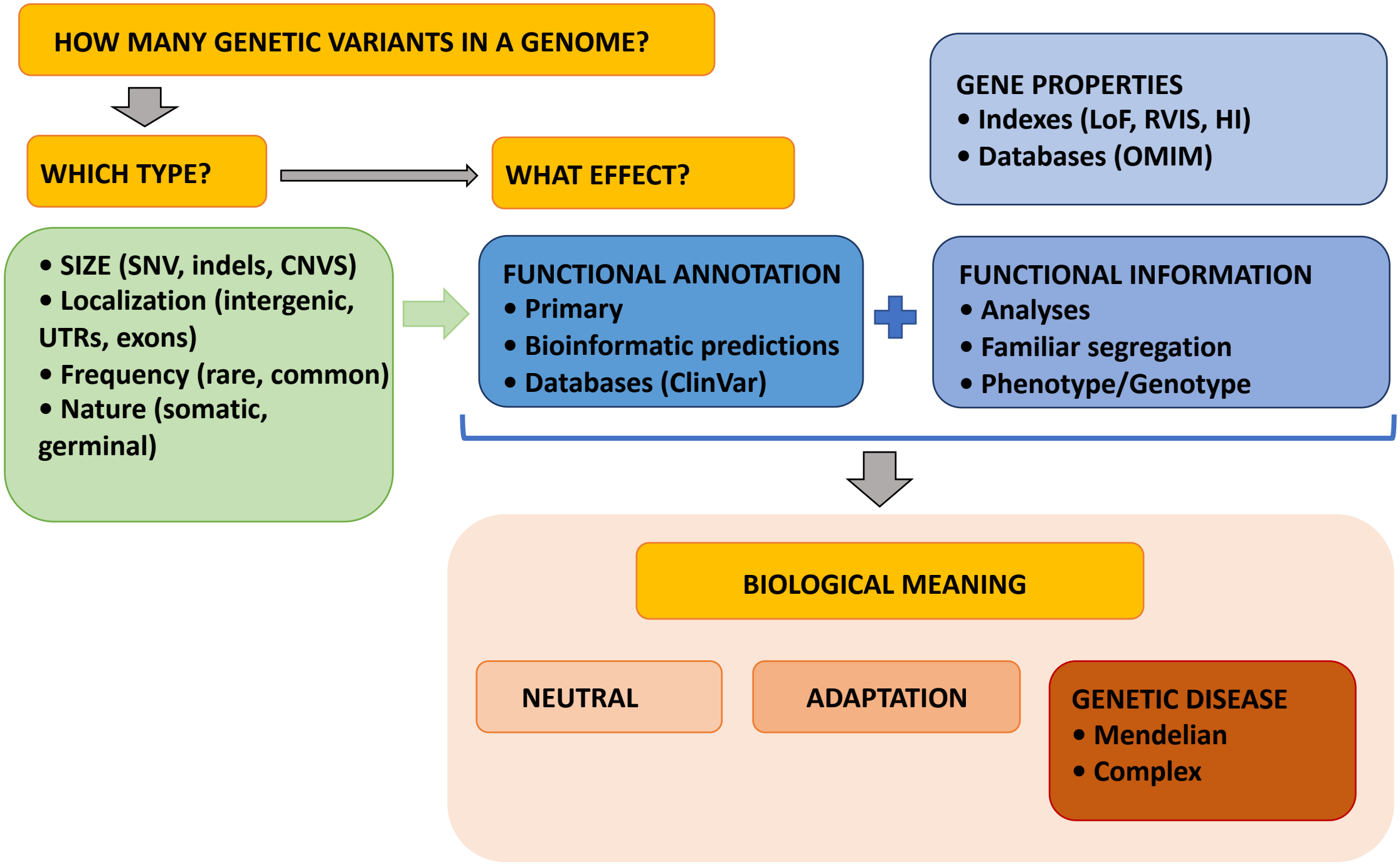
GENETIC DISEASE
- **Mendelian**
- **Complex**

ADAPTATION

NEUTRAL

GENOME WIDE ASSOCIATION STUDIES



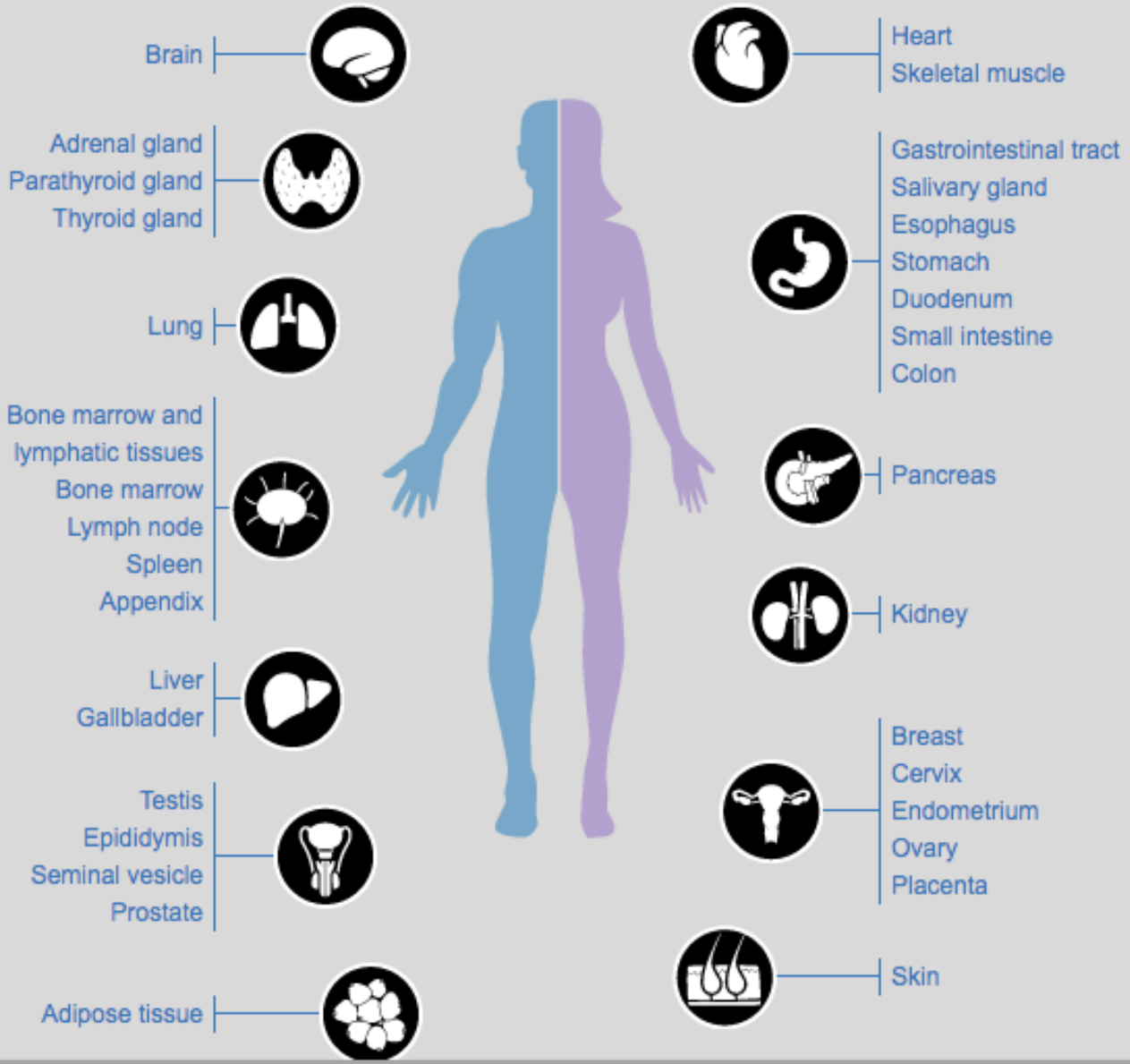

HOW MANY GENETIC VARIANTS IN A GENOME?
WHICH TYPE?
WHAT EFFECT?
GENE PROPERTIES
• Indexes (LoF, RVIS, HI)
• Databases (OMIM)
• SIZE (SNV, indels, CNVS)
• Localization (intergenic, UTRs, exons)
• Frequency (rare, common)
• Nature (somatic, germinal)
FUNCTIONAL ANNOTATION
• Primary
• Bioinformatic predictions
• Databases (ClinVar)
FUNCTIONAL INFORMATION
• Analyses
• Familiar segregation
• Phenotype/Genotype
BIOLOGICAL MEANING
NEUTRAL
ADAPTATION
GENETIC DISEASE
• Mendelian
• Complex

THE HUMAN PROTEIN ATLAS

THE TISSUE AND ORGAN PROTEOMES

Explore the proteomes of specific tissues and organs

The expression for all protein-coding genes in all major tissues and organs in the human body can be explored in this interactive database, including numerous catalogues of proteins expressed in a tissue-restricted manner.



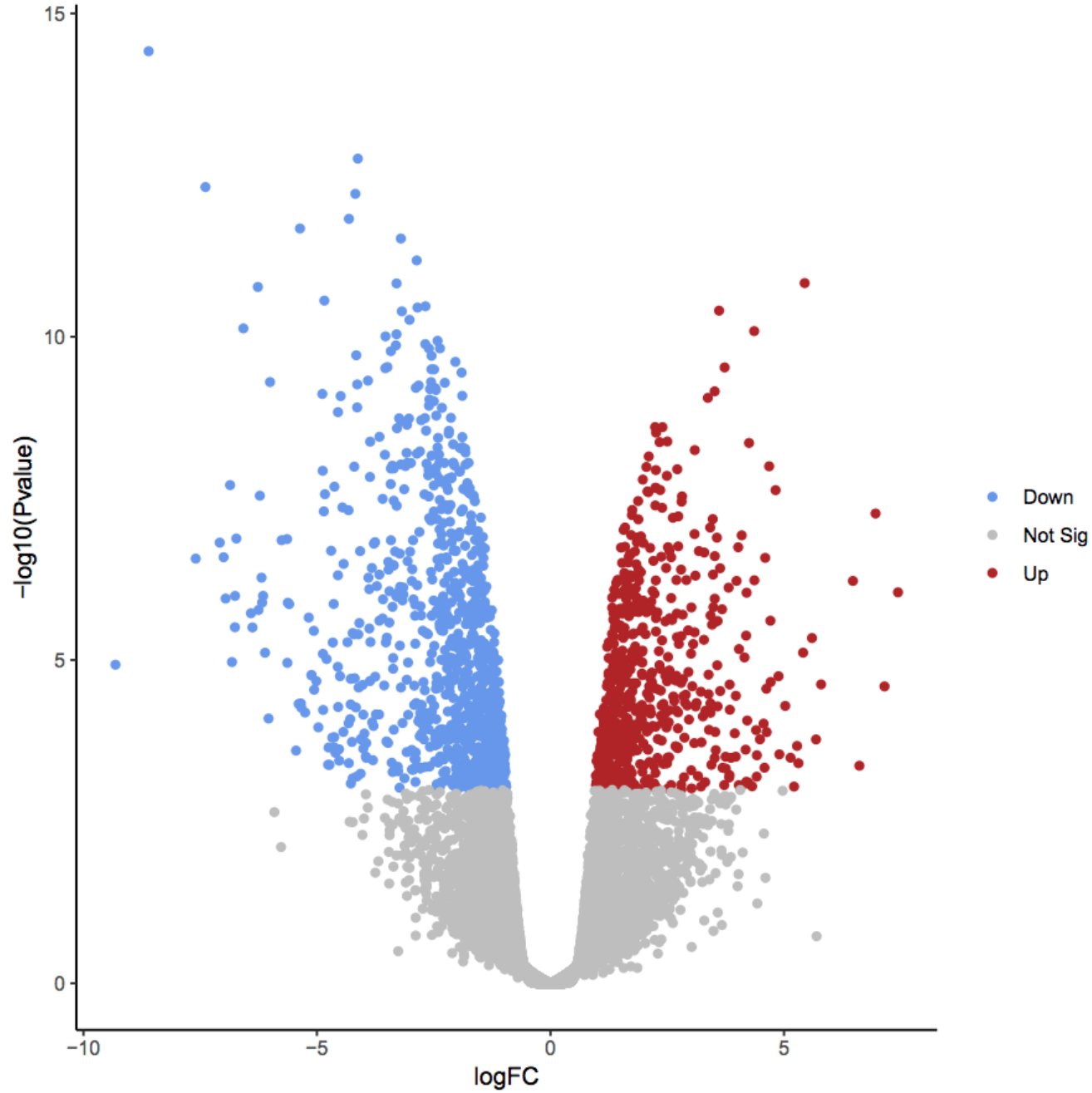
Differential Expression Analysis

1. Mapping & transcript reconstruction

2. Abundance estimation

3. Differential expression (gene level)

4. Enrichment analysis


15
10
5
0
−log10(Pvalue)
−10
−5
0
5
logFC
Down
Not Sig
Up

::: How works GSEA?