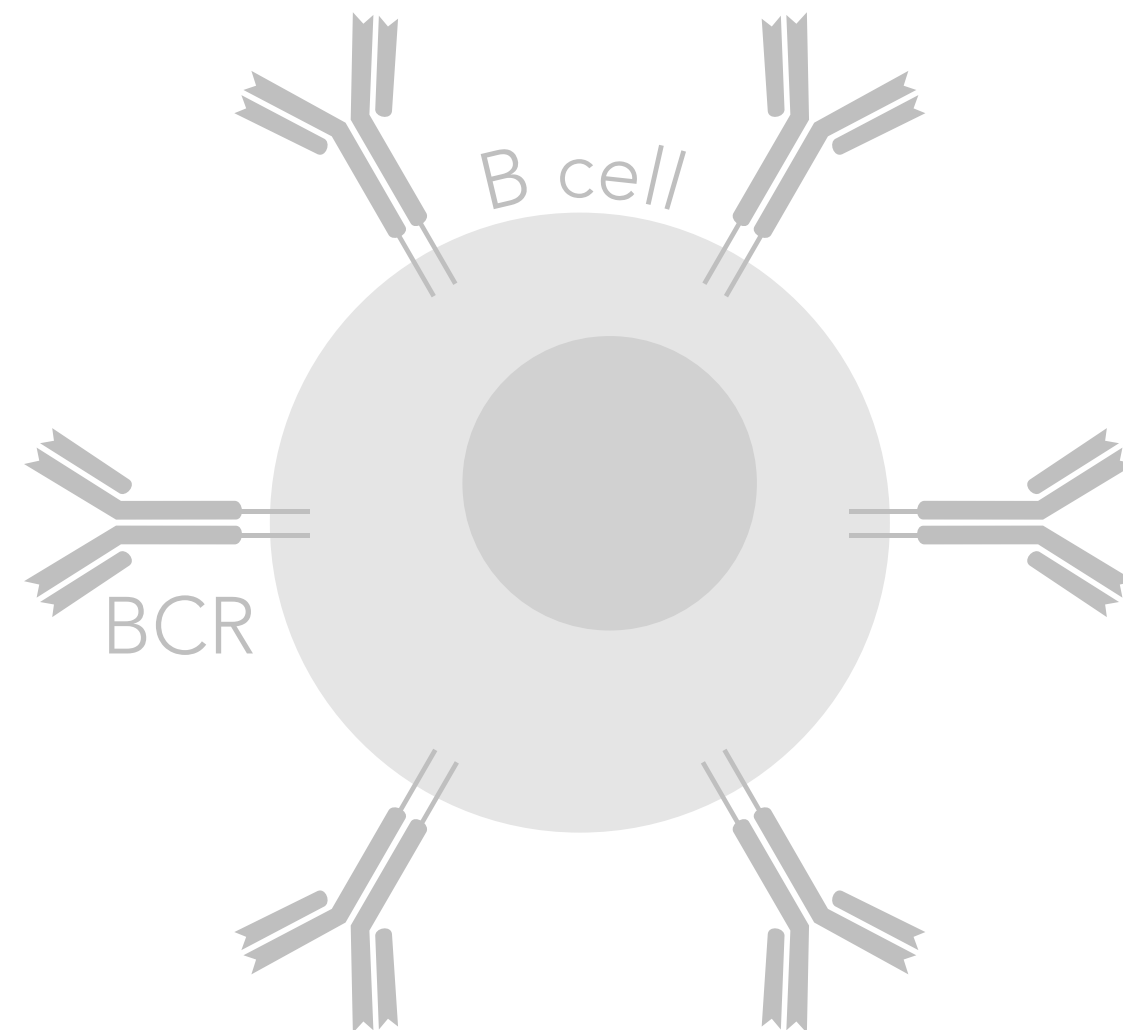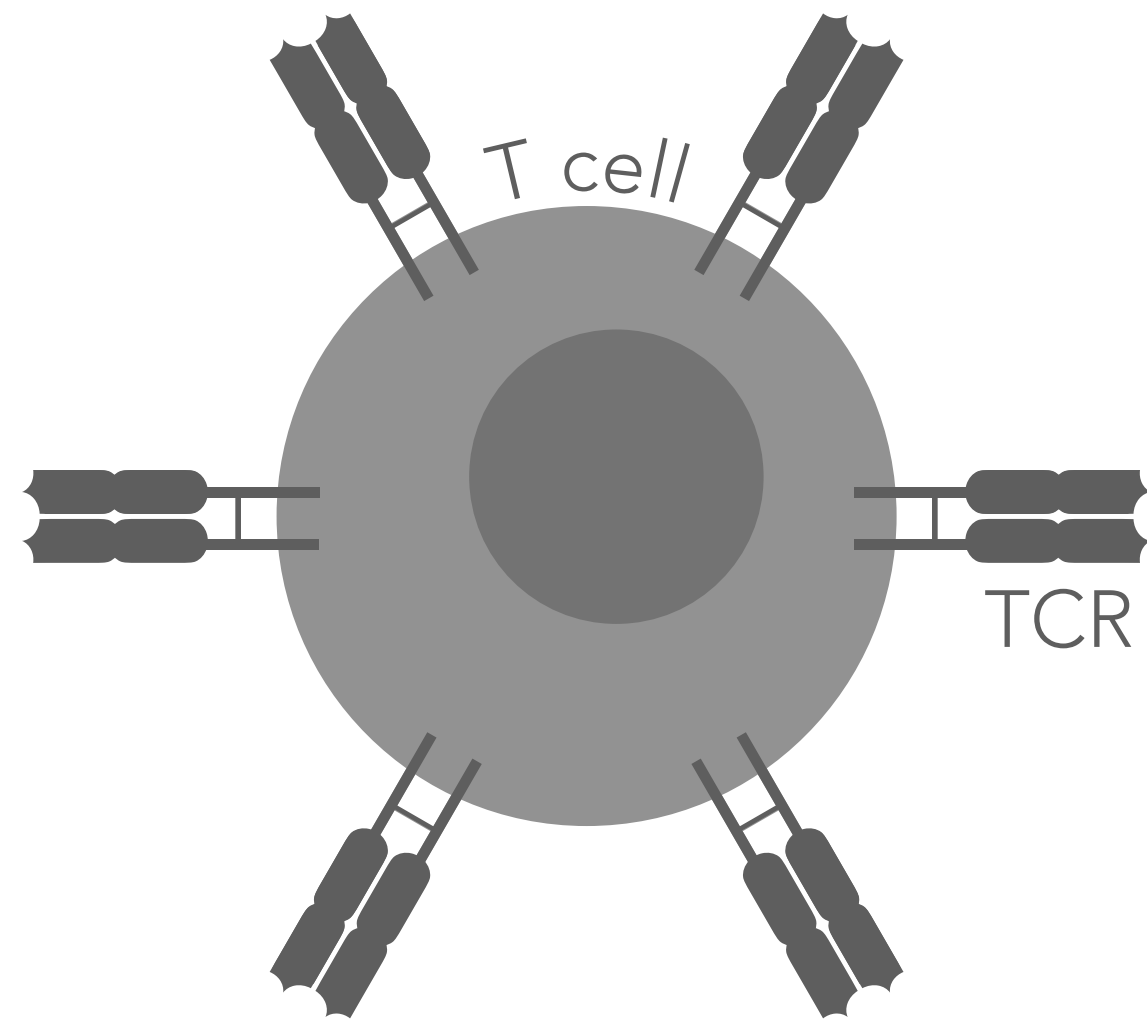# TFCB LECTURE 17:
## INTRO TO IMMUNE REPERTOIRE SEQUENCING AND ANALYSIS

MAGGIE RUSSELL

Adaptive immunity protects us from infection

- B and T cell receptors recognize pathogen-derived antigens

- Our repertoires contain **millions** of these unique receptor sequences

- Repertoires reveal immune history, vaccine response, clonal dynamics, etc.
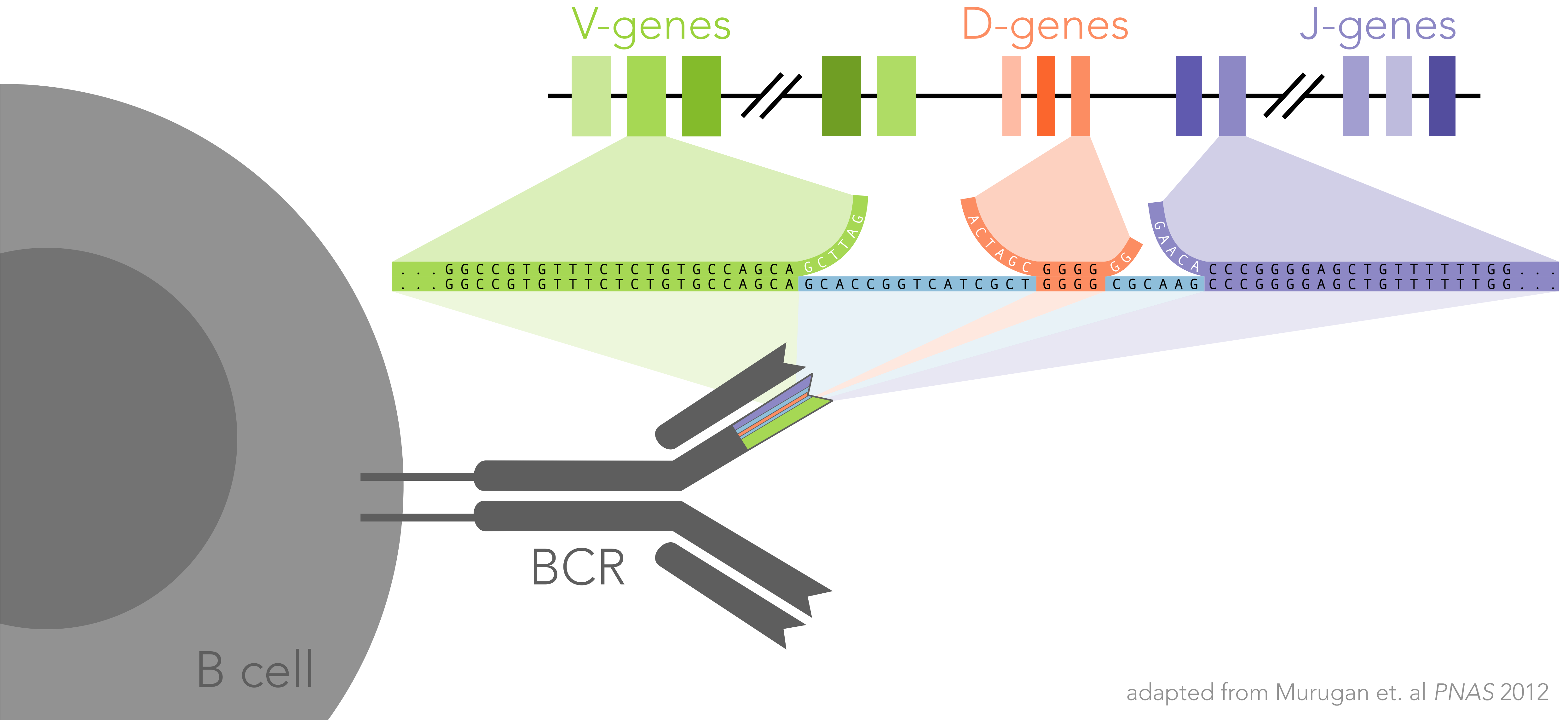
- Computational analysis is required to interpret them…

B cells express BCRs and can secrete antibodies
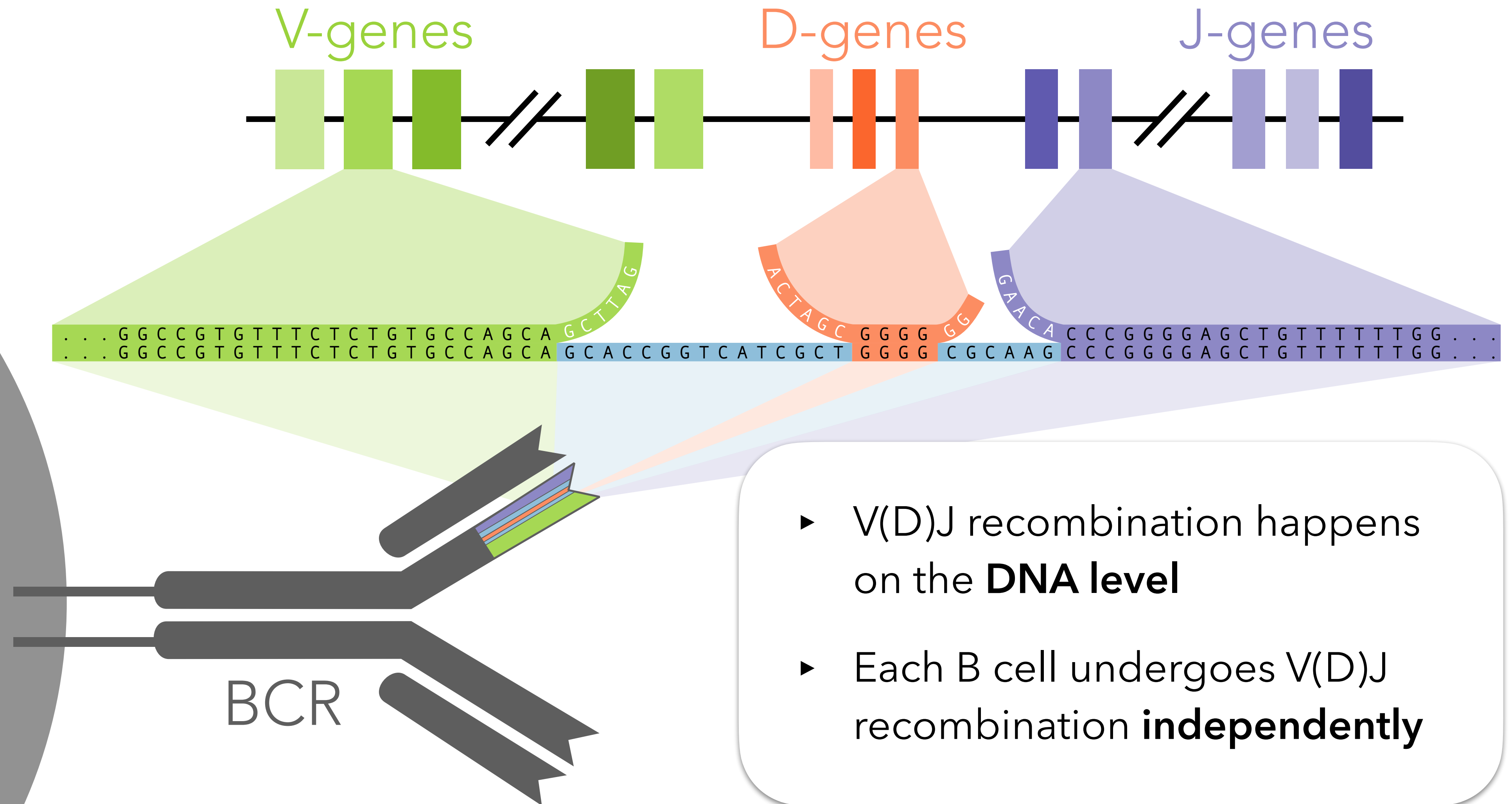

B cell

antigen

BCR

antibody

3
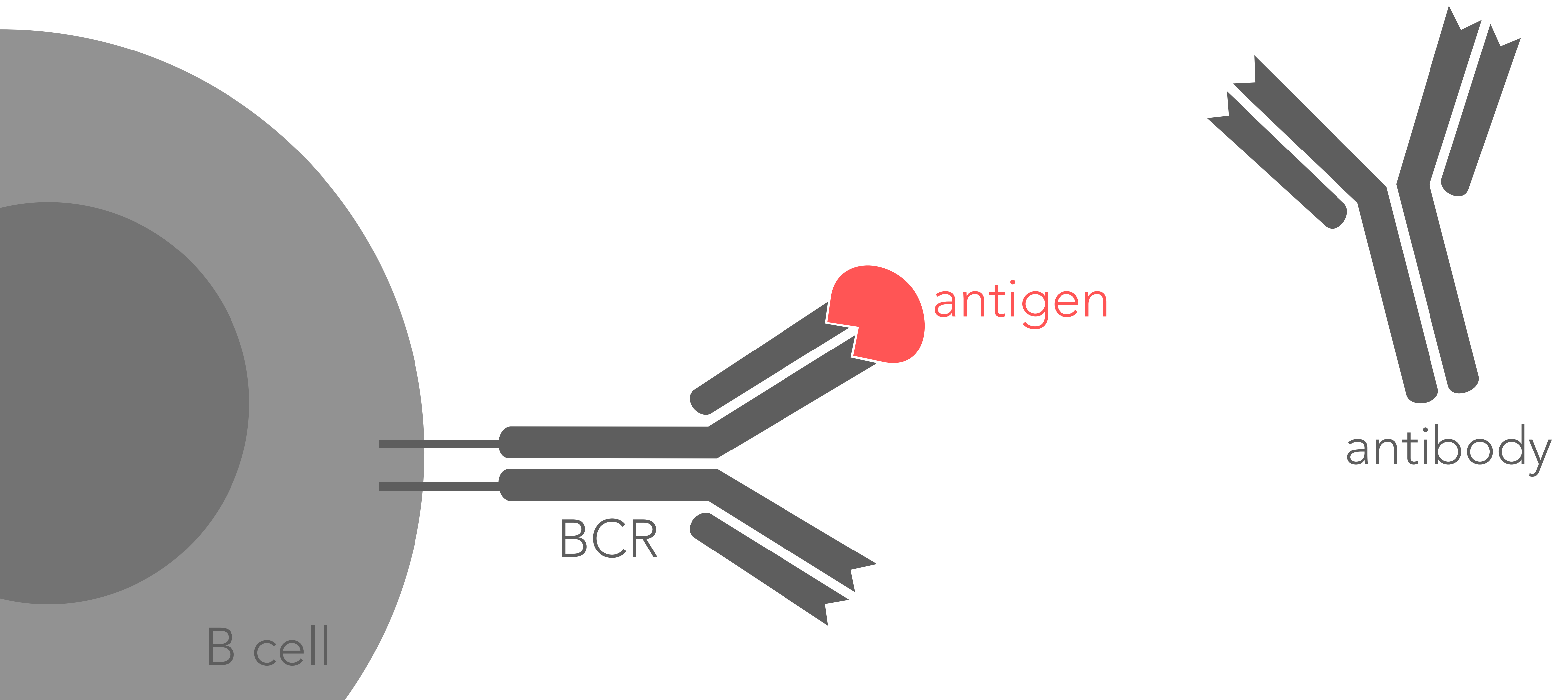
Antibodies consist of several distinct regions
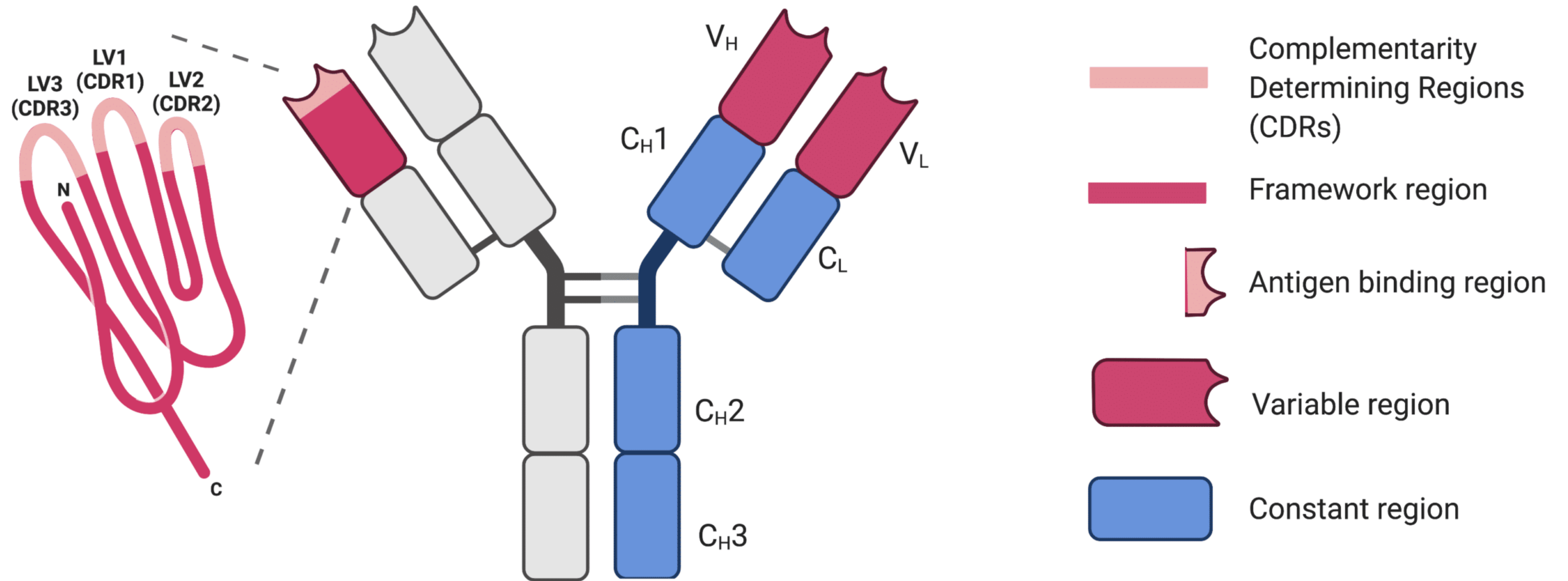
V(D)J recombination generates the variable region

# V(D)J recombination generates the variable region

V-genes D-genes J-genes

...GGCCGTGTTTCTCTGTGCCAGCA GCTTAG
...GGCCGTGTTTCTCTGTGCCAGCA GCACCGGTCATCGCT GGGG CGCAAG CCCGGGGAGCTGTTTTTTGG...

ACTAGC GGGG GG
GAACA CCCGGGGAGCTGTTTTTTGG...

B cell

BCR

▸ V(D)J recombination happens on the **DNA level**

▸ Each B cell undergoes V(D)J recombination **independently**

naive B cells often have **low affinity** for antigen
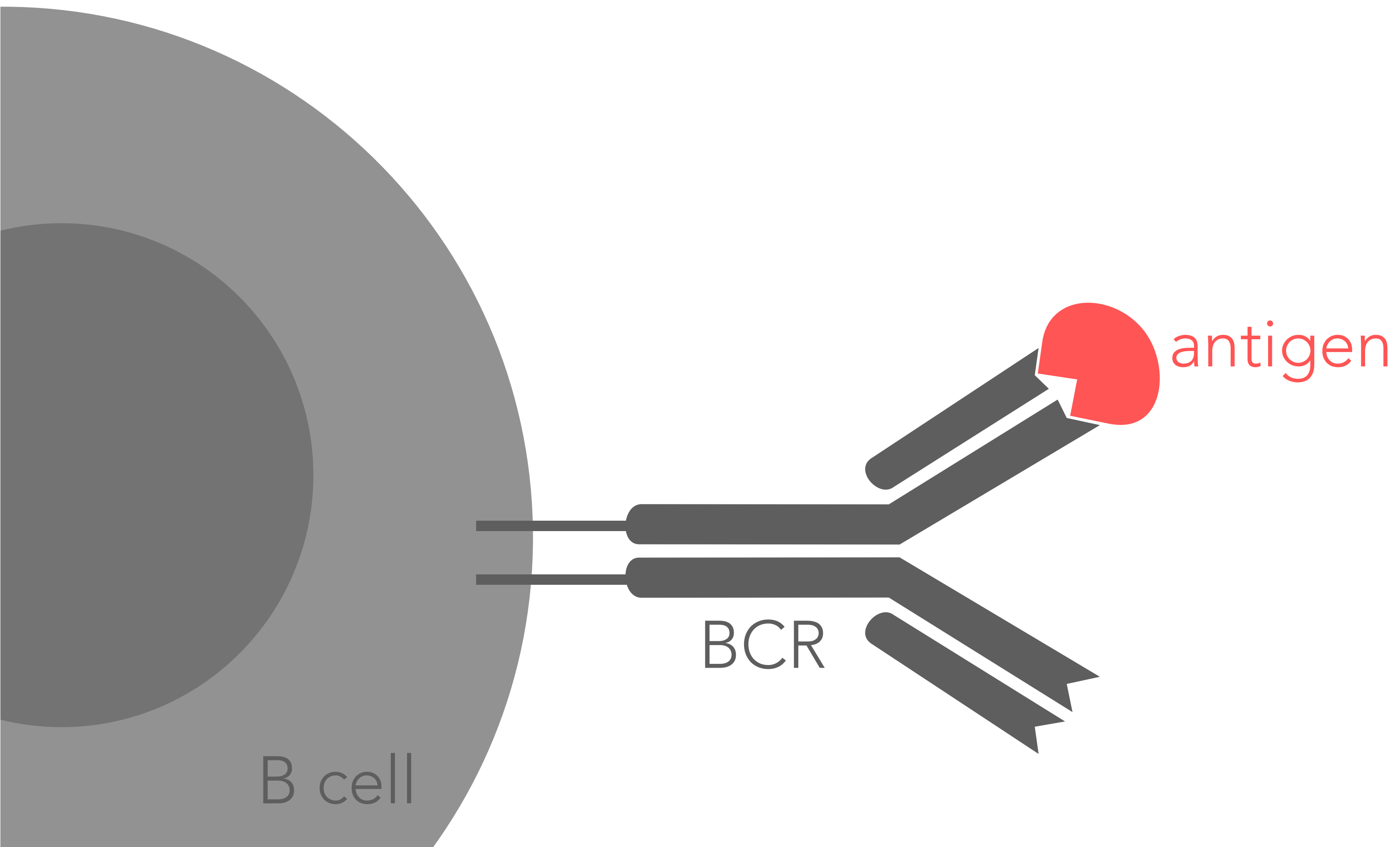
antigen

BCR

B cell

B cells evolve within **germinal centers** to **increase affinity**

↑ *affinity*
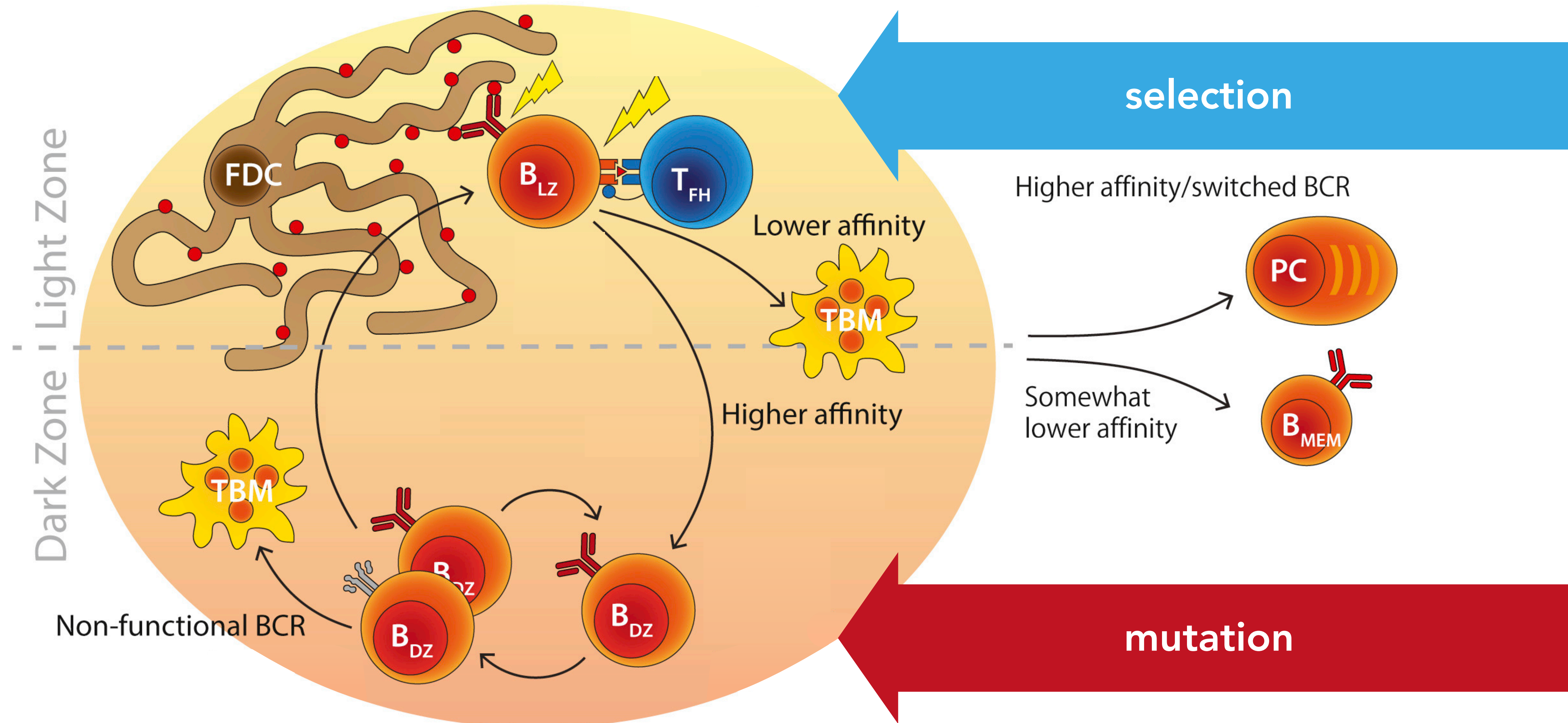
antigen

BCR

B cell
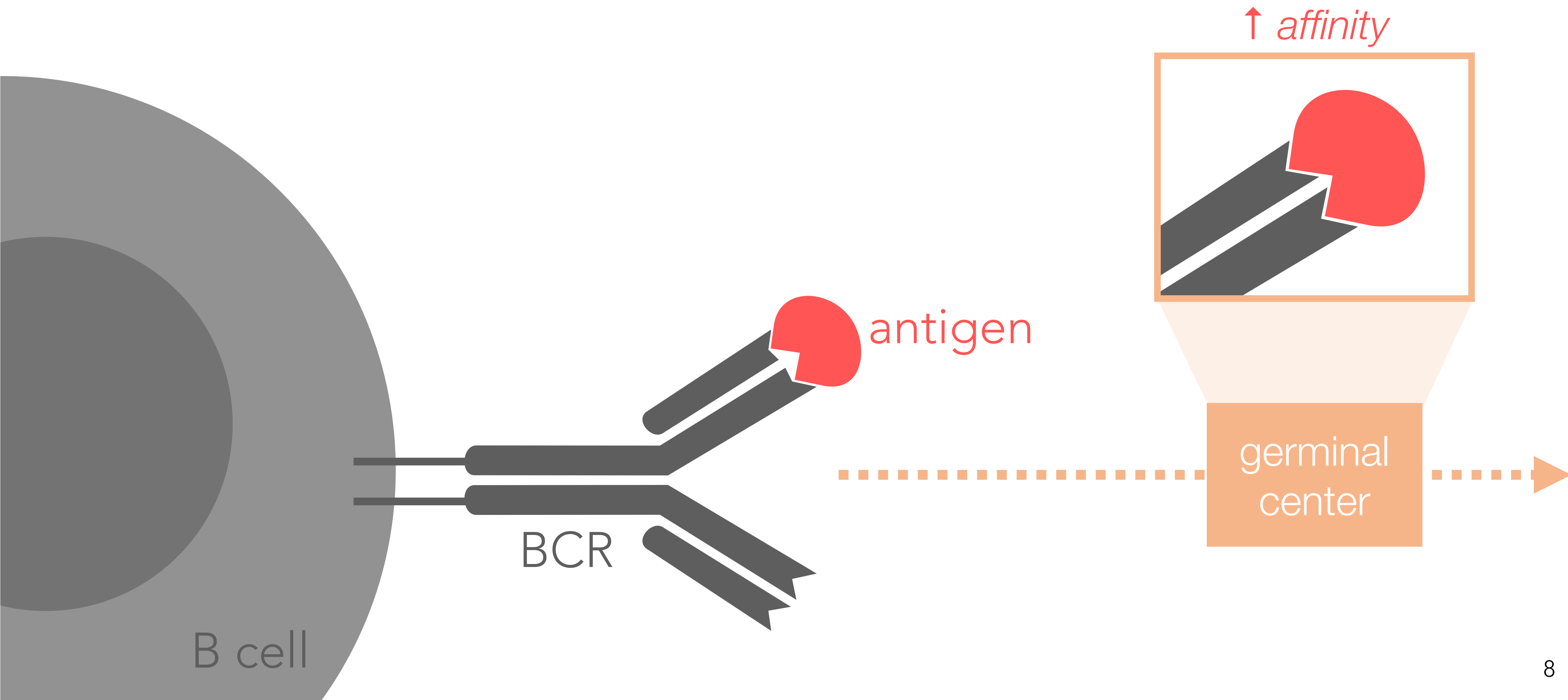
germinal center

B cells evolve within **germinal centers** to **increase affinity**
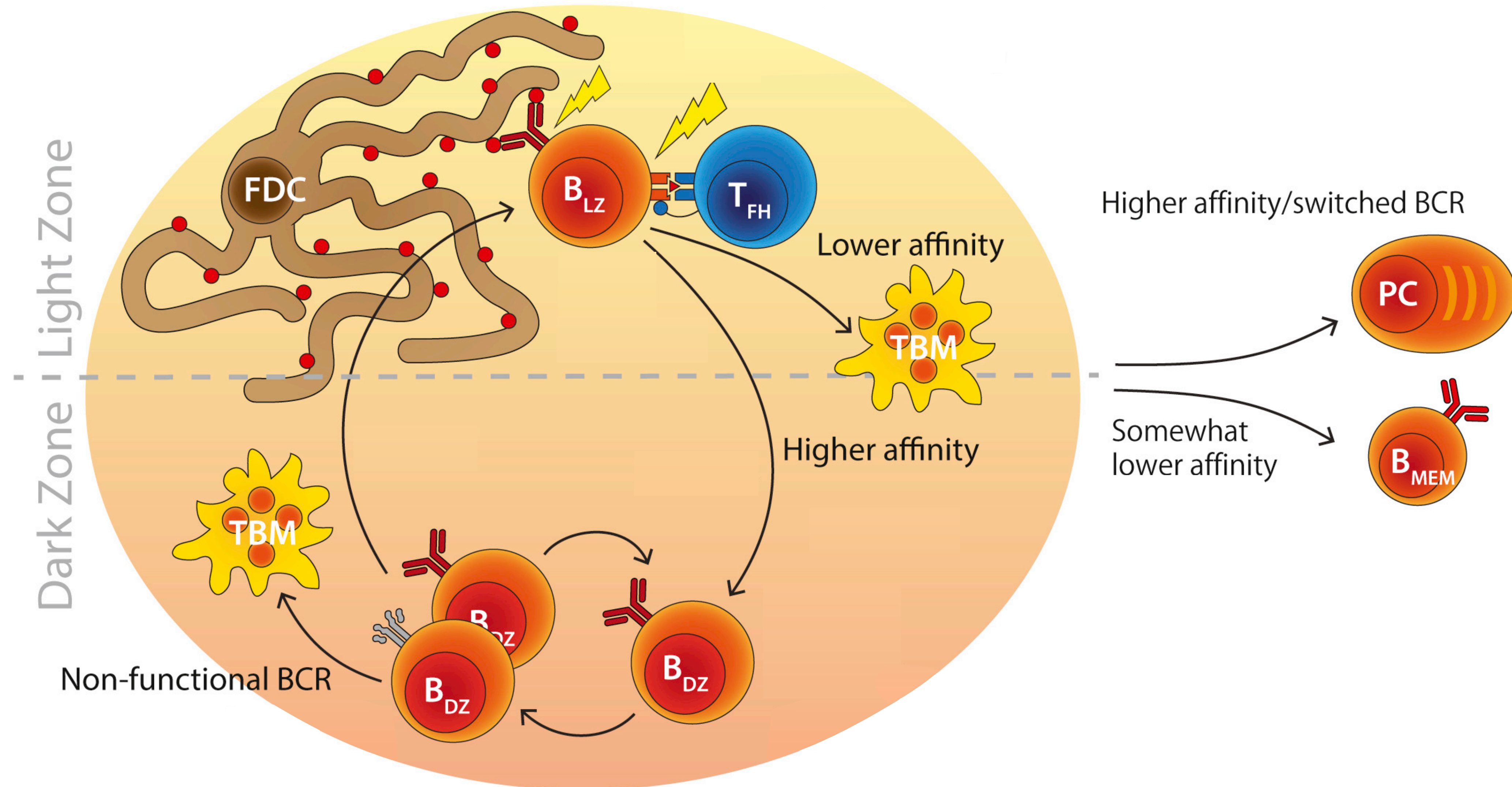
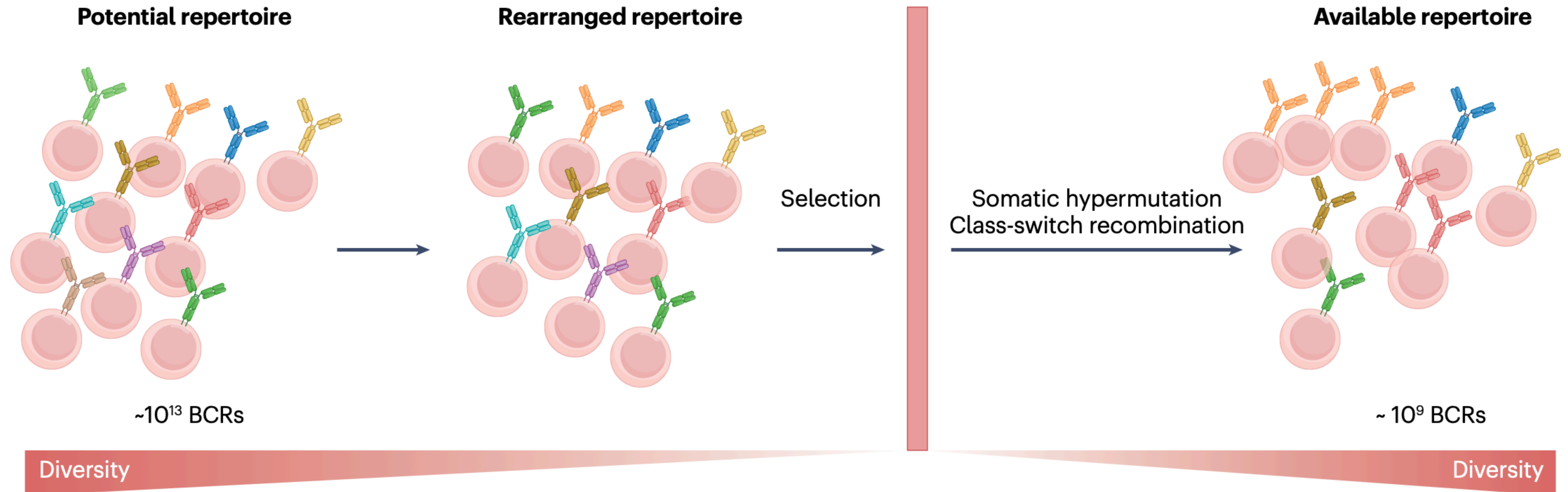# affinity maturation consists of iterative **mutation** and **selection**

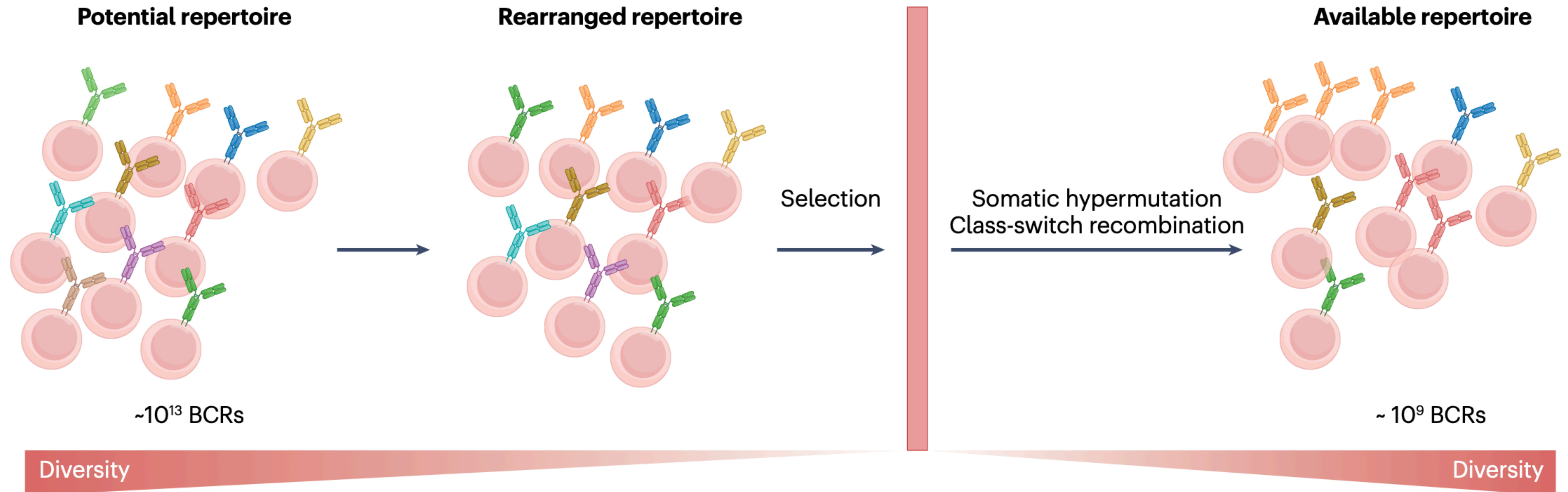# Together, these process generate diverse repertoires

Bone marrow

Potential repertoire

Rearranged repertoire

Selection

Periphery

Somatic hypermutation
Class-switch recombination

Available repertoire

~10^{13} BCRs

~10^{9} BCRs

Diversity

Diversity

Together, these process generate diverse repertoires

Bone marrow

Potential repertoire

Rearranged repertoire

Periphery

Available repertoire

Selection

Somatic hypermutation
Class-switch recombination

$\sim 10^{13}$ BCRs

$\sim 10^9$ BCRs

Diversity

Diversity

We can learn about **repertoire biases and dynamics** through **sequencing + statistics...**

We can sequence *mutated* BCRs…

mutated BCRs
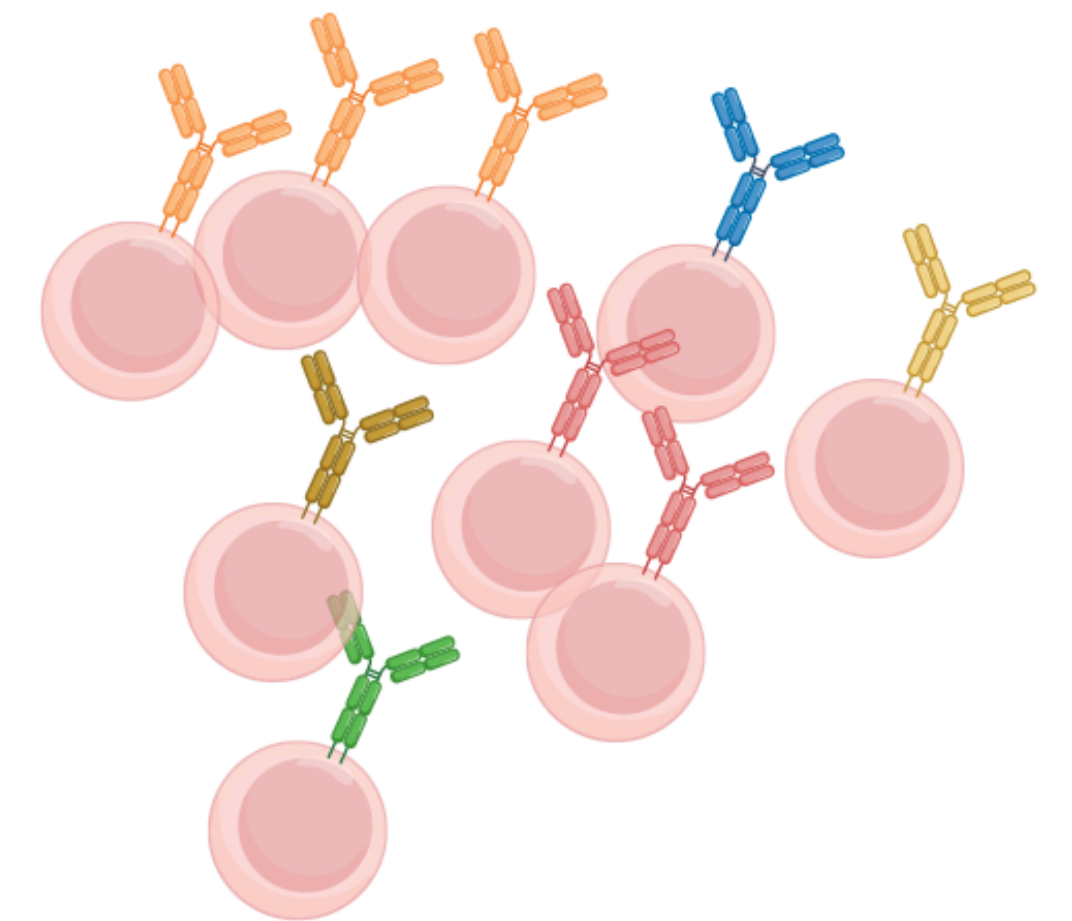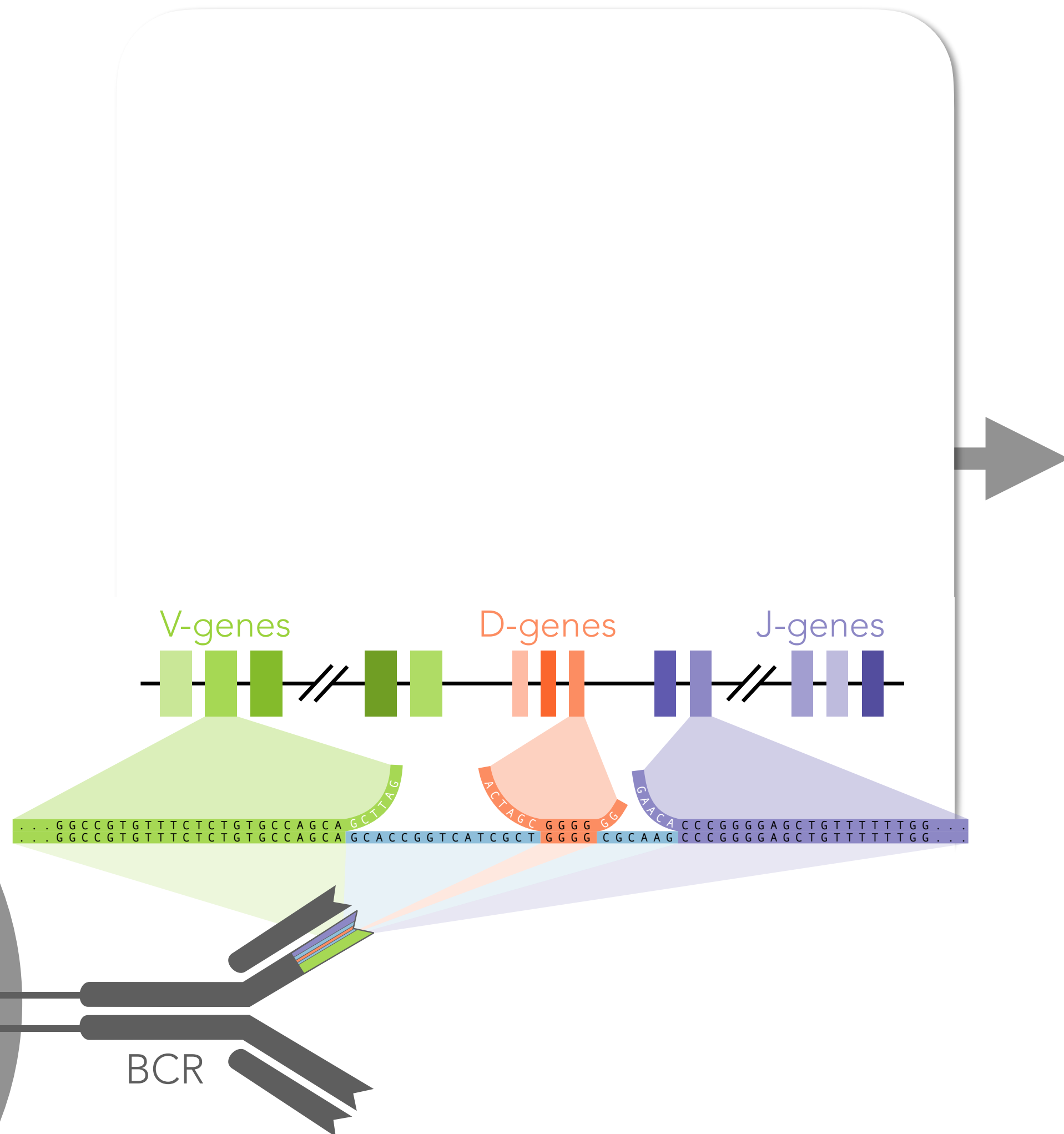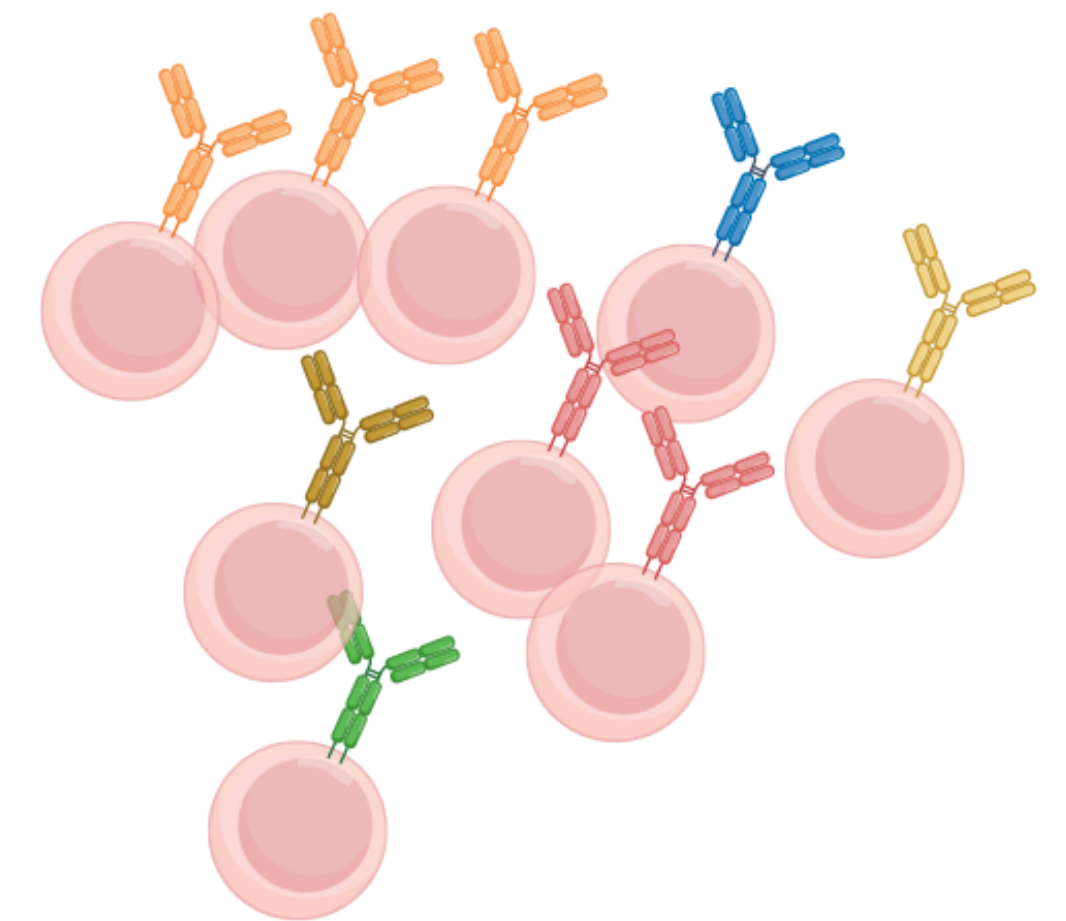
We can sequence *mutated* BCRs…

rearranged germline BCRs

mutated BCRs

We can sequence *mutated* BCRs…

rearranged germline BCRs
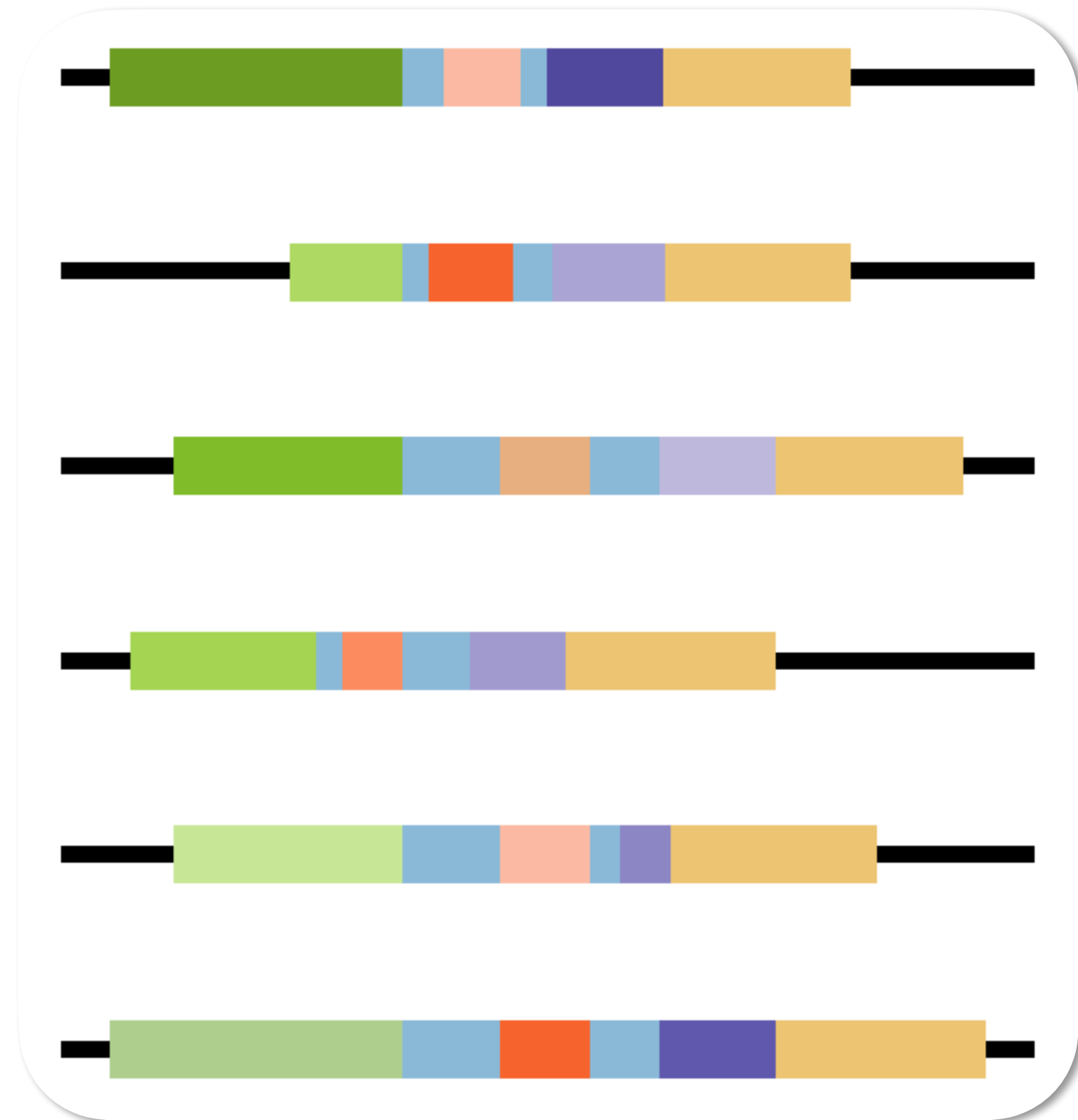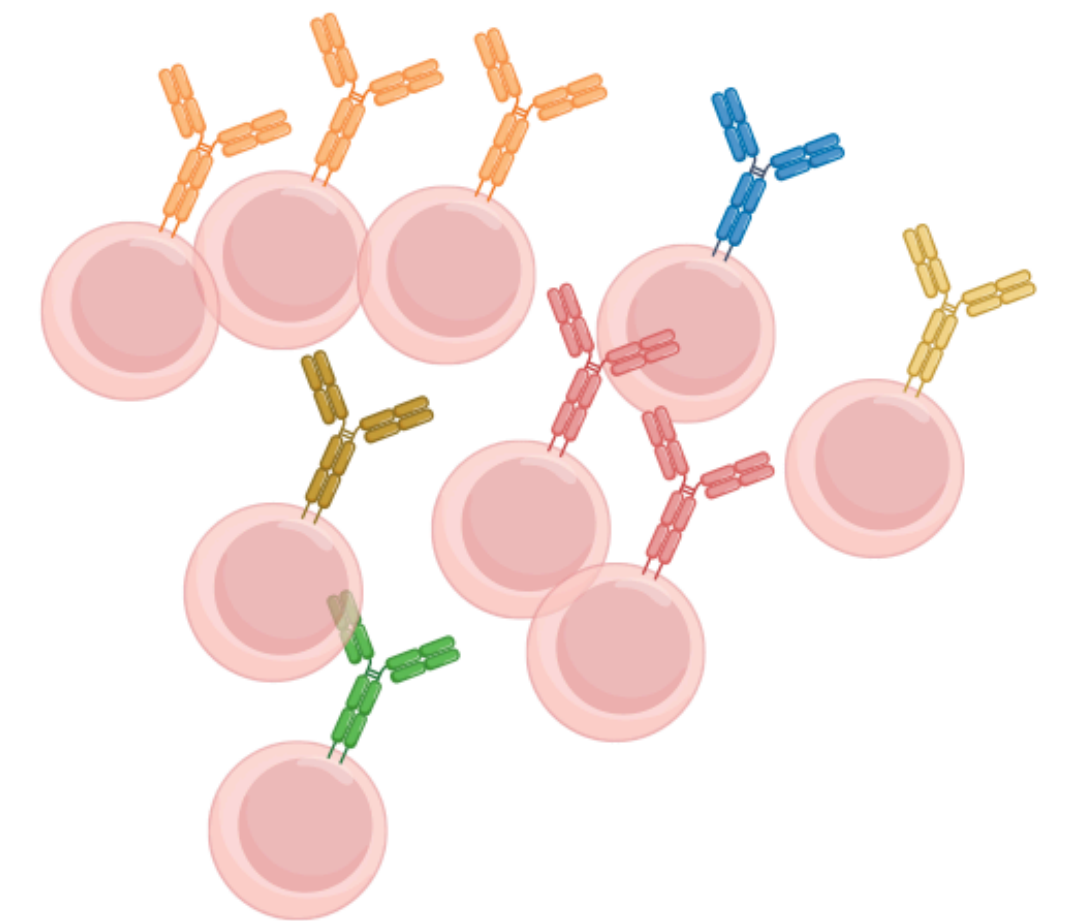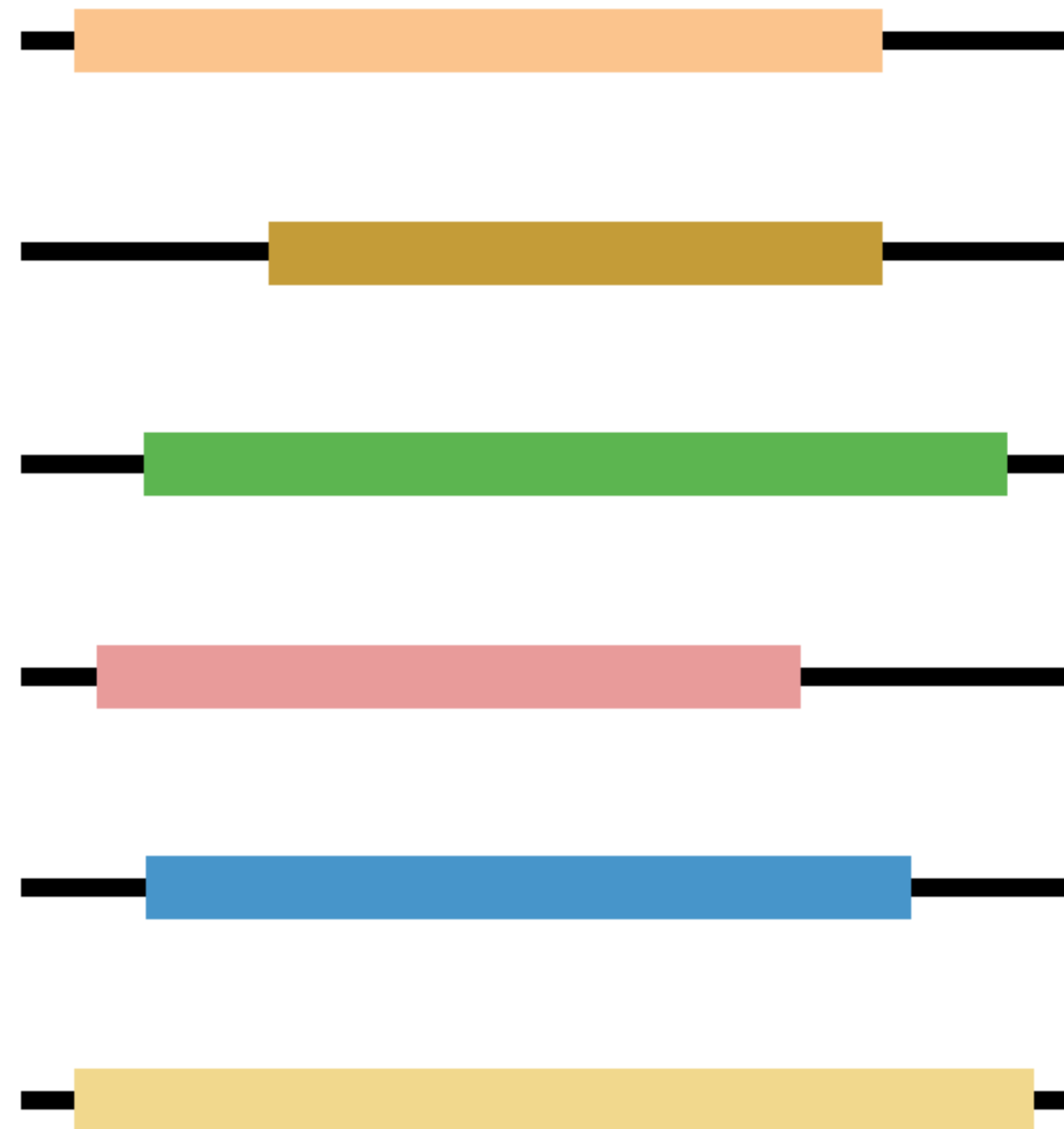
mutated BCRs

We can sequence *mutated* BCRs…

rearranged germline BCRs
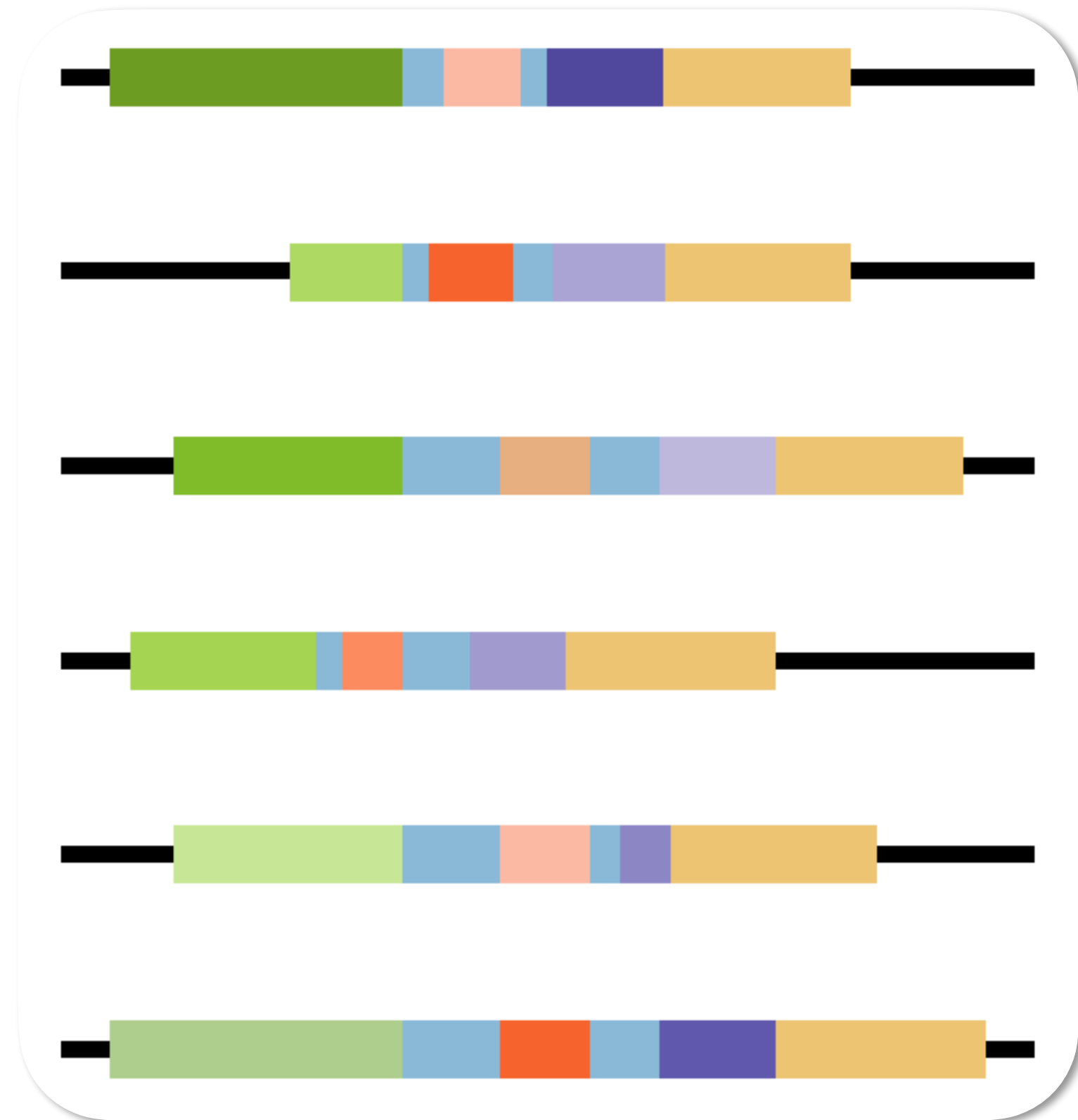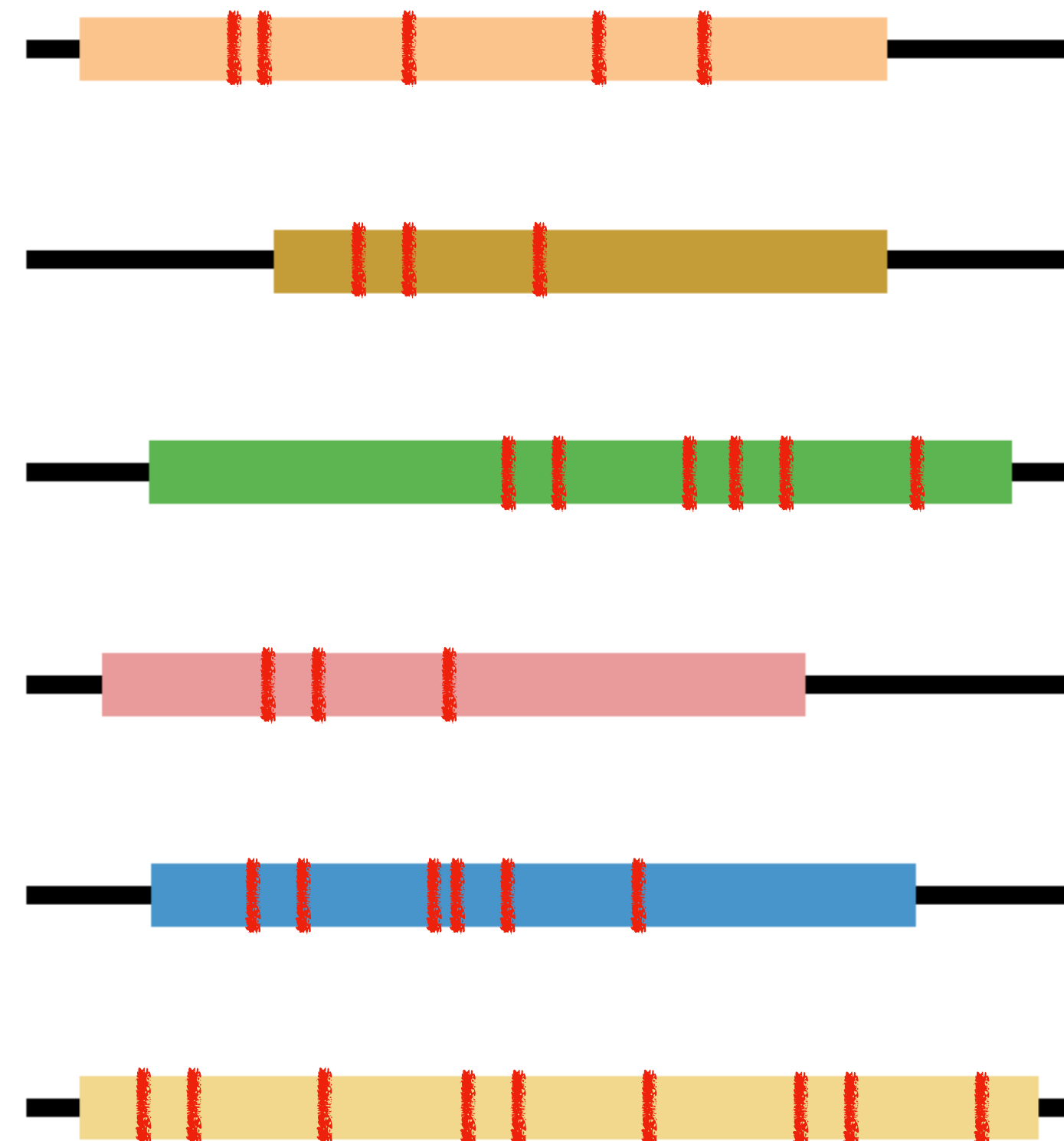
mutated BCRs

inferred mutations!
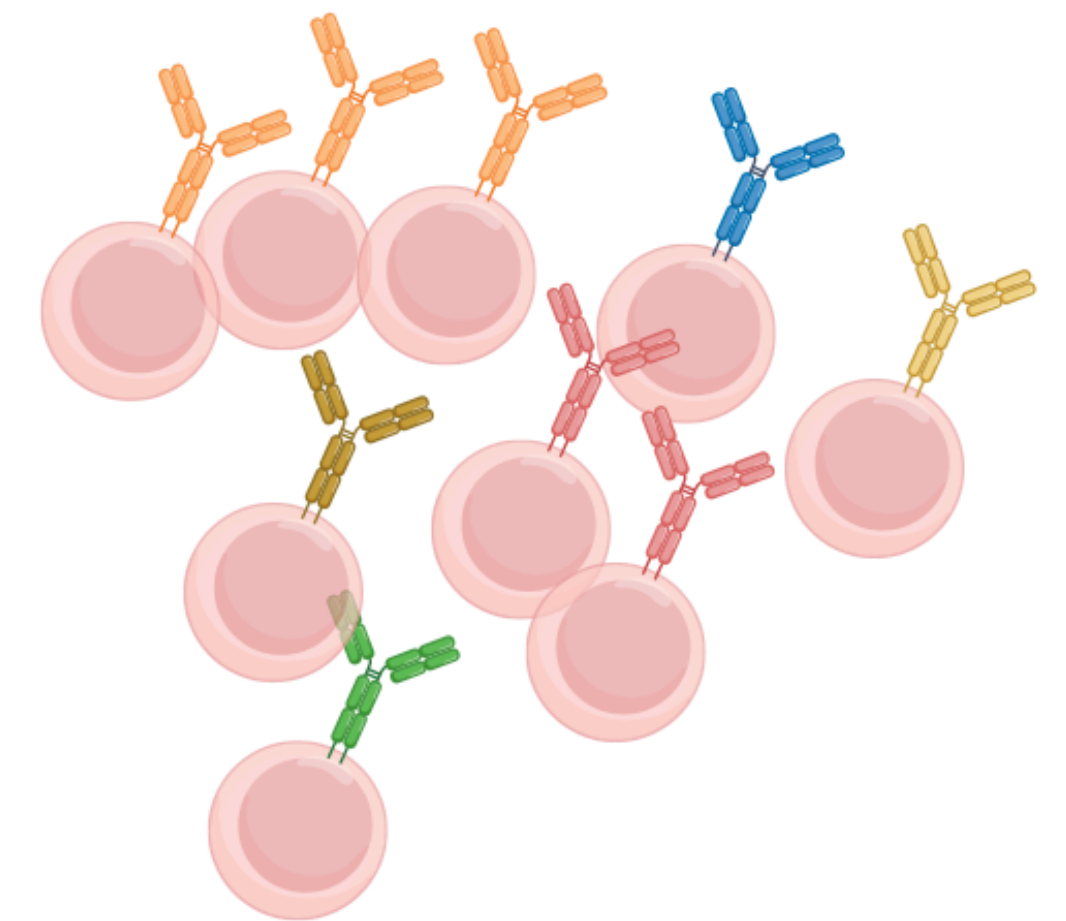
Key computational problems:

rearranged germline BCRs

mutated BCRs
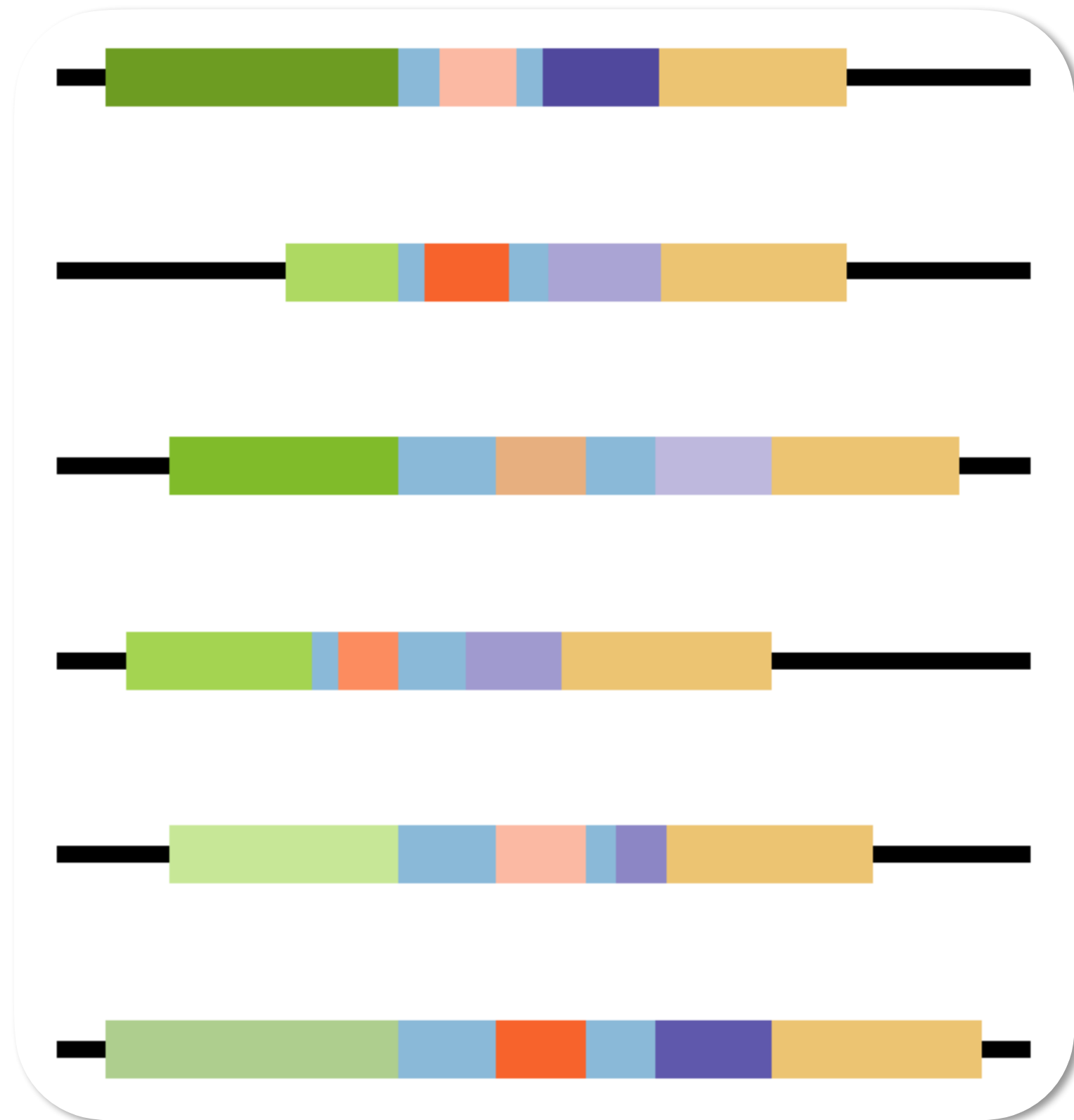
inferred mutations!
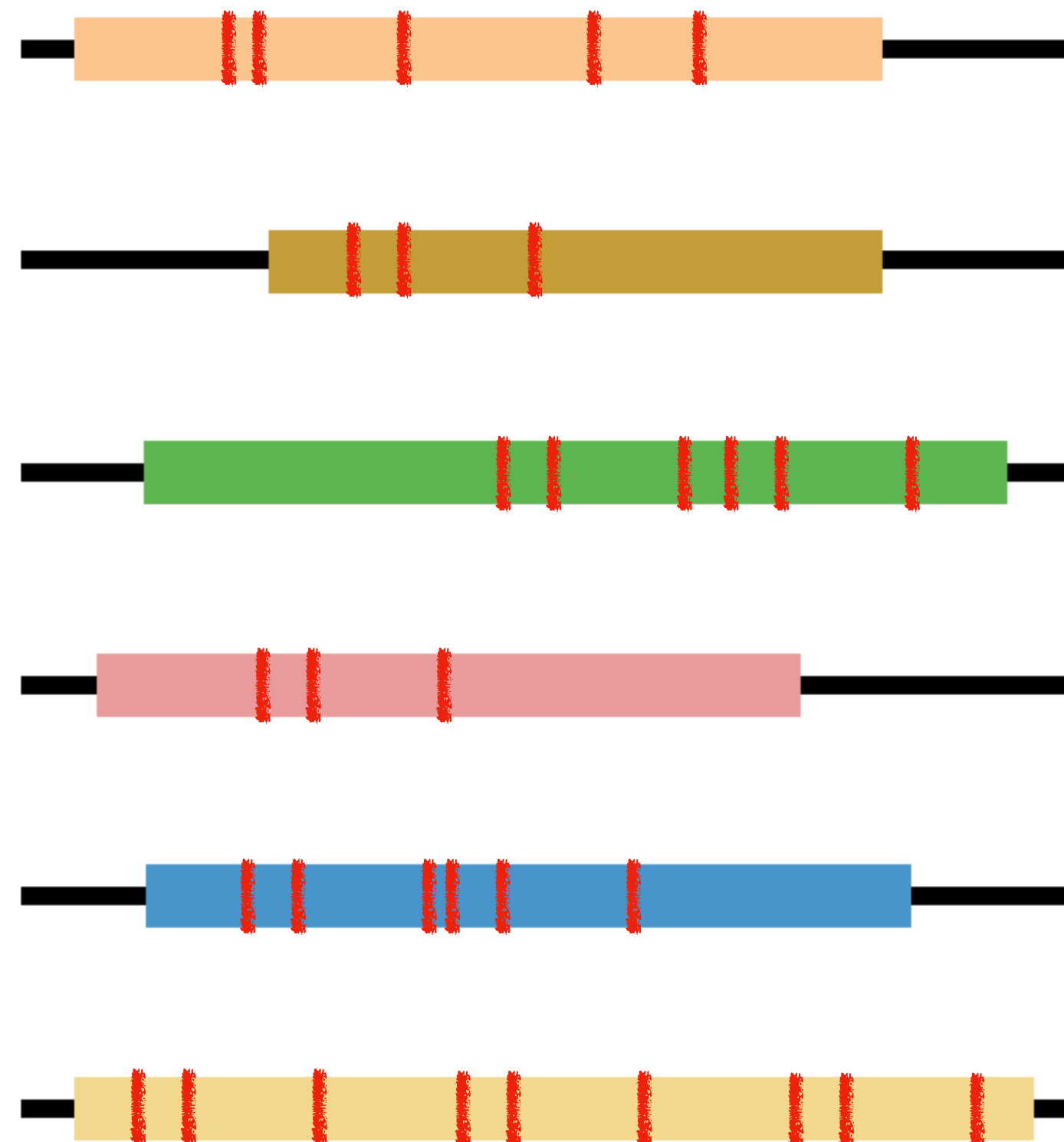
Given a sequence, how do we determine:

1. Which V, D, and J genes were used?

2. Where are the mutations?

3. Which mutations are synonymous/ nonsynonymous?

4. Which are in hotspot motifs?

This annotation is difficult!

- Many similar germline genes

- Mutation introduces additional mismatches

- Junctions (random gene deletions and insertions) disrupt alignment

- Need robust, specialized tools…

ANARCI

IMGT/V-quest
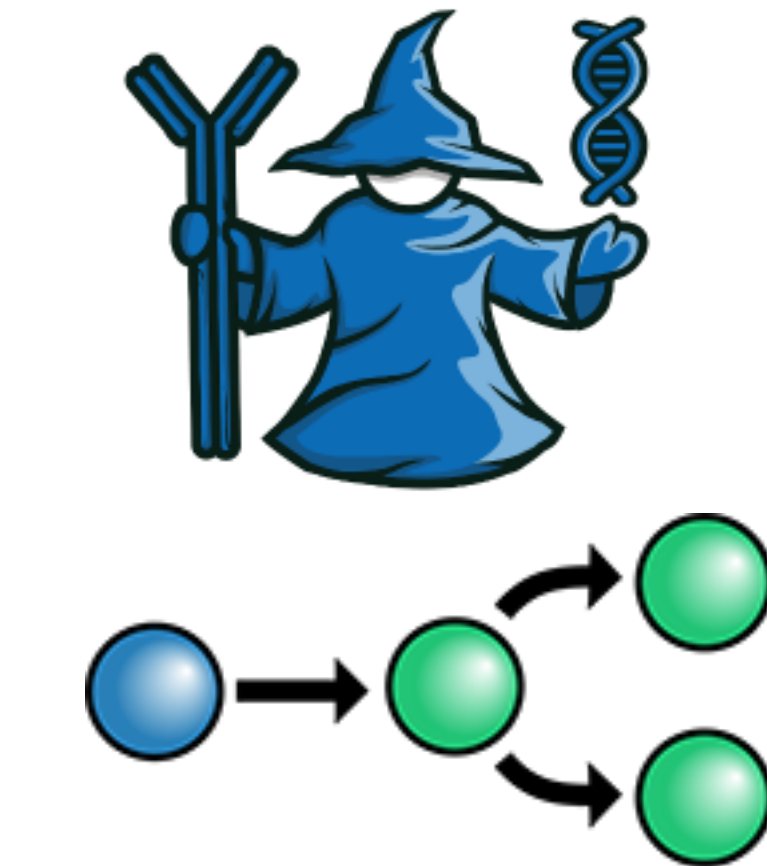
Immcantation / Change-o

MiXCR

NCBI IgBLAST

partis

RIOT

Immune repertoire analyses often focus on diversity, architecture, evolution, or convergence
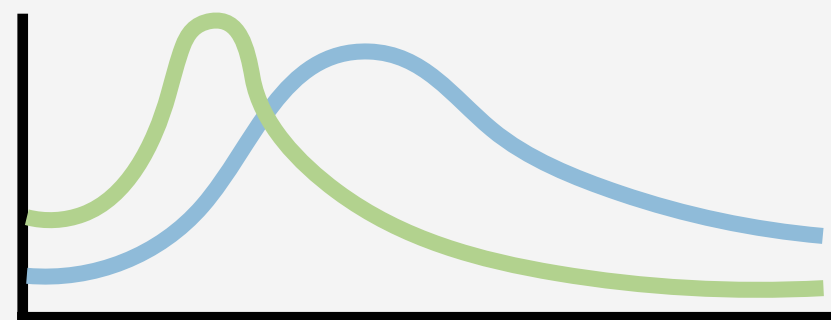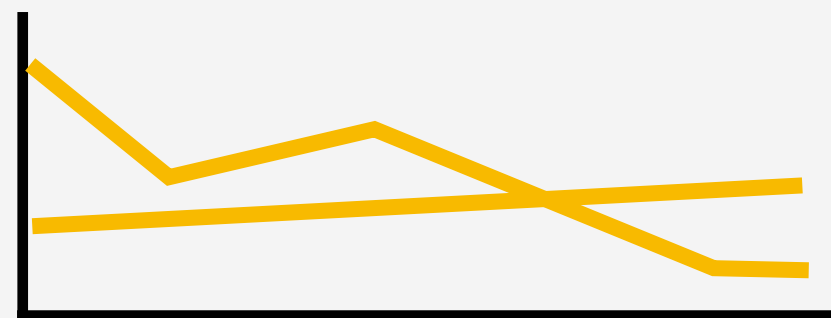
Immune repertoire analyses often focus on **diversity**, architecture, evolution, or convergence
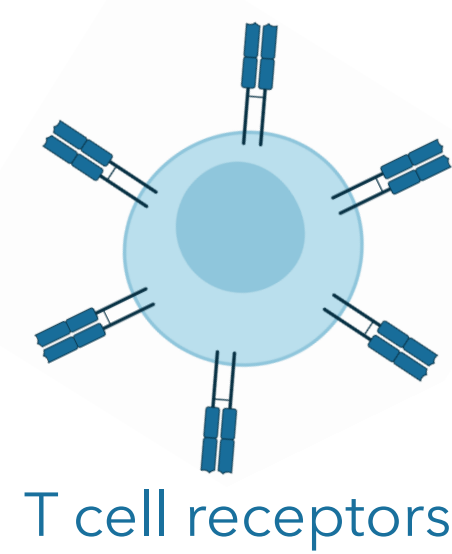
Immune repertoire analyses often focus on diversity, **architecture**, evolution, or convergence
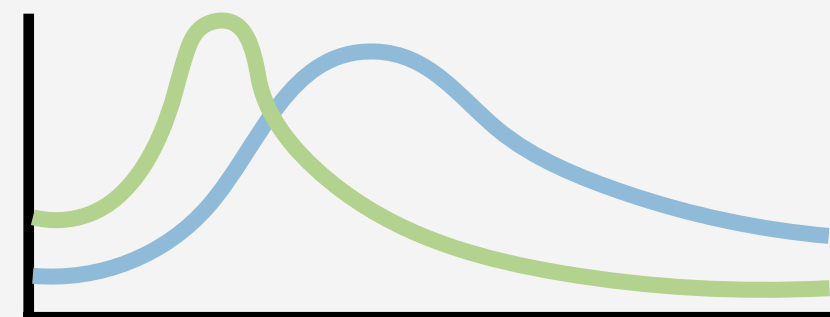

Repertoire architecture
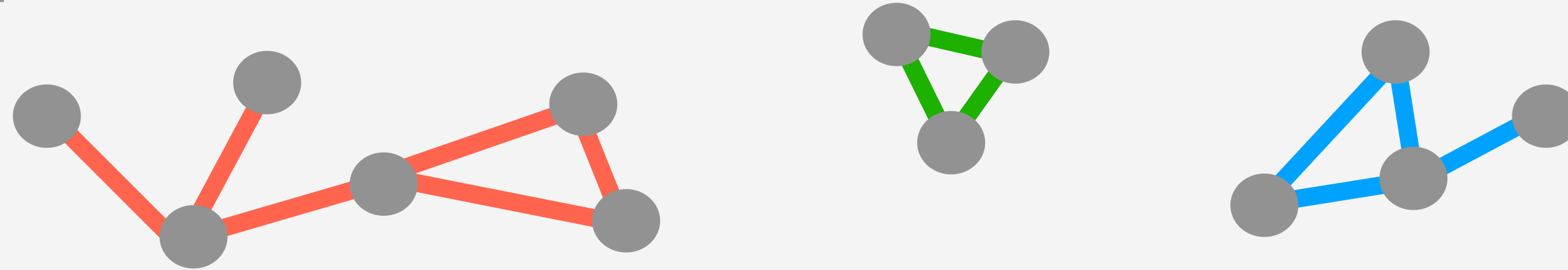
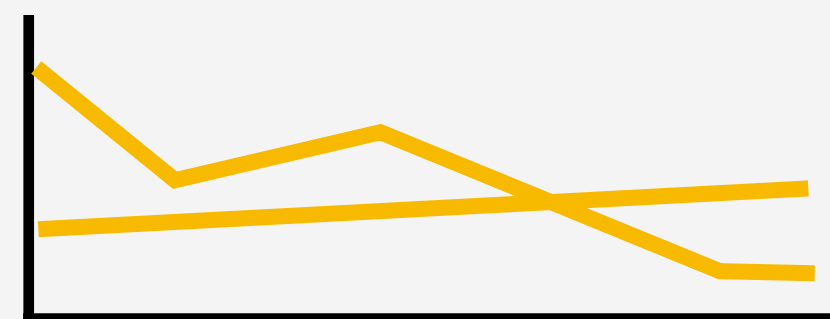defining antigen recognition breadth using network analysis
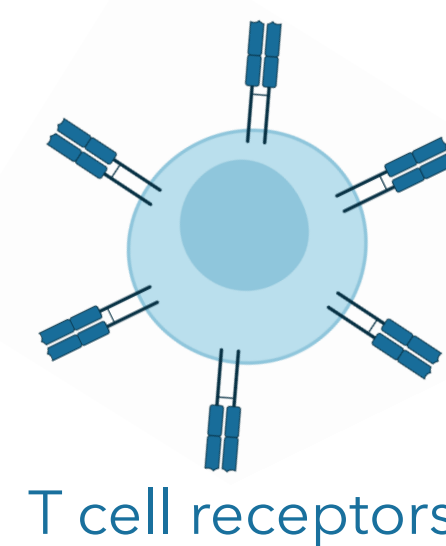
Underlying mechanisms of diversity generation

probabilistic sequence annotation

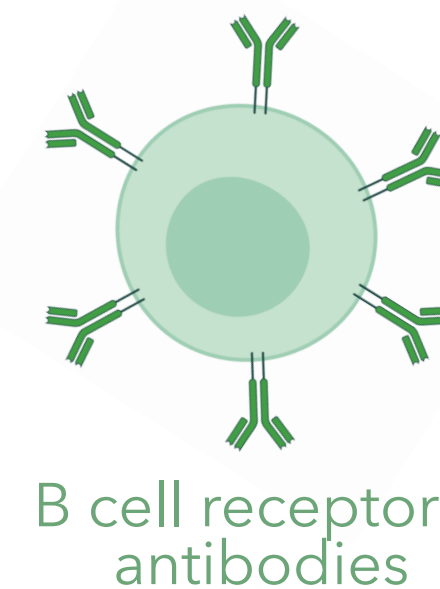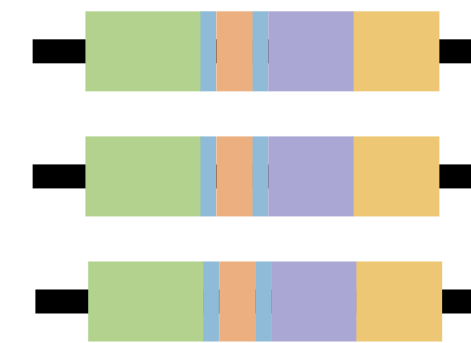recombination statistics to learn about generation and selection
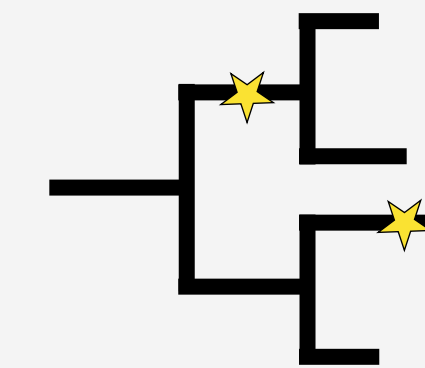
clonotype diversity dynamics

Immune repertoire data

T cell receptors

B cell receptors/ antibodies

Immune repertoire analyses often focus on diversity, architecture, **evolution**, or convergence

Repertoire architecture

defining antigen recognition breadth using network analysis
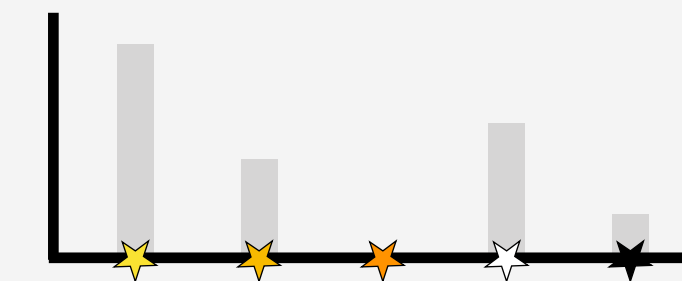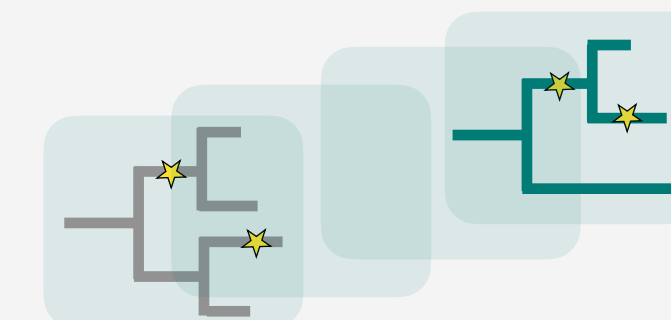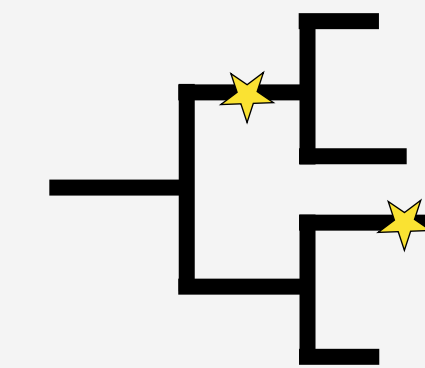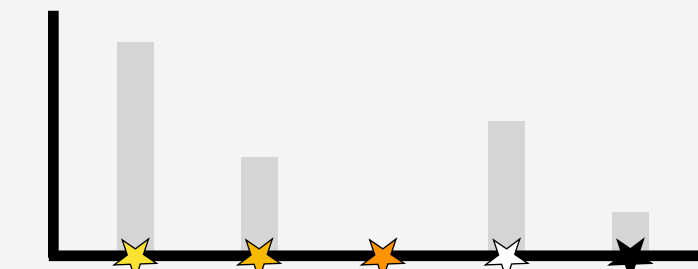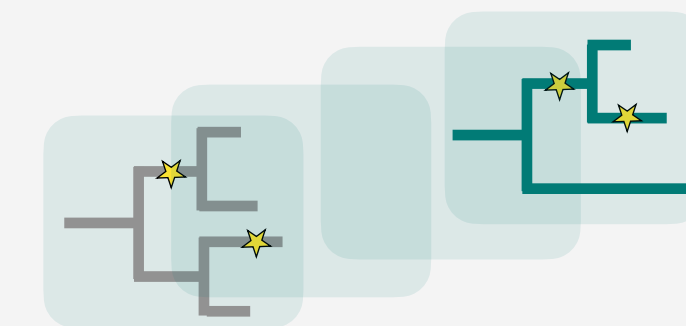
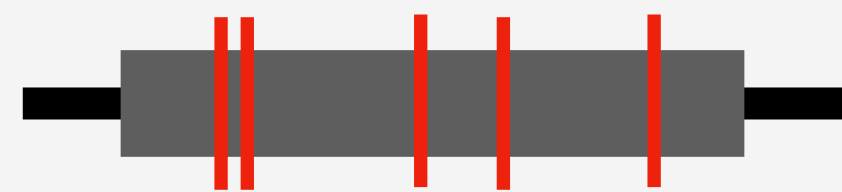**Antibody evolution**

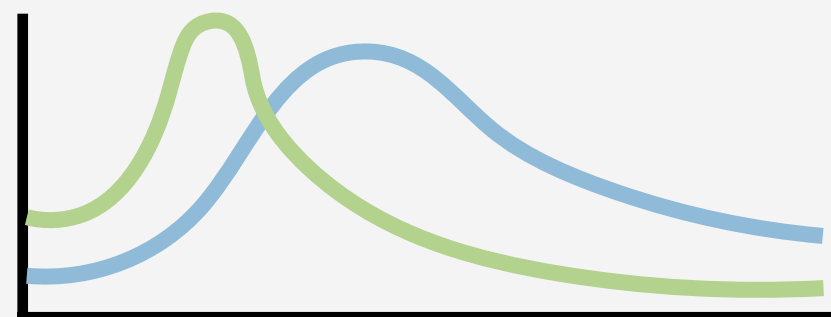reconstruction of phylogenetic trees

inferring mutation statistics
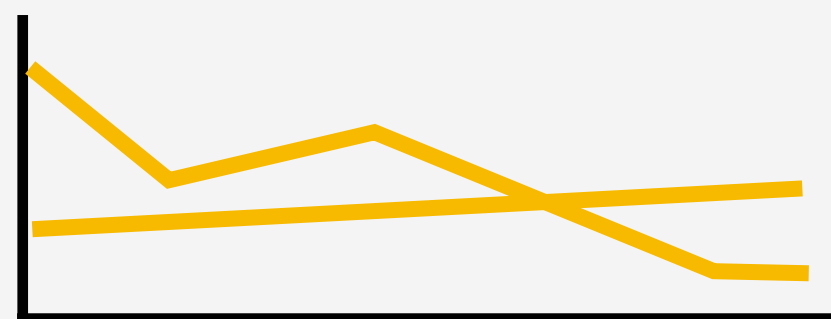
simulating antibody repertoire evolution
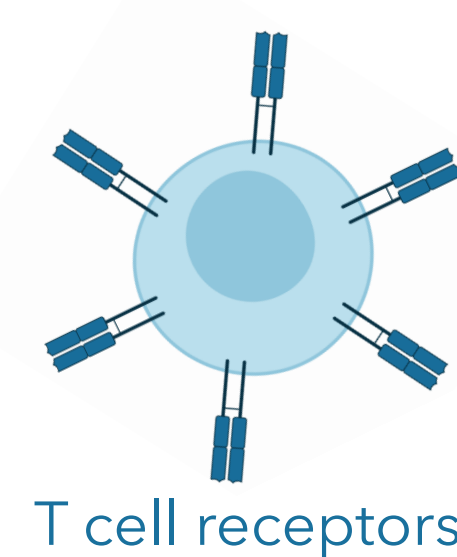
Underlying mechanisms of diversity generation

probabilistic sequence annotation

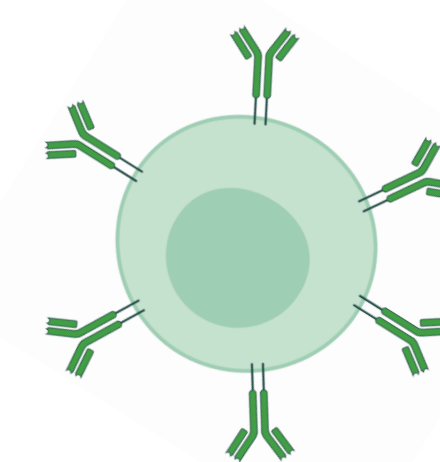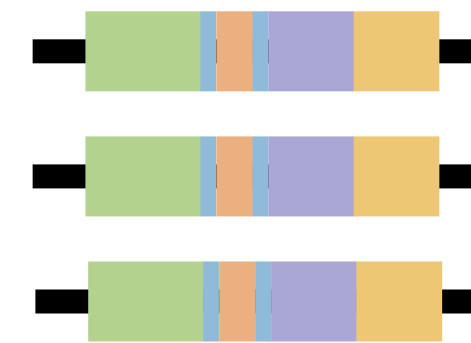recombination statistics to learn about generation and selection

clonotype diversity dynamics

Immune repertoire data
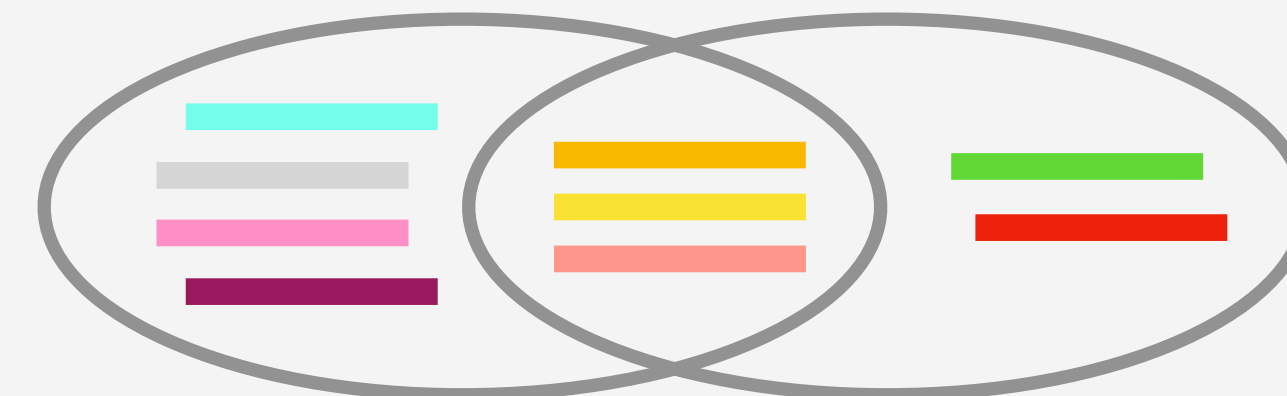
T cell receptors

B cell receptors/ antibodies

Immune repertoire analyses often focus on
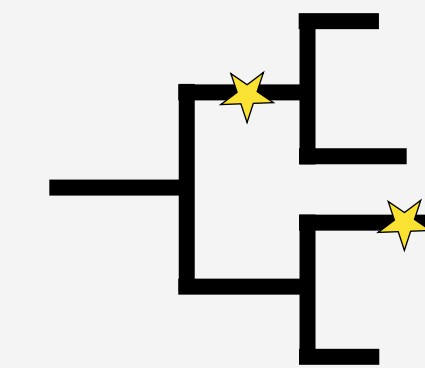# diversity, architecture, evolution, or **convergence**

Immune repertoire analyses often focus on diversity, architecture, evolution, or **convergence**
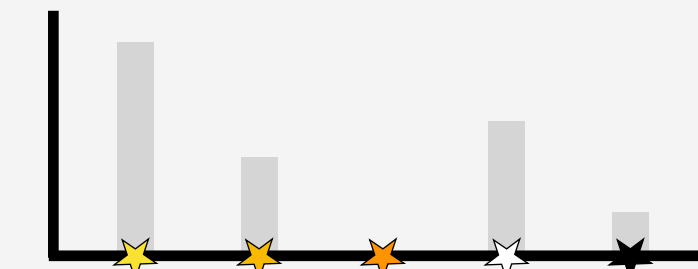

Repertoire architecture

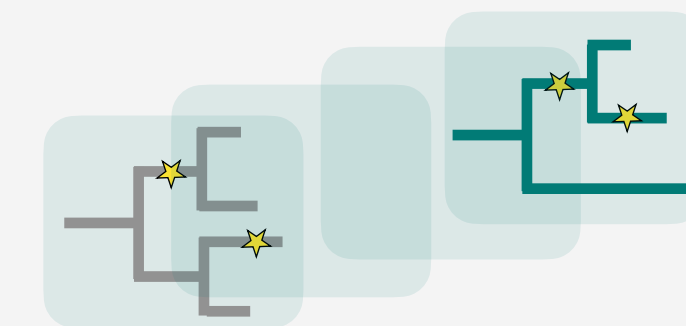defining antigen recognition breadth using network analysis

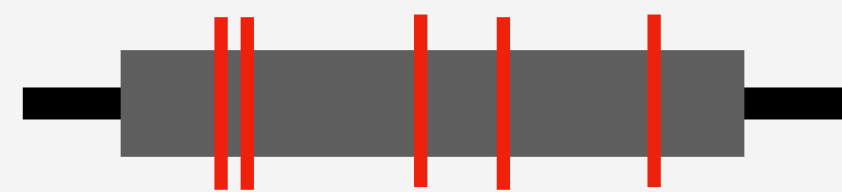**Antibody evolution**

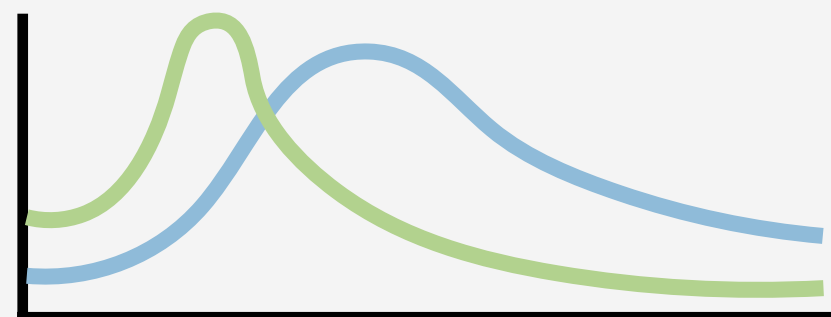reconstruction of phylogenetic trees

inferring mutation statistics
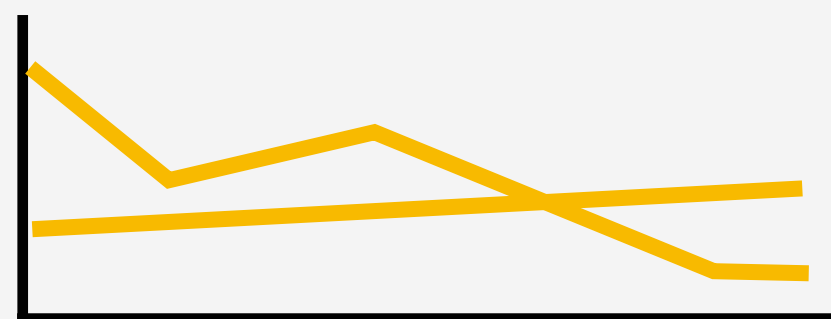
simulating antibody repertoire evolution

Underlying mechanisms of diversity generation
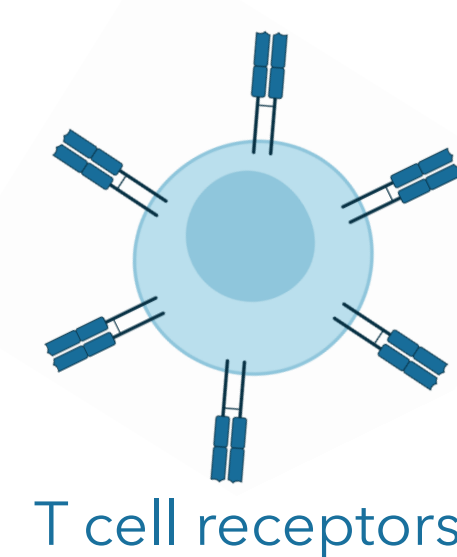
probabilistic sequence annotation

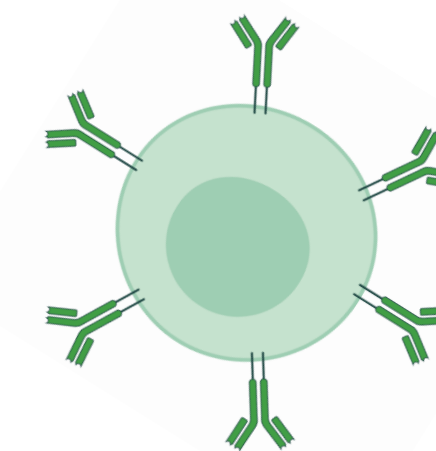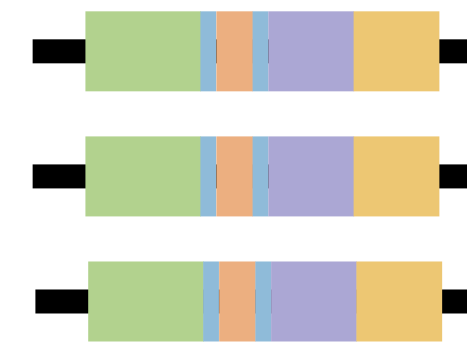recombination statistics to learn about generation and selection

clonotype diversity dynamics

Immune repertoire data

T cell receptors

B cell receptors/ antibodies

Molecular convergence

exploring cross-individual sequence similarity and convergence