

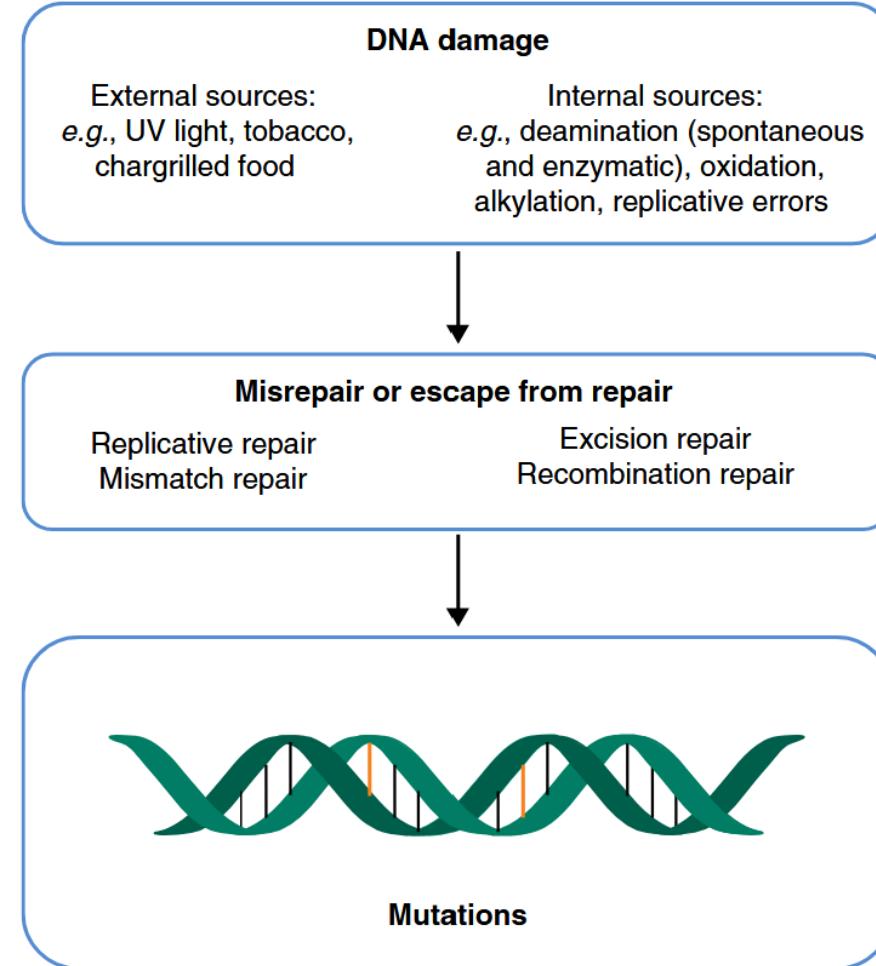
# **Module 3: Human genetic variation**

## Key concepts

- Origins of sequence variation
- Mechanisms of pathogenesis
- Pathogenic variants
- Protein polymorphism
- Variable expression of polygenic diseases

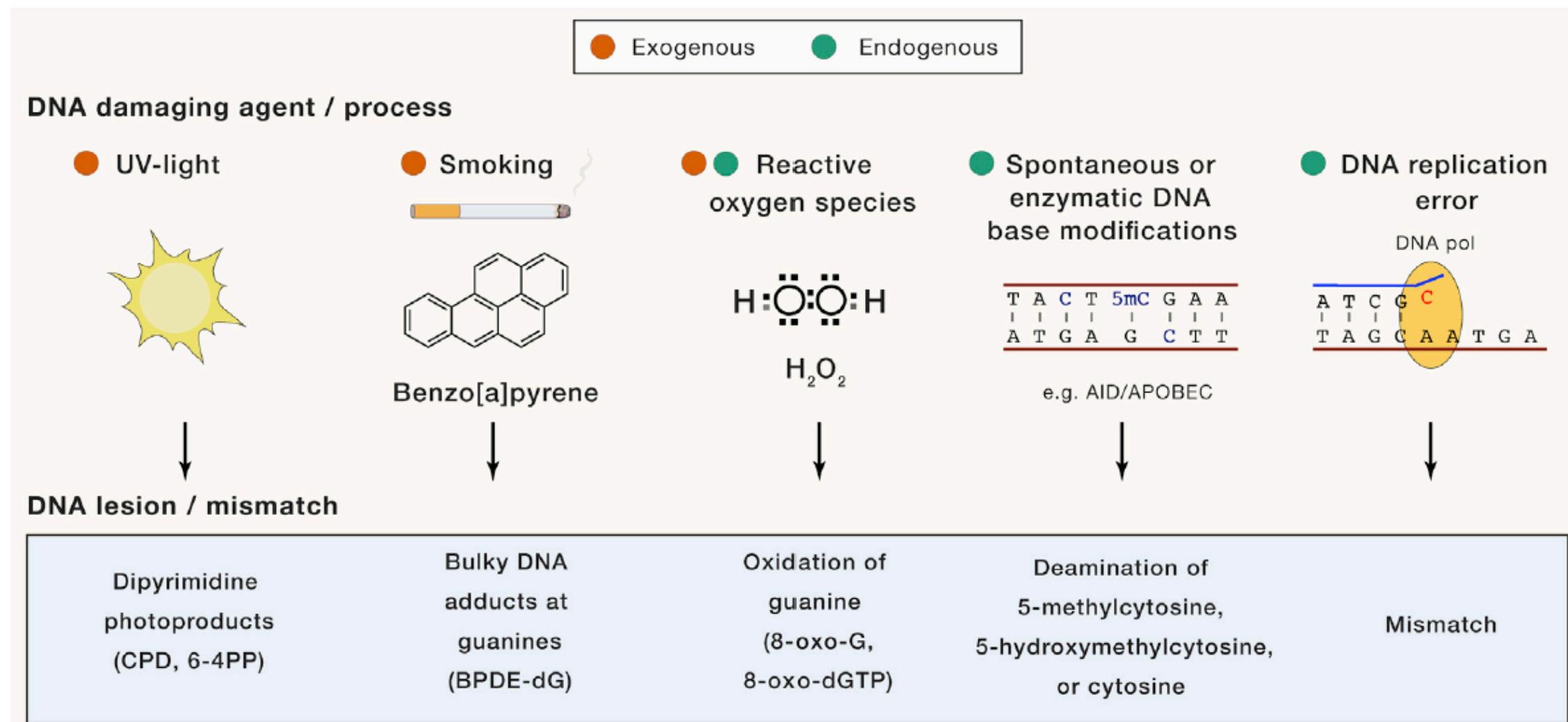
# Origins of sequence variation <sup>1</sup>

1. Chemical damage to DNA
2. DNA replication/repair errors
3. Chromosome segregation and recombination errors



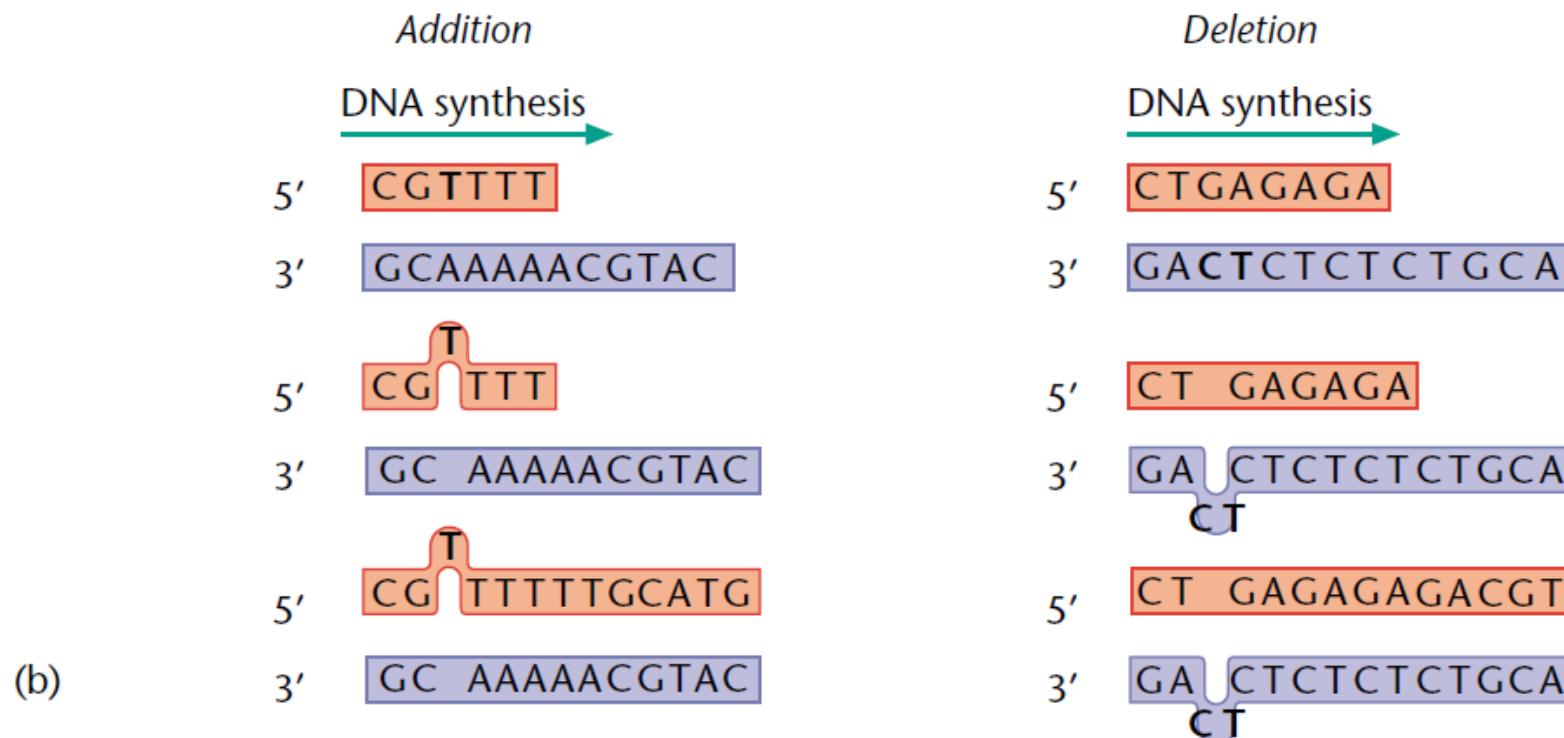
**Figure 1 External and internal sources of mutation in cancer.** A schematic depiction of major external and internal sources of DNA damage, a variety of DNA repair mechanisms that serve to counteract damage, and mutation as an outcome of unrepaired DNA damage.

# 1. Chemical damage to DNA <sup>2</sup>

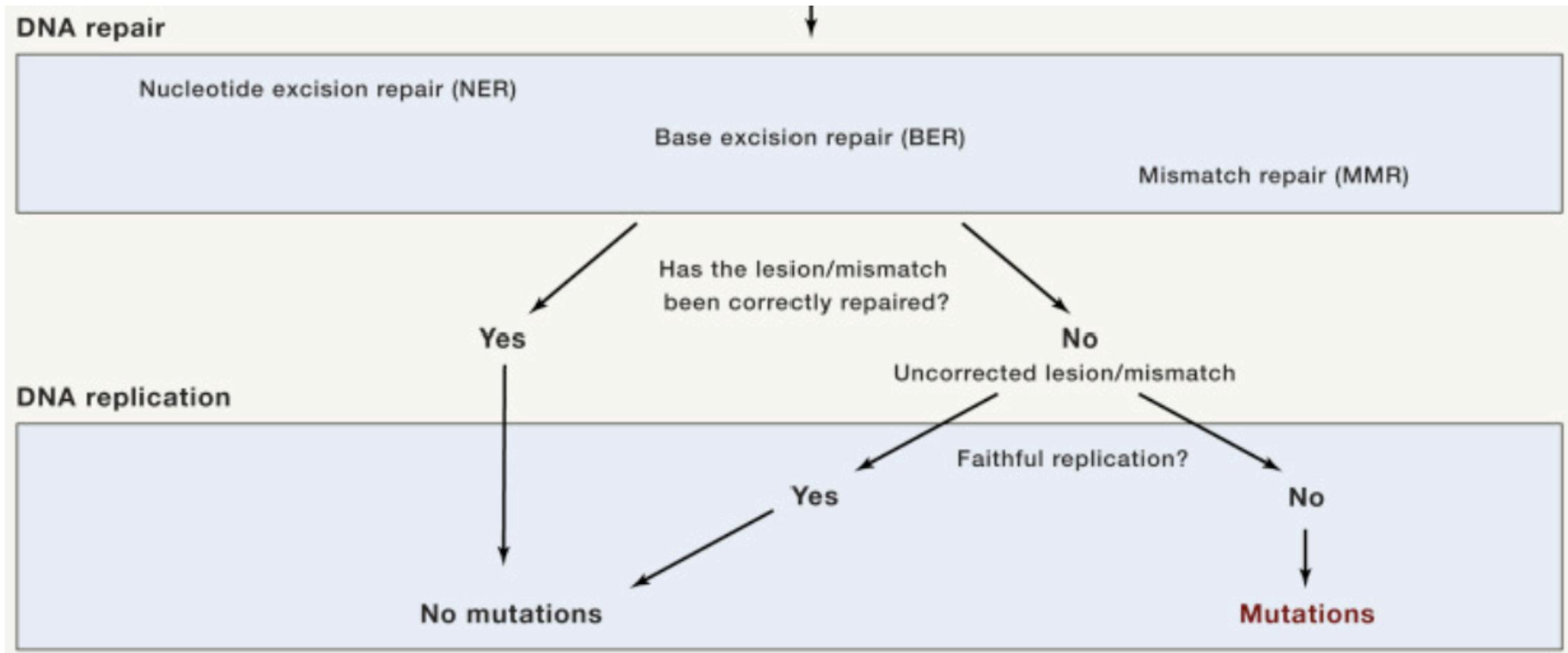


## 2. DNA replication and repair errors <sup>3</sup>

- Replication slippage: DNA polymerases occasionally insert the wrong nucleotide resulting in a base mismatch leading to a misalignment between template and newly synthesized strands



## 2. DNA replication and repair errors <sup>2</sup>



### 3. DNA replication/repair errors <sup>4</sup>

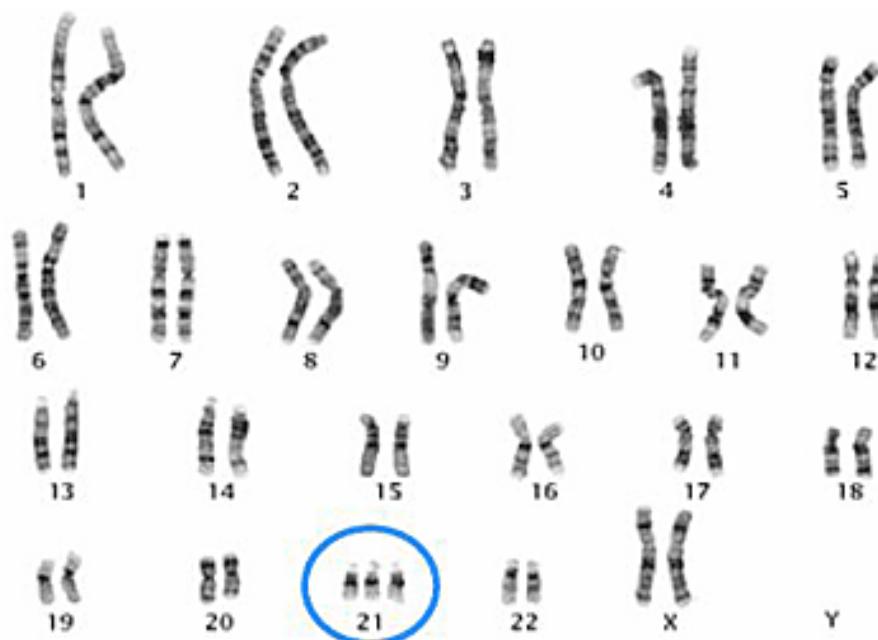
- Dysregulation in DNA polymerases contribute to the mutational load associated with tumors.

**Table 2.** Expression of TLS pathway genes correlates with mutational load in cancer.

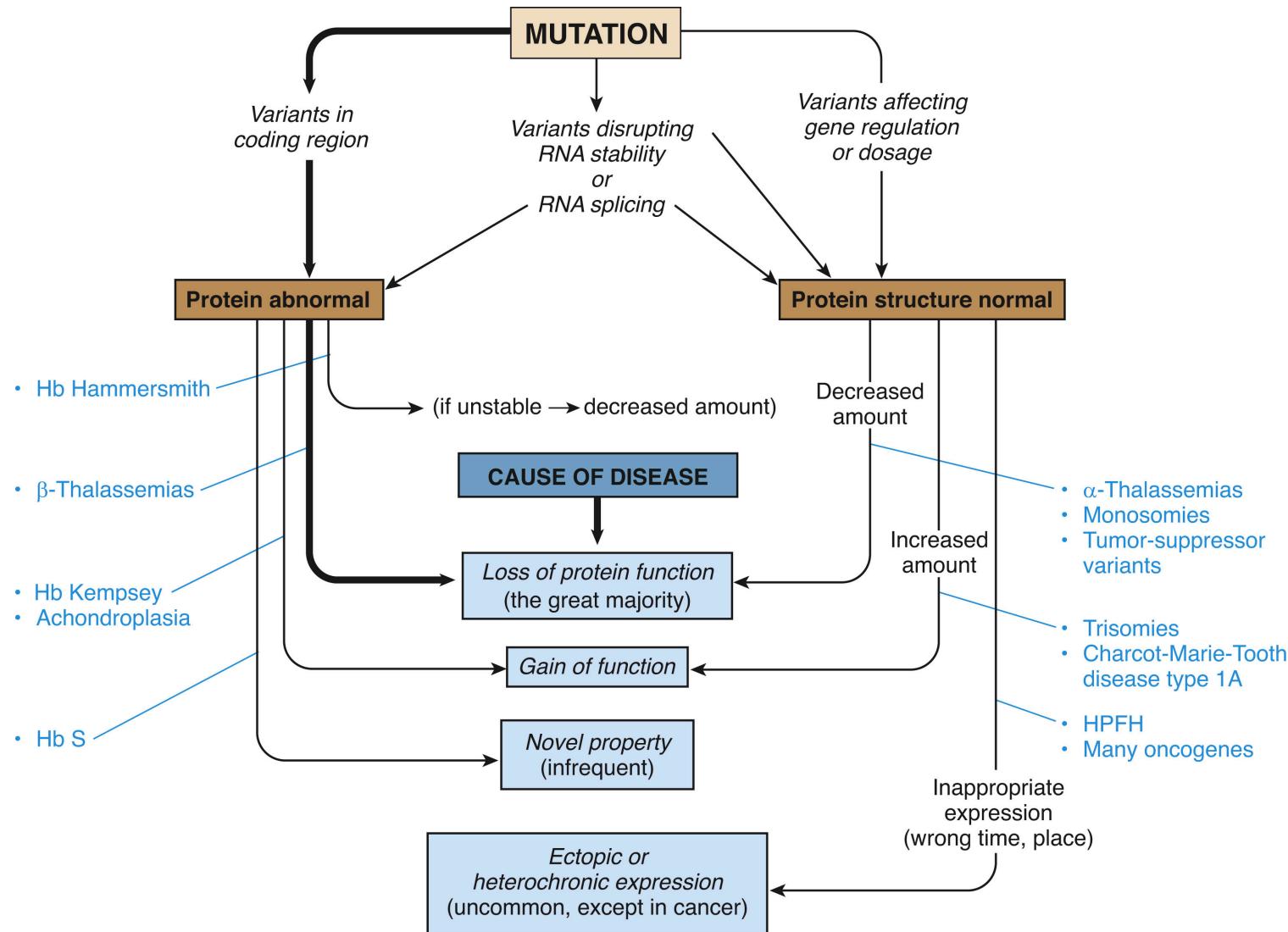
	<b>Bladder cancer (BLCA)</b>	<b>Lung Adenocarcinoma (LUAD)</b>	<b>Lung squamous cell carcinoma (LUSC)</b>
Pol η		<i>Negatively correlated with SNVs in smokers.</i>	Downregulated in tumors <i>Negatively correlated with SNVs in non-smokers.</i>
Pol ε		Downregulated in tumors <i>Negatively correlated with SNVs in smokers.</i>	Downregulated in tumors.
Pol κ	Downregulated in tumors.	Downregulated in tumor. <i>Negatively correlated with SNVs in all tumors.</i>	Downregulated in tumors.
REV1	Downregulated in tumors.	Downregulated in tumors. <i>Negatively correlated with SNVs in smokers.</i>	
MAD2L2	Overexpressed in tumors.	Overexpressed in tumors. <i>Positively correlated with SNVs in smokers.</i>	Overexpressed in tumors.
REV3L	Downregulated in tumors.	Downregulated in tumors.	Downregulated in tumors.
RAD18	Overexpressed in tumors. <i>Positively correlated with SNVs.</i>	Overexpressed in tumors. <i>Positively correlated with SNVs in all tumors irrespective of smoking history.</i>	Overexpressed in tumors. <i>Positively correlated with SNVs in smokers.</i>

### 3. Chromosome segregation and recombination errors <sup>5</sup>

- Errors in chromosome segregation changes chromosomal DNA copy number
- Sometimes chromatids can misalign during recombination and subsequent crossovers results in duplication/deletion
- Example: Individuals with Down syndrome has a third copy of chromosome 21.



# Mechanisms of pathogenesis

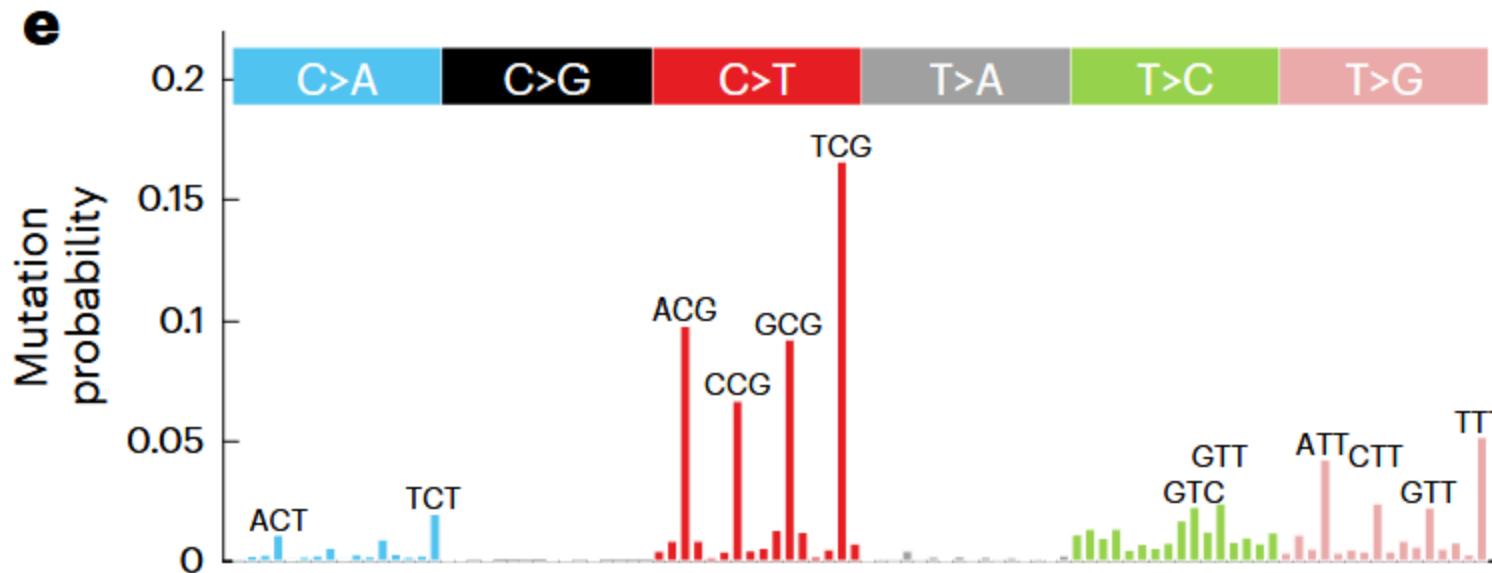
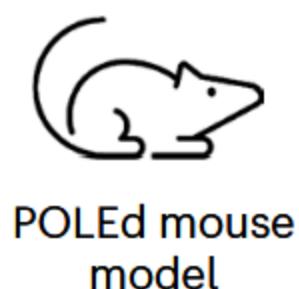


# Types of Variation in Human Genetic Disease

Type of Variation	Percentage
Missense variants (amino acid substitutions)	40%
Nonsense variants (premature stop codons)	10%
RNA processing variants (destroy consensus splice sites, cap sites, and polyadenylation sites or create cryptic sites)	10%
Splice-site variants leading to frameshift mutations and premature stop codons	10%
Long-range regulatory variants	Rare

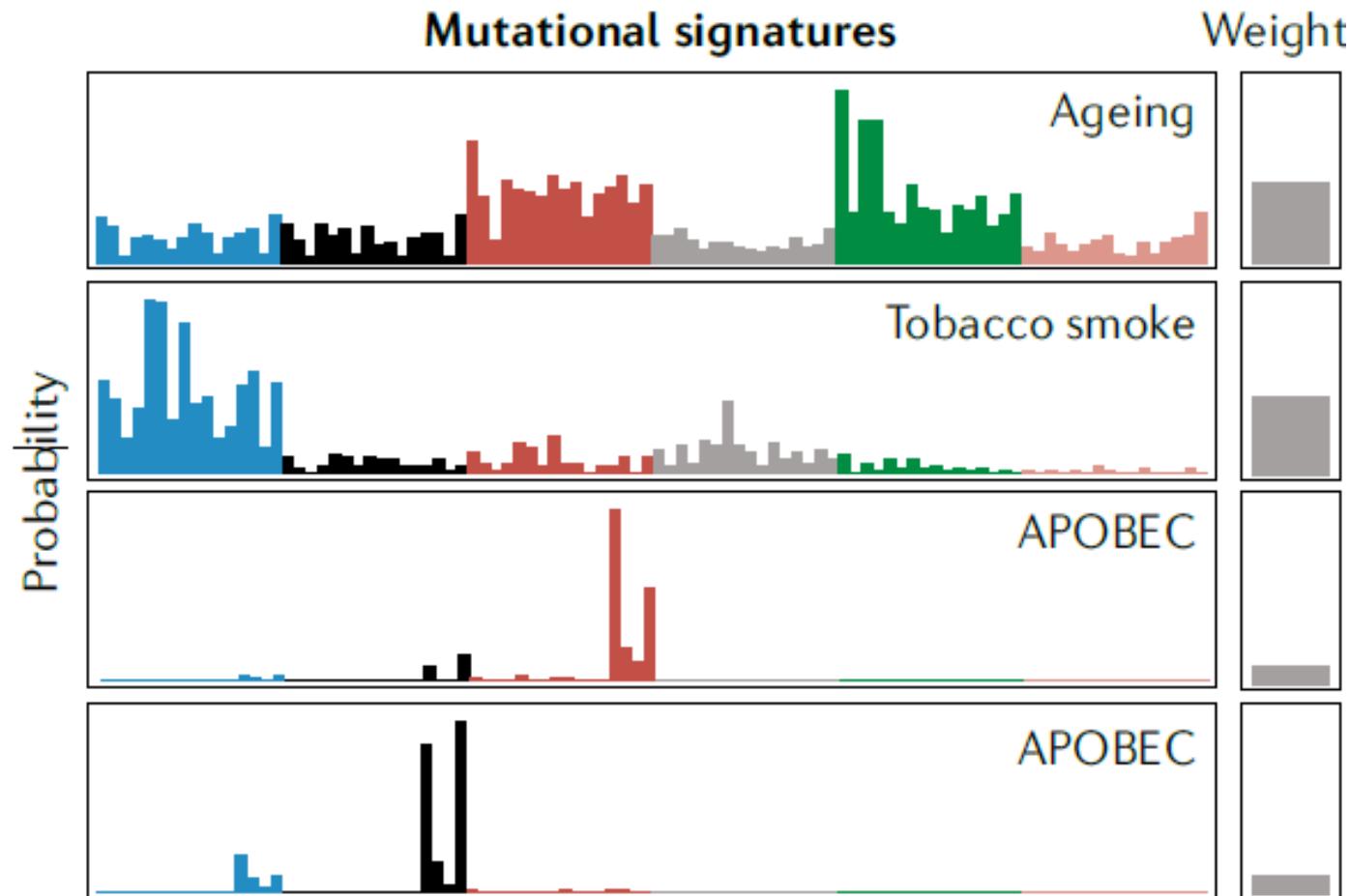
## Pathogenic Variants - Indels <sup>6</sup>

- Example: The most common mutation type for tumor, caused by DNA polymerase errors, is a substitution from cytosine to thymine in a CpG dinucleotide (CpG>TpG)

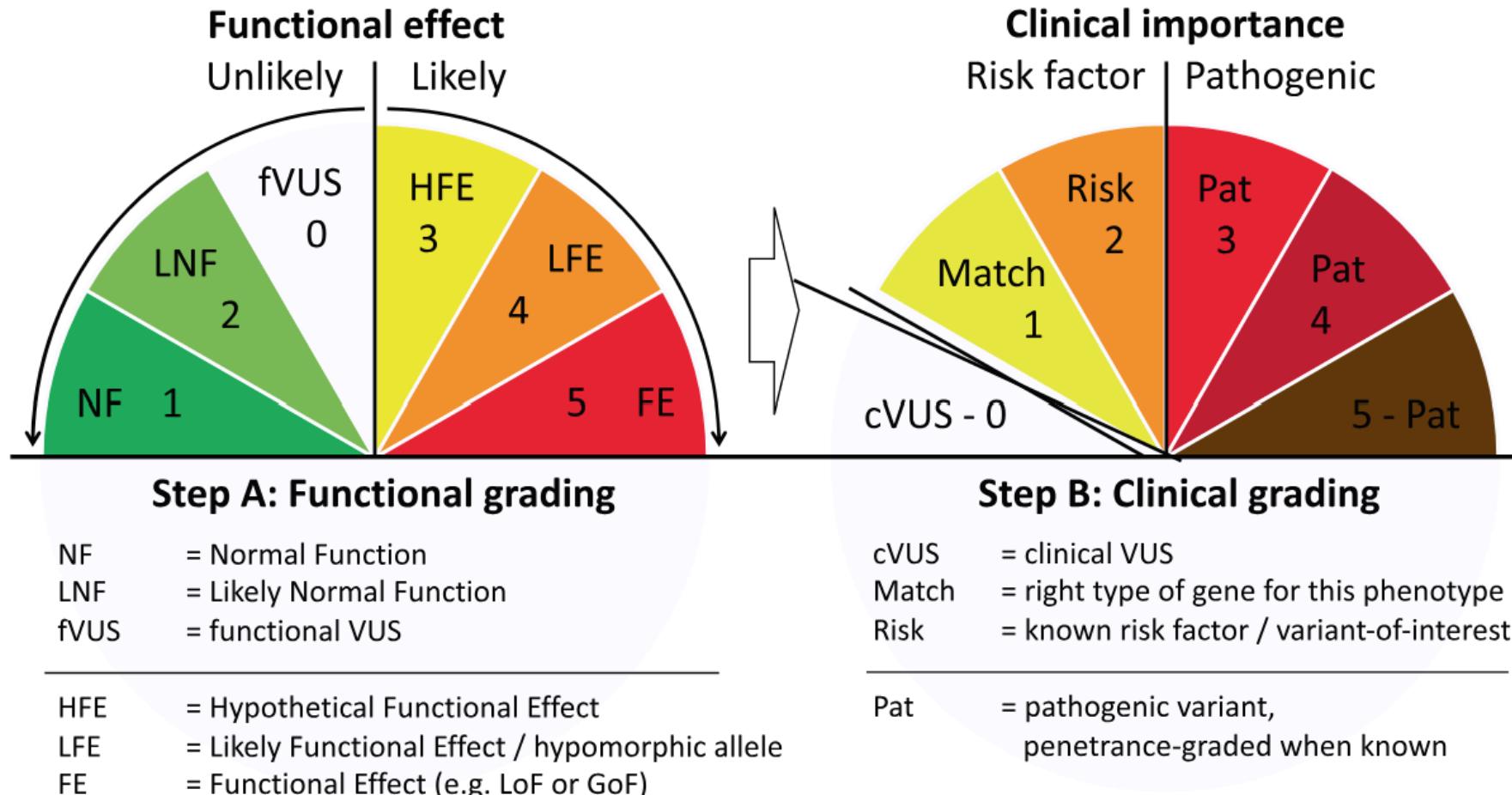


# Pathogenic Variants - Indels<sup>6</sup>

- Each polygenic disease has a unique mutational signature.

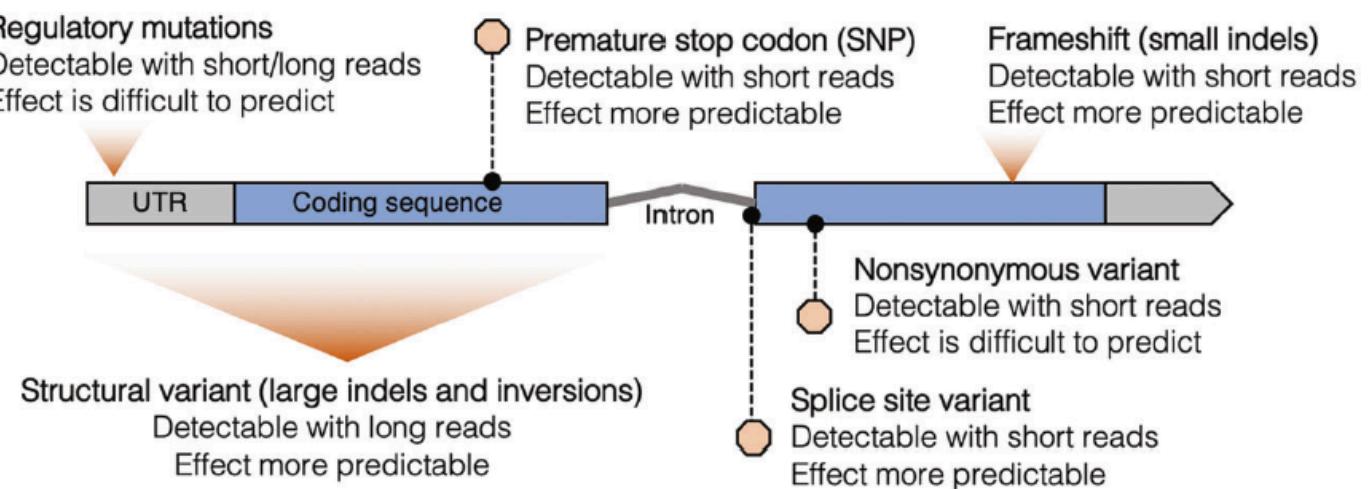


# Why does the type of pathogenic variant matter? <sup>7</sup>



# Loss of Function Variants

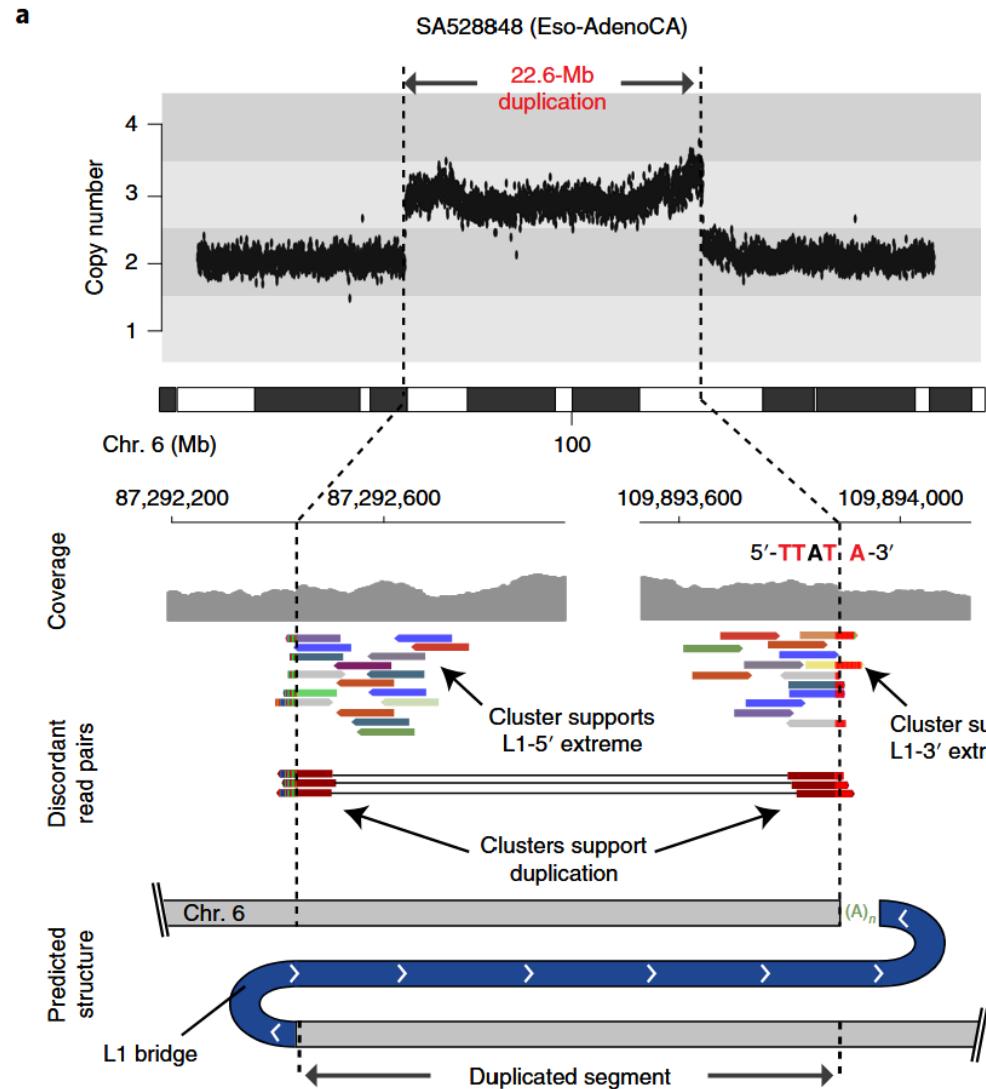
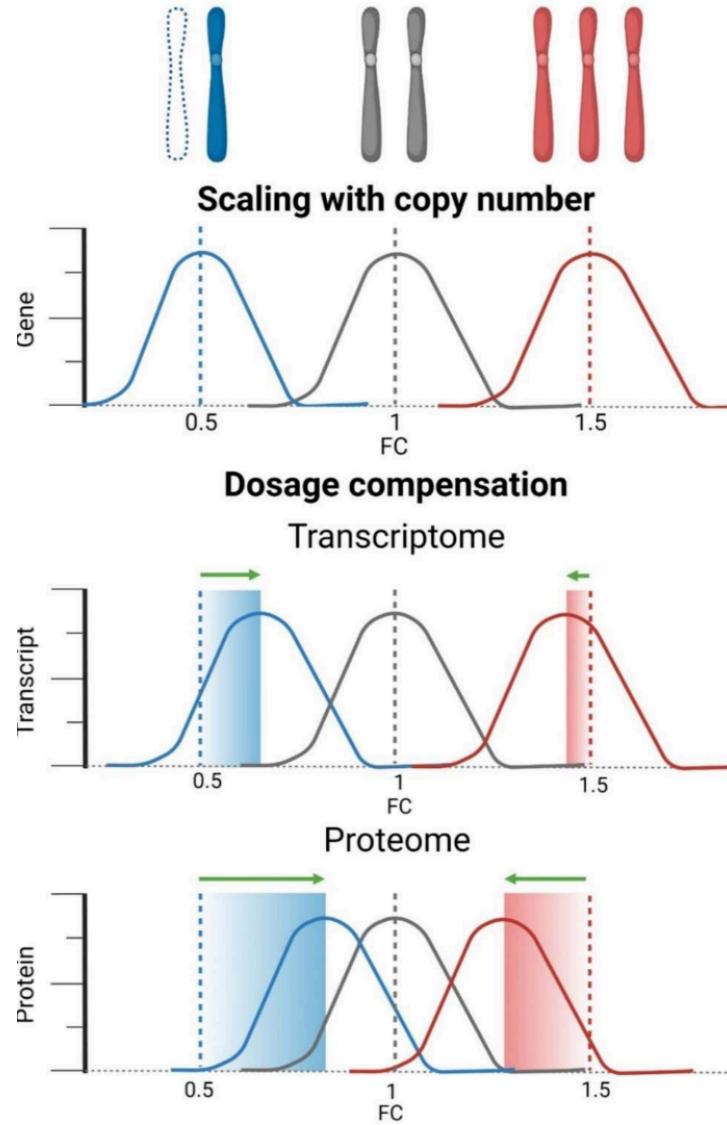
- **Loss of function (LoF) variant:** the complete loss of a protein encoded from the allele, due to loss of the allele, unstable mRNA or unstable and inactive protein.



## Gain of Function Variants

1. Variants that increase the production of a normal protein.
  - **Example:** Extra copy of chromosome 21 in Down Syndrome
2. (Rare) Variants that improves the normal function of a protein.
  - Example: a missense variant that creates hemoglobin Kempsey locks hemoglobin into its high oxygen affinity state, thereby reducing oxygen delivery to tissues.

Gene dosage is the number of copies of a particular gene present in a genome.<sup>8,9</sup>



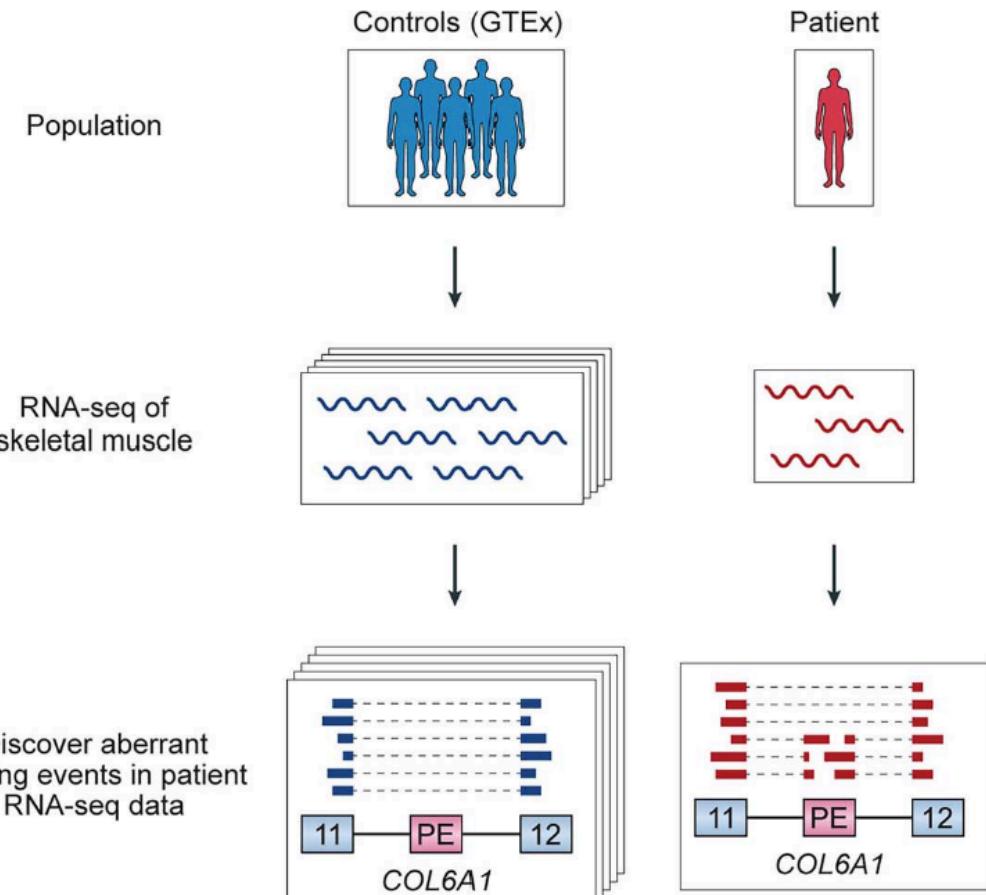
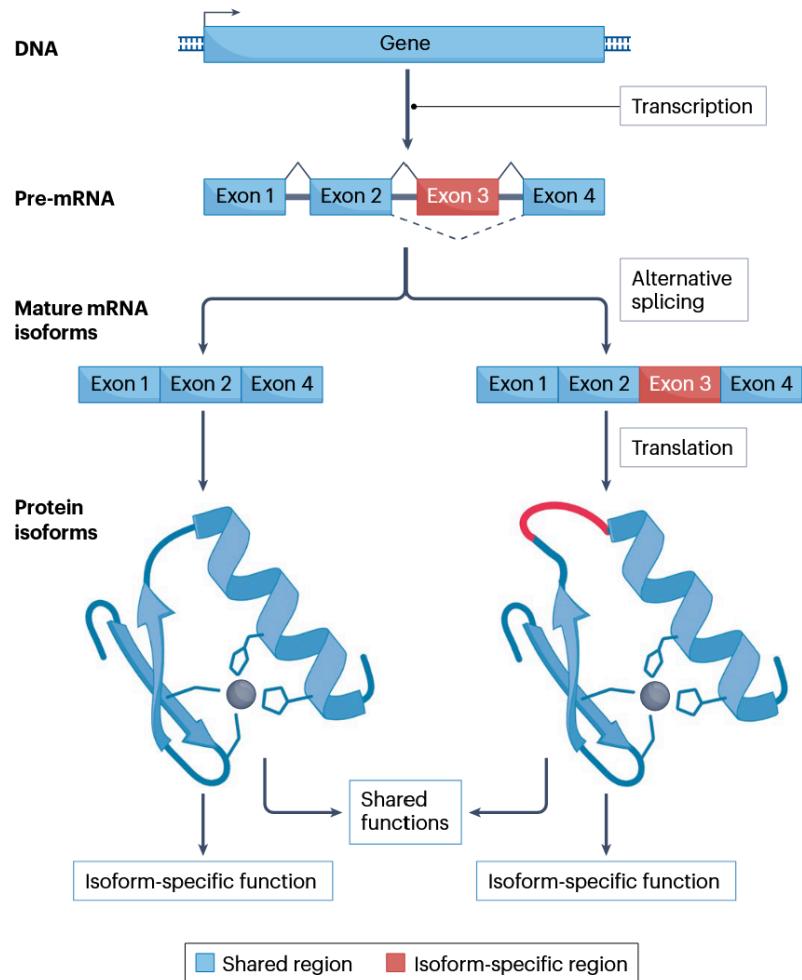
# Mutation hierarchy<sup>8</sup>

- Functional consequence determines the **mutation hierarchy** in GenVisR

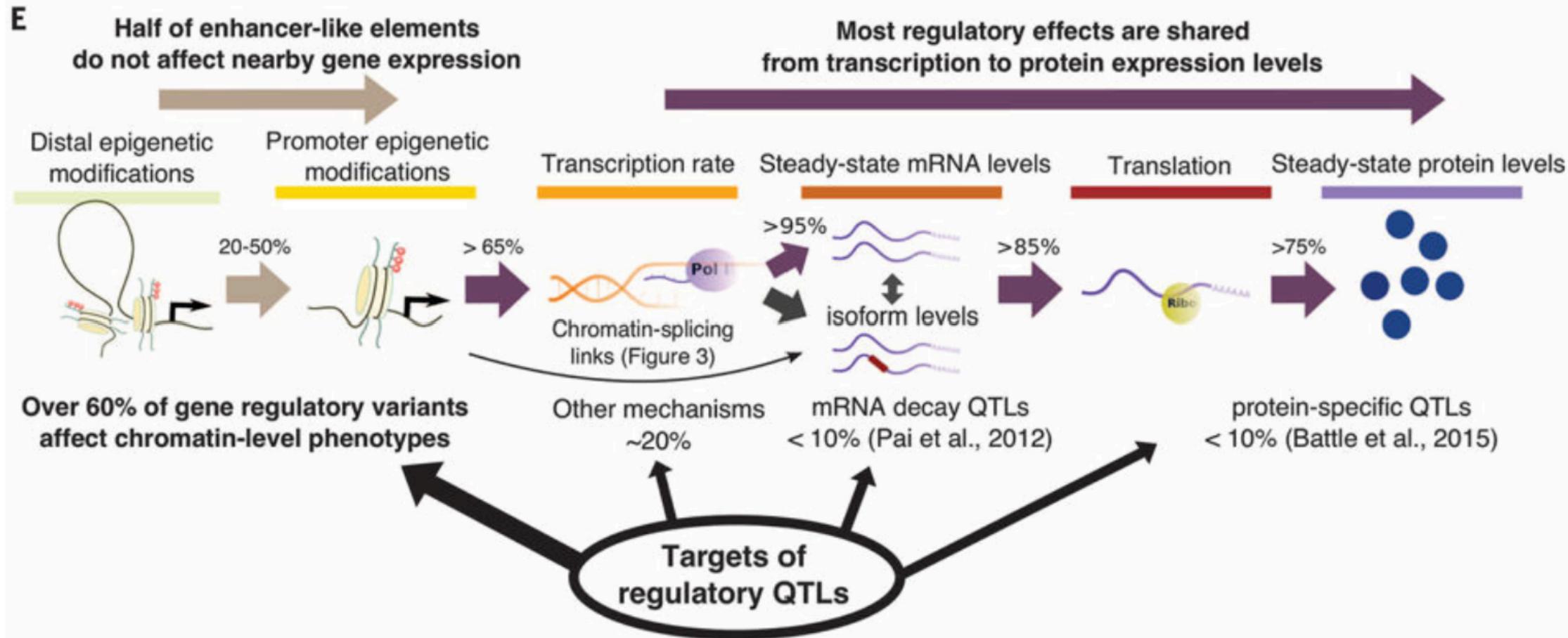
```
mutHierarchy <- c("nonsense", "frameshift", "splice_site", "missense")
waterfall(brcaMAf, ..., mutationHierarchy = mutHierarchy)
```



# Protein polymorphism <sup>9,10</sup>

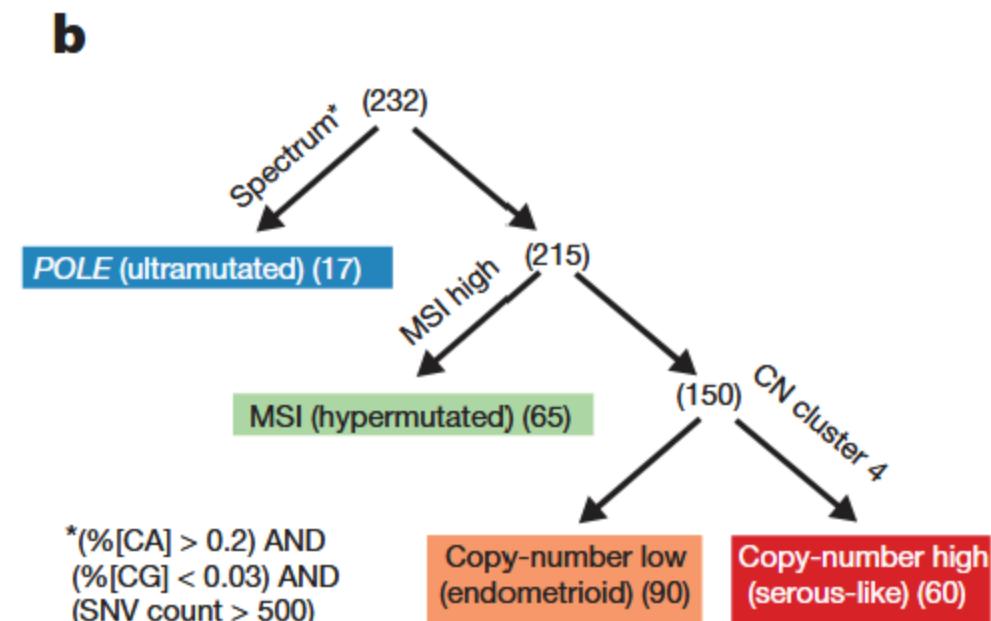
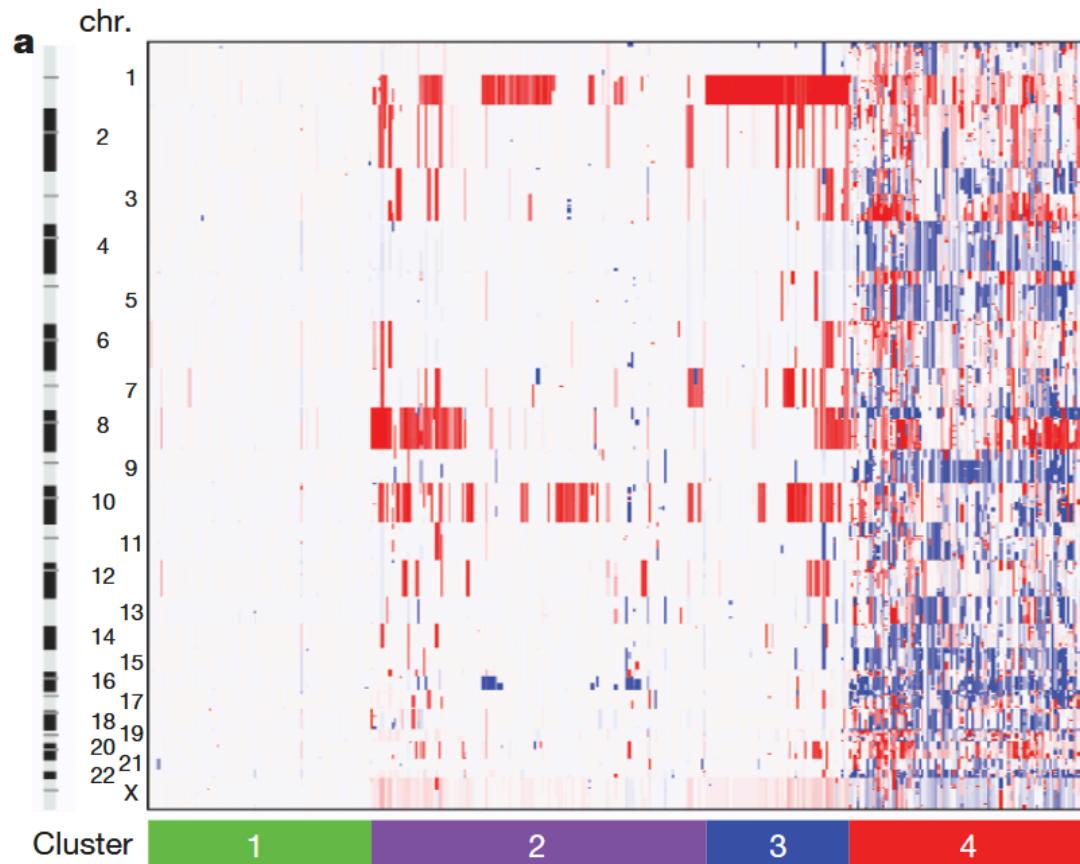


# Protein polymorphism<sup>11</sup>

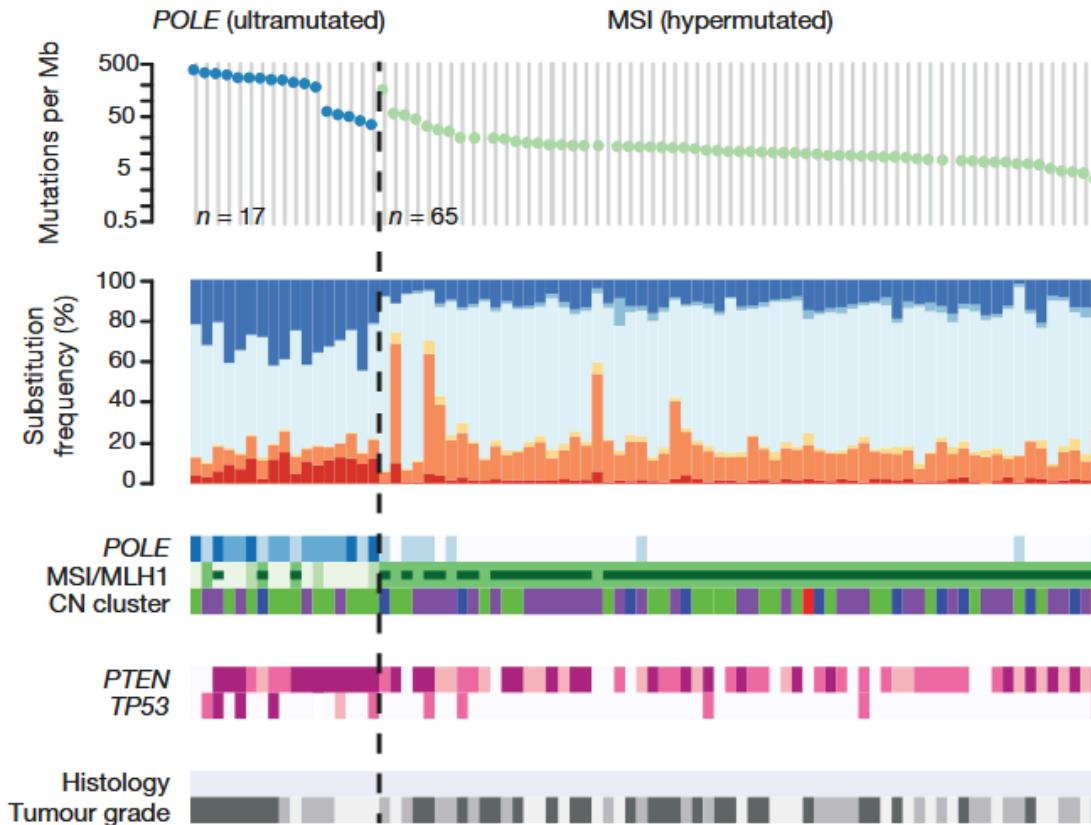


# Case Study: Variable expression of complex diseases

- An integrated genomic, transcriptomic and proteomic characterization of 373 endometrial carcinomas <sup>11</sup>



# Mutation spectra across four groups of endometrial carcinoma<sup>11</sup>

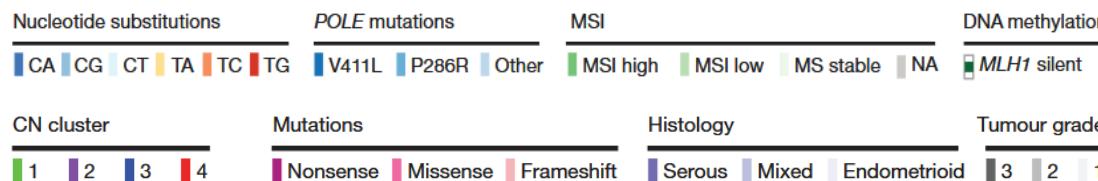


## 1. POLE (ultramutated):

- unusually high mutation rate
- high frequency of C>A transversions

## 2. MSI (hypermutated):

- high mutation rate
- MLH1 promoter methylation



# Mutation spectra across four groups of endometrial carcinoma<sup>11</sup>



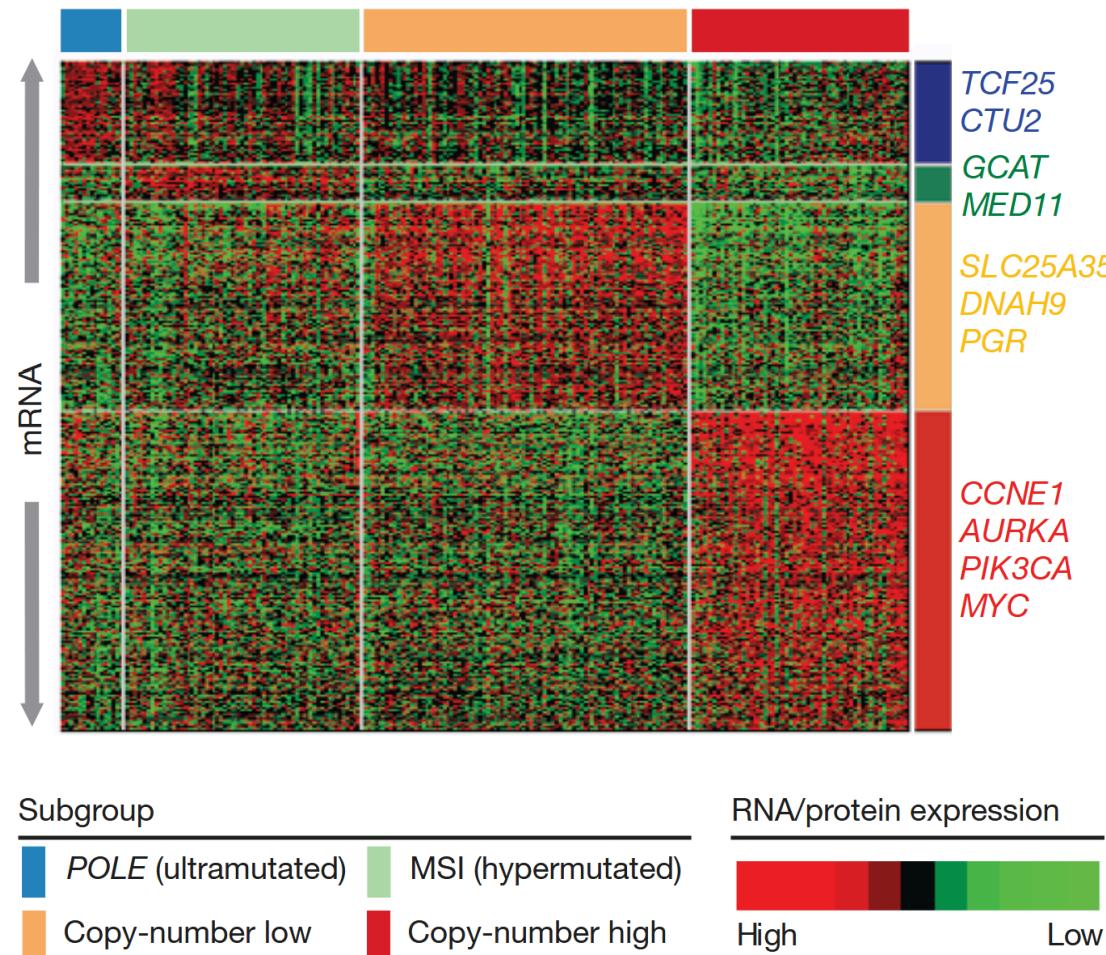
## 3. Copy-number low:

- lower mutation frequency
- microsatellite stable (MSS)

## 4. Copy-number high:

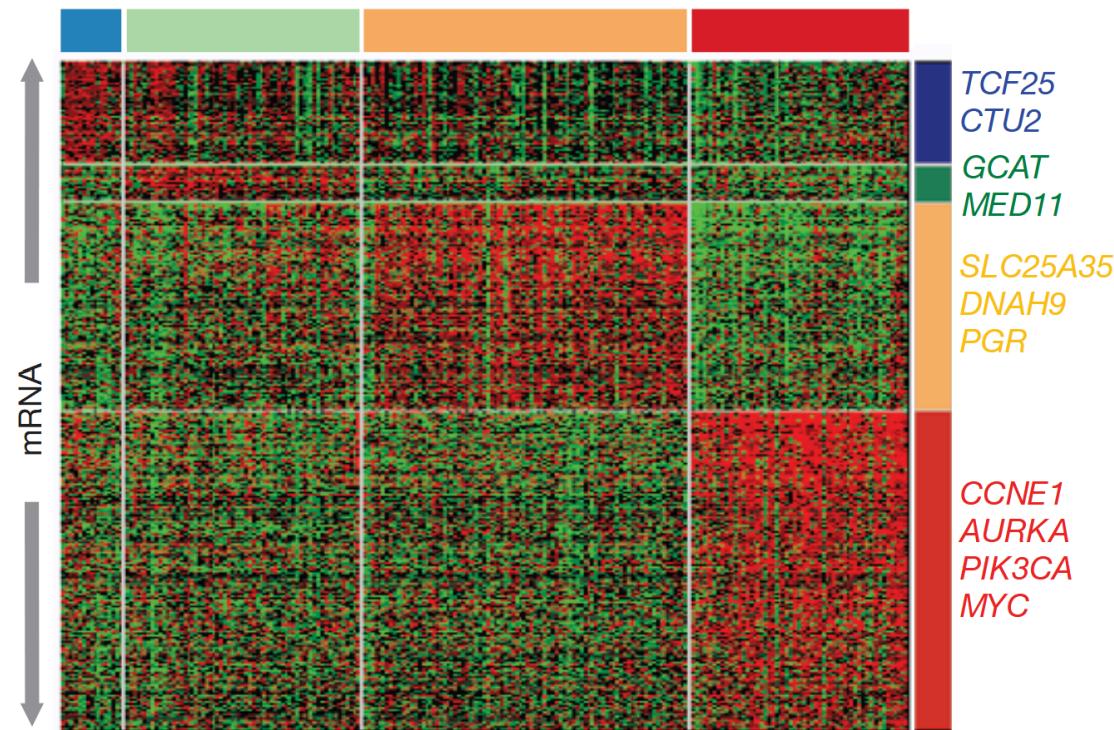
- low mutation rate
- high copy number (CN cluster 4)
- TP53 mutation

# Gene expression across subtype clusters in endometrial carcinomas<sup>11</sup>

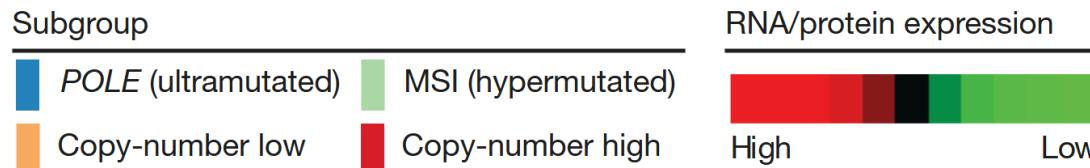


- **POLE cluster:** signature genes mostly involved in cellular metabolism
- **MSI cluster:** decreased *MLH1* expression (probably due to its promoter methylation)
- **Copy-number low cluster:** increased progesterone receptor (*PGR*) expression suggests responsiveness to hormonal therapy

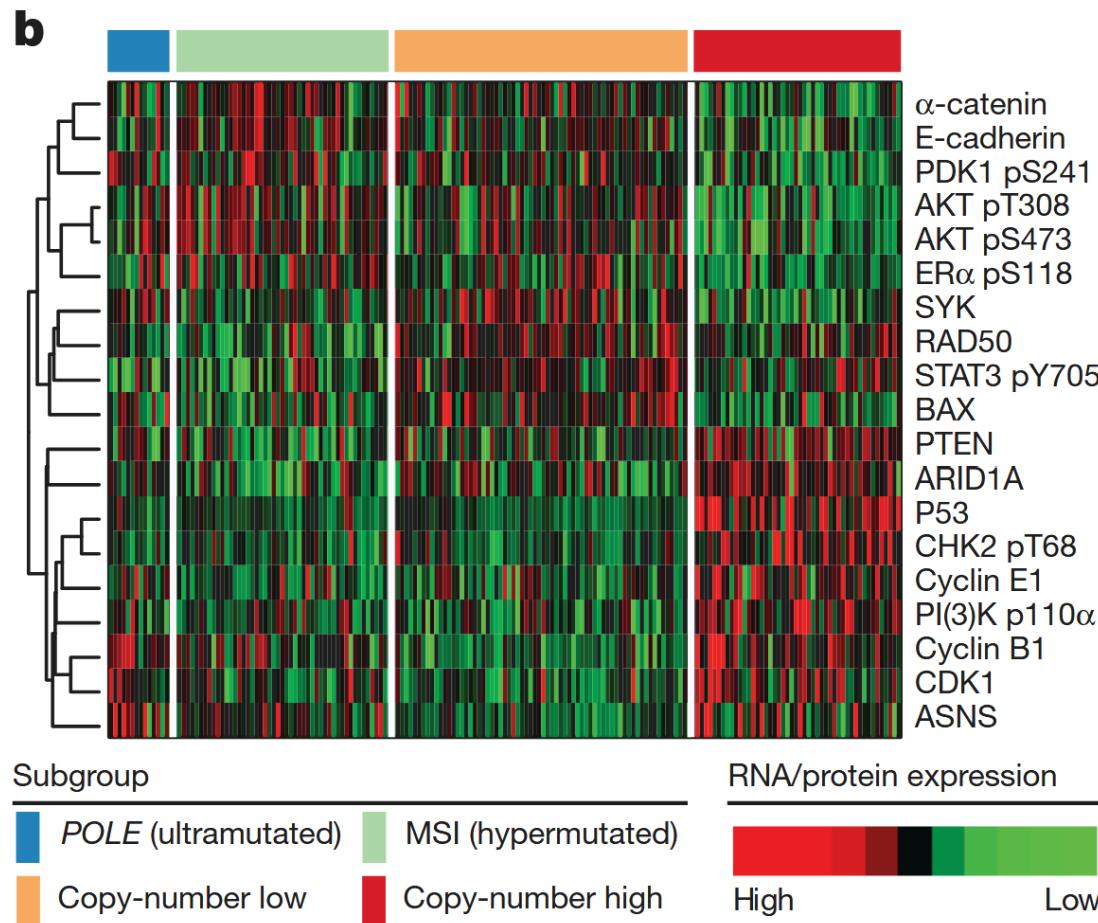
# Gene expression across subtype clusters in endometrial carcinomas<sup>11</sup>



- **Copy-number high cluster:** increased expression in genes involved with cell cycle deregulation (*CCNE1*, *PIK3CA*, *MYC*, *CDKN2A*)

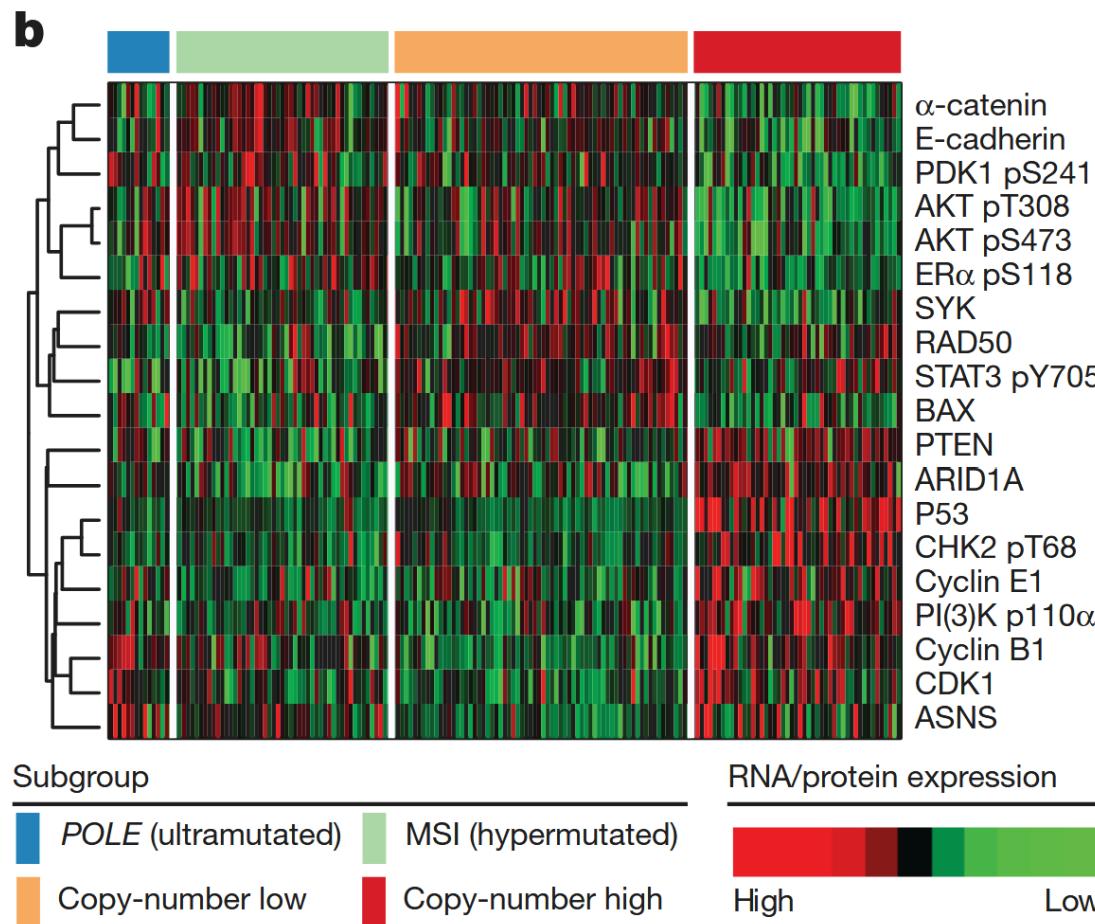


# Protein expression across subtype clusters in endometrial carcinomas<sup>9</sup>



- TP53 was frequently mutated in the copy-number high group and its protein expression was also increased.
- (Class poll Q1) These mutations are probably \_\_\_\_ of function variants.

# Protein expression across subtype clusters in endometrial carcinomas<sup>11</sup>



- PTEN had high mutation rates in the remaining groups, but their protein expression was decreased.
- (Class poll Q2) The mutations in these genes are probably \_\_\_\_\_ of function variants.