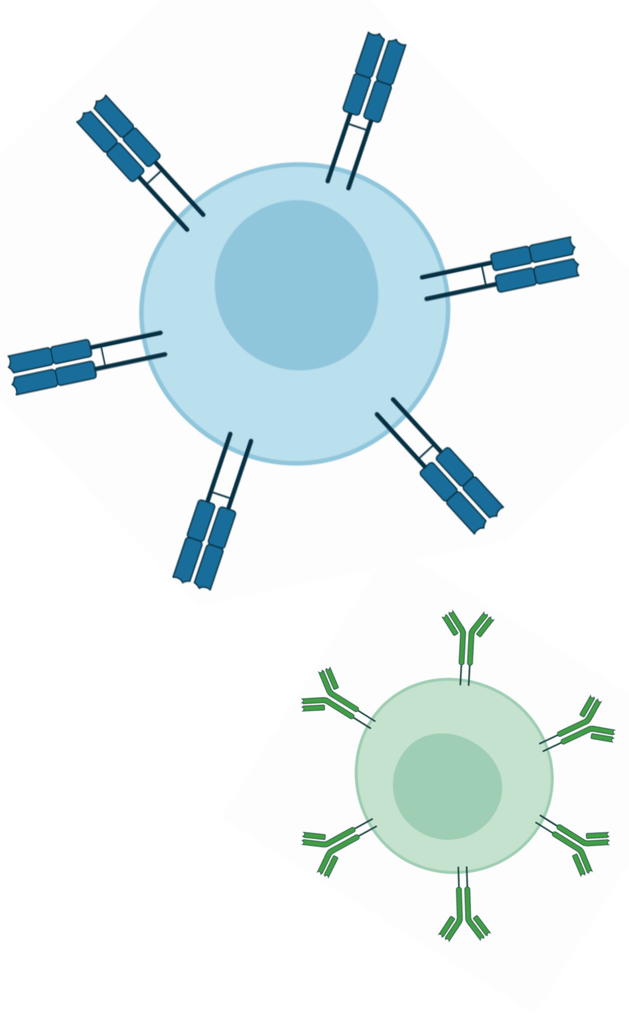
Intro to immune repertoire sequencing and analysis

Maggie Russell

TFCB 2024

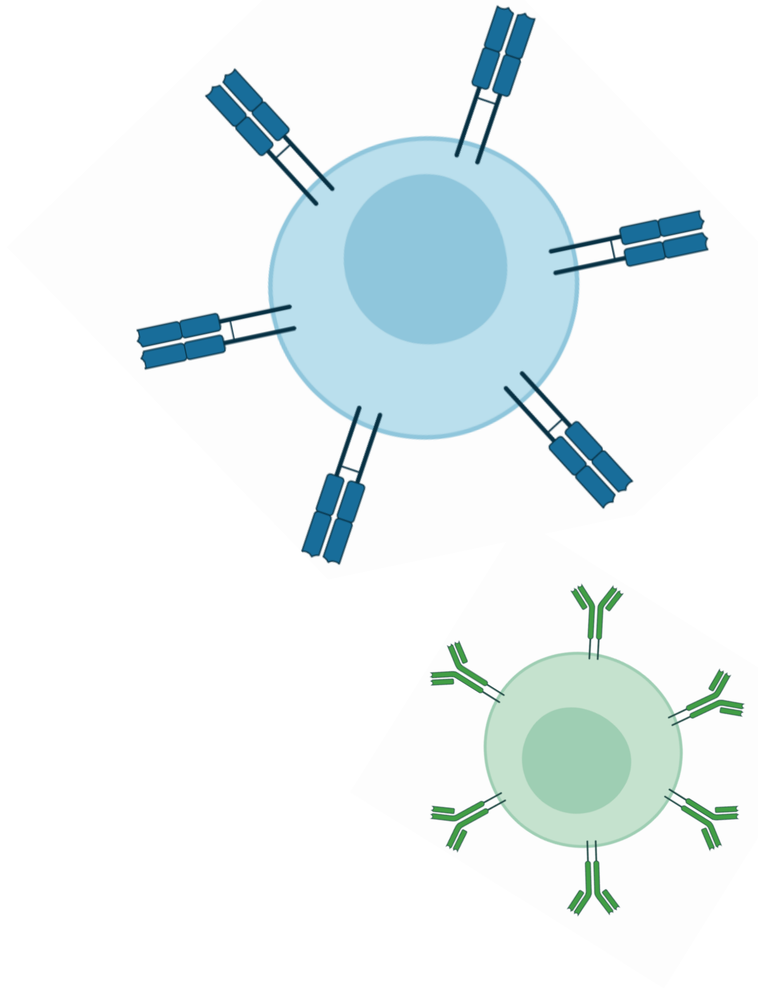
Lecture goals:

1. learn about immune repertoire sequencing
2. familiarize with immune repertoire data
3. work through an example analysis



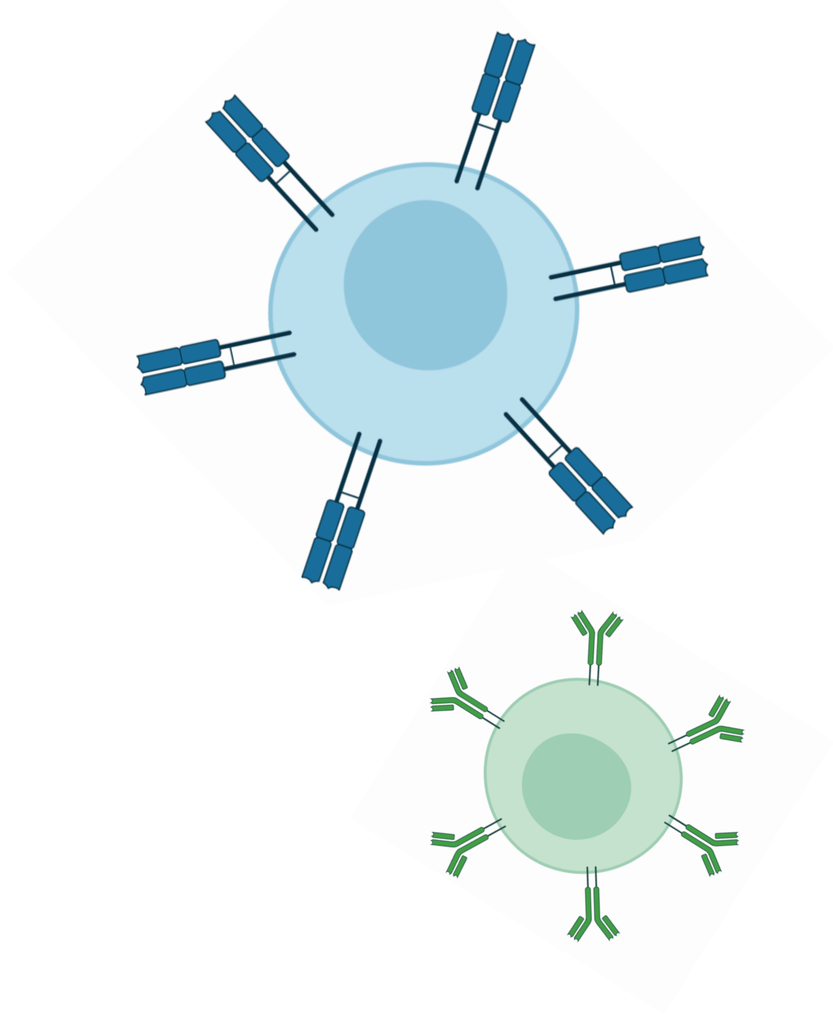
Lecture goals:

1. learn about immune repertoire sequencing

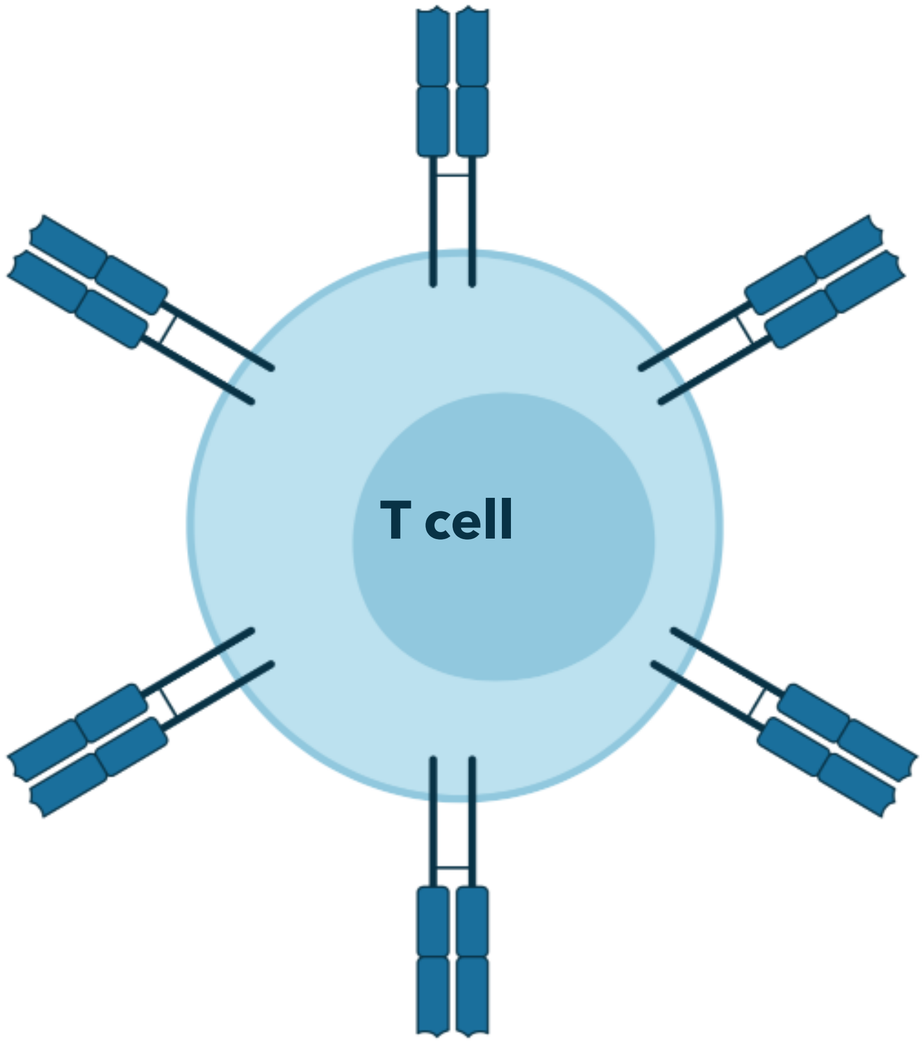


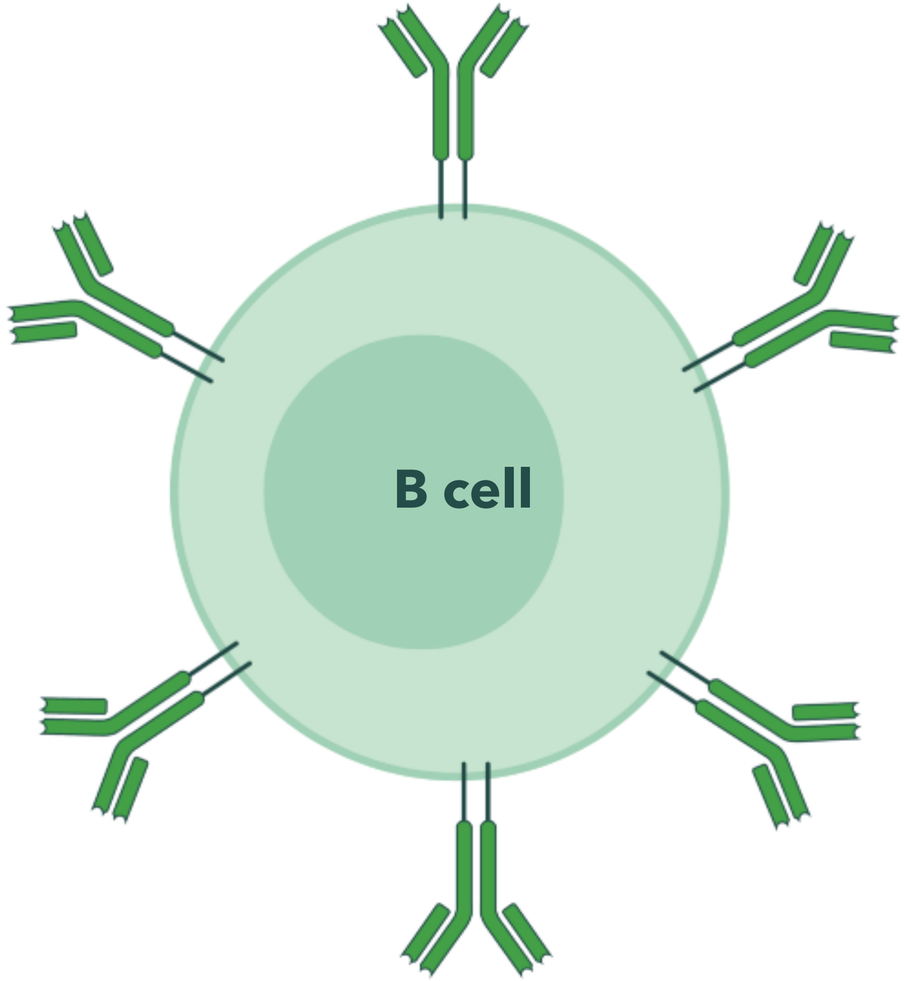
Lecture goals:

1. learn about immune repertoire sequencing
    - **what are immune repertoires?**

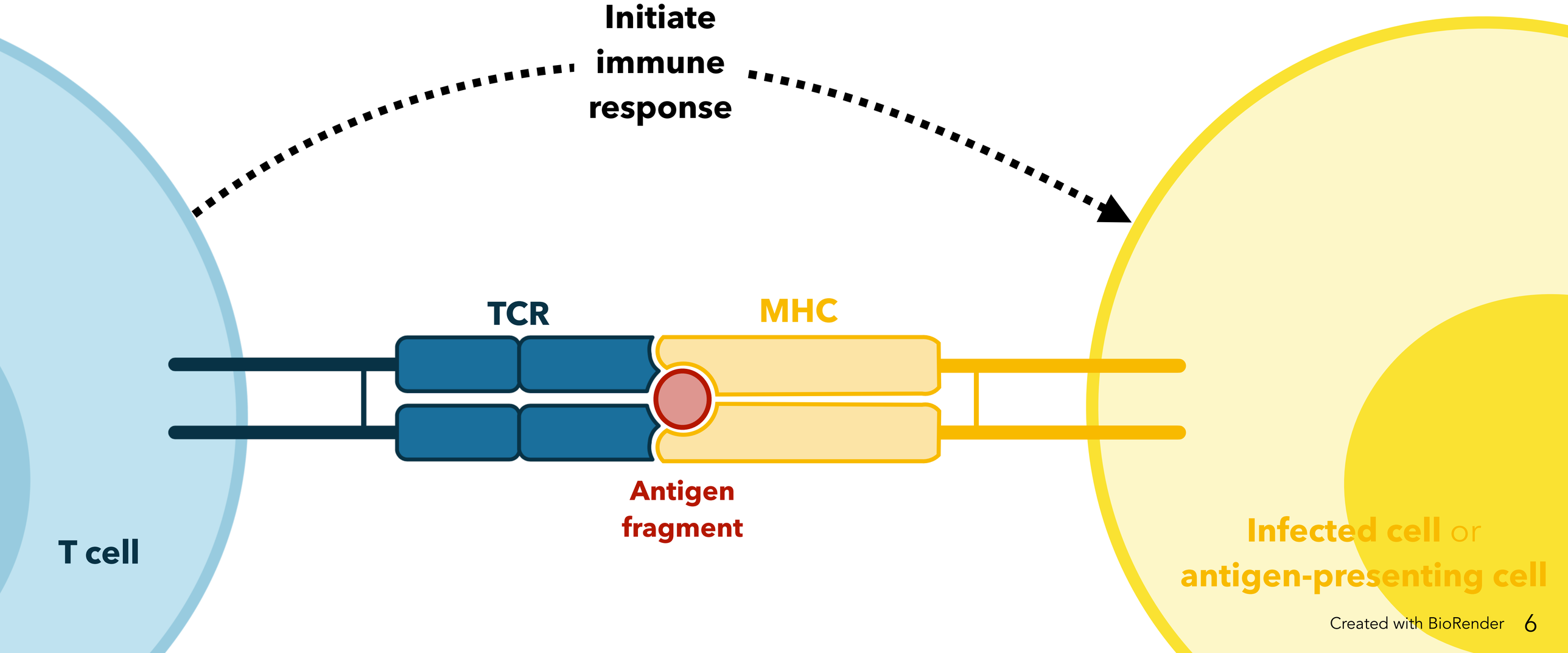


Adaptive immunity is essential for defending against pathogens

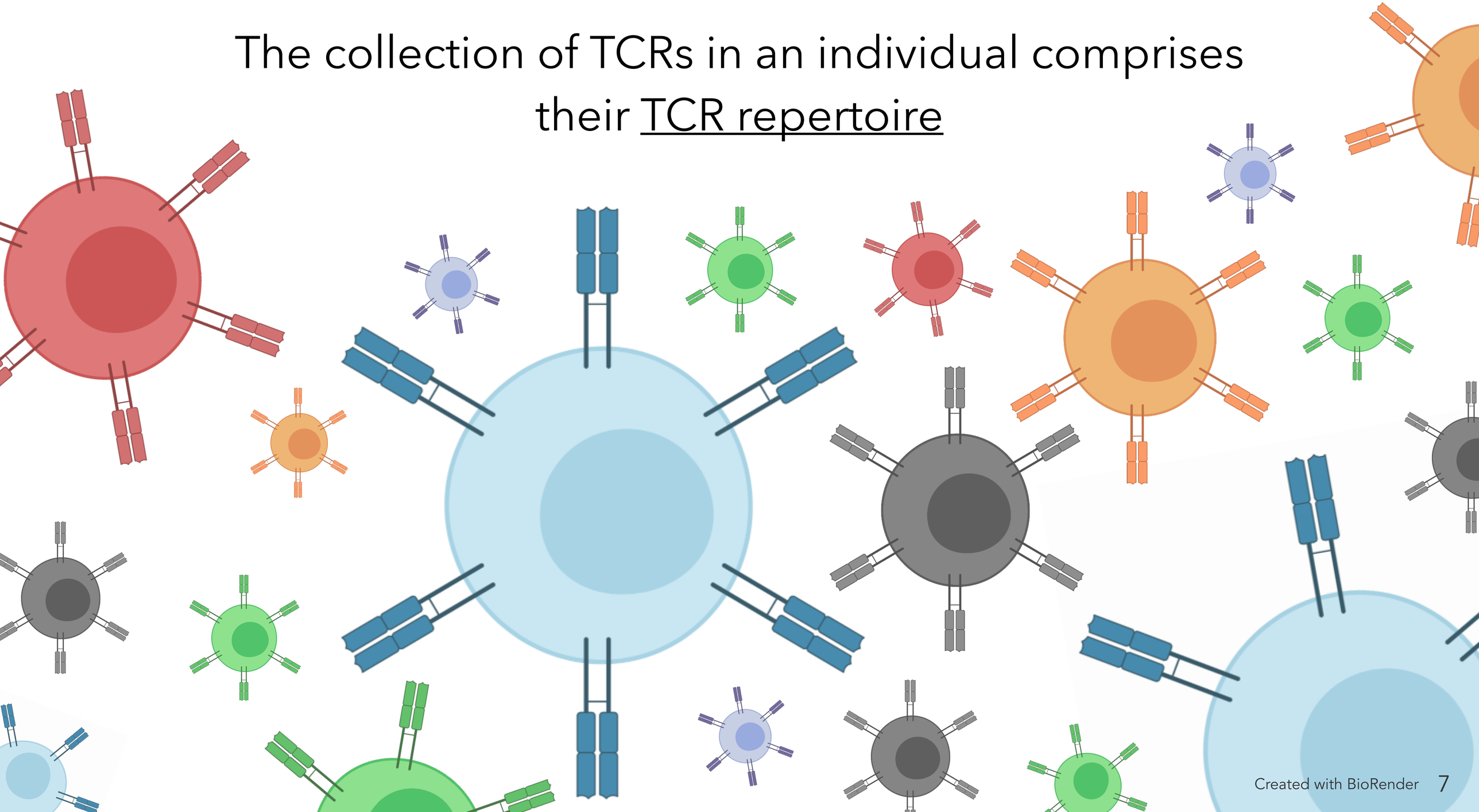




T cell receptors recognize antigen fragments bound to MHC

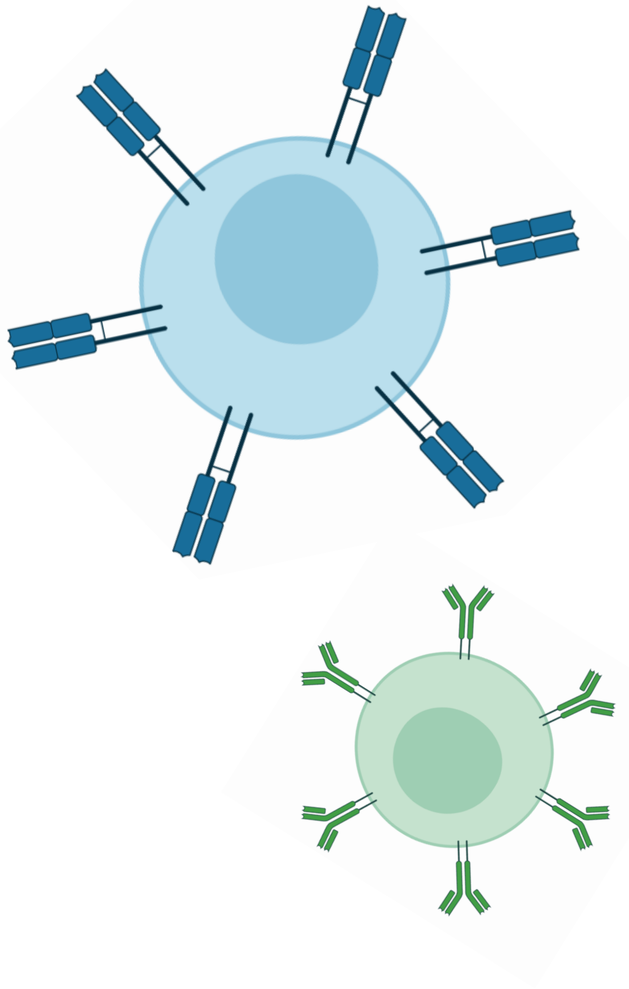


The collection of TCRs in an individual comprises their TCR repertoire



Lecture goals:

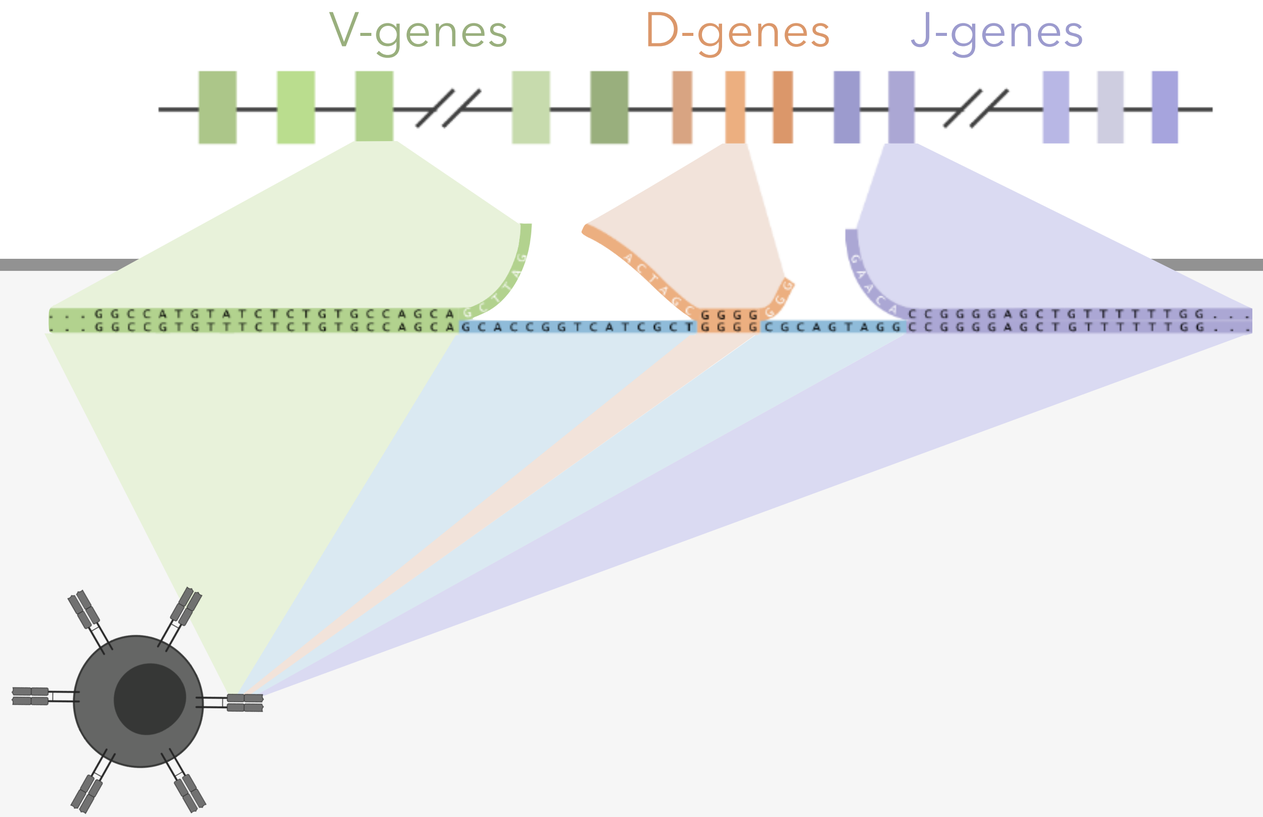
1. learn about immune repertoire sequencing
    - what are immune repertoires?
    - **how are they formed?**



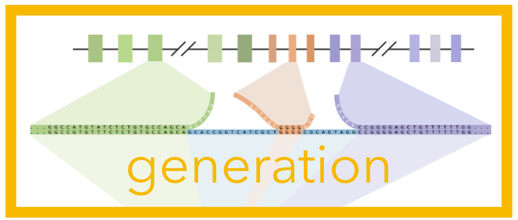
Repertoire composition is influenced by generation, selection, and exposures

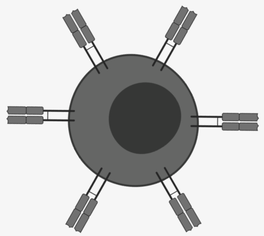
Let's use a water pipe as an analogy for TCR repertoire formation…

Repertoire composition is influenced by generation, selection, and exposures



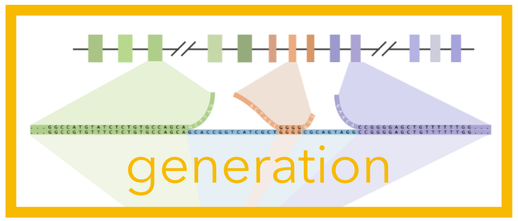
Repertoire composition is influenced by **generation**, selection, and exposures





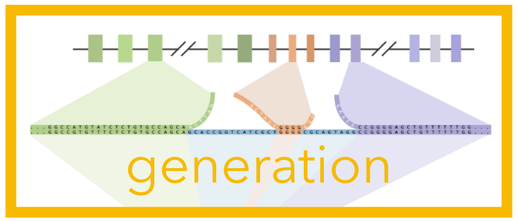
modified from Murugan et. al *PNAS* 2012
Created with BioRender

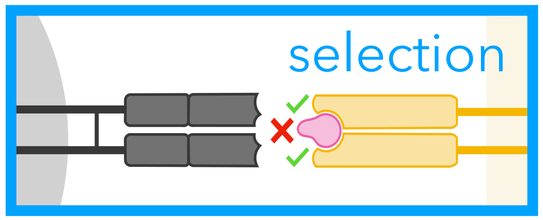
Repertoire composition is influenced by **generation**, selection, and exposures

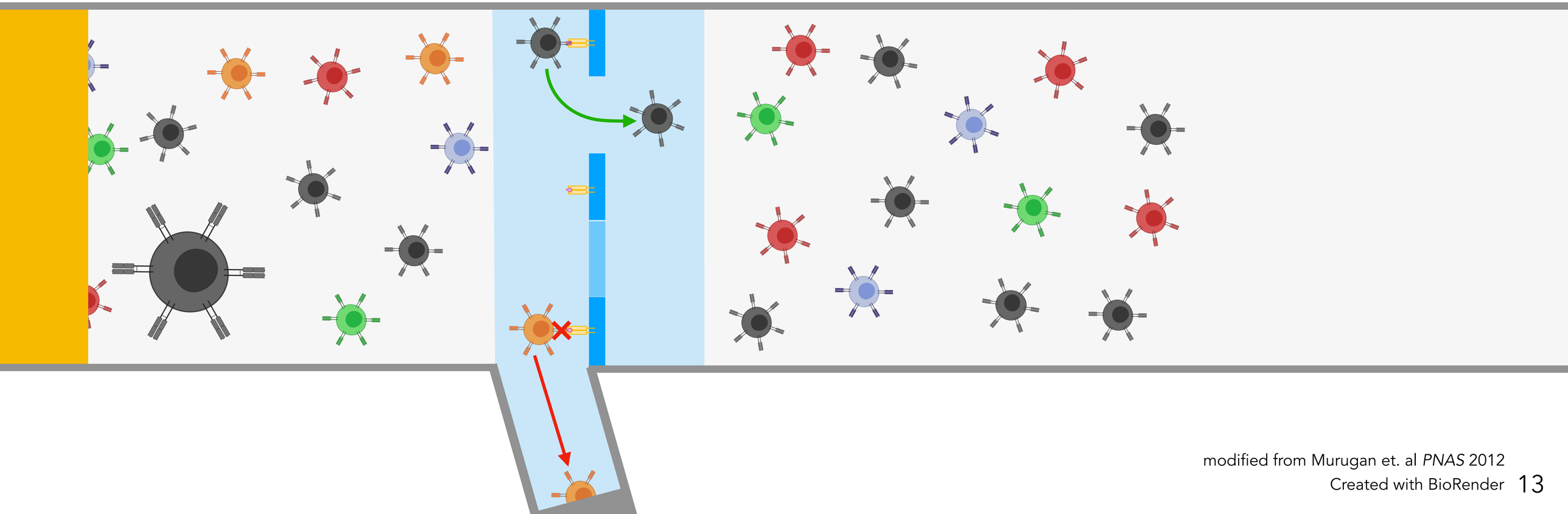


modified from Murugan et. al *PNAS* 2012
Created with BioRender

Repertoire composition is influenced by generation, **selection**, and exposures



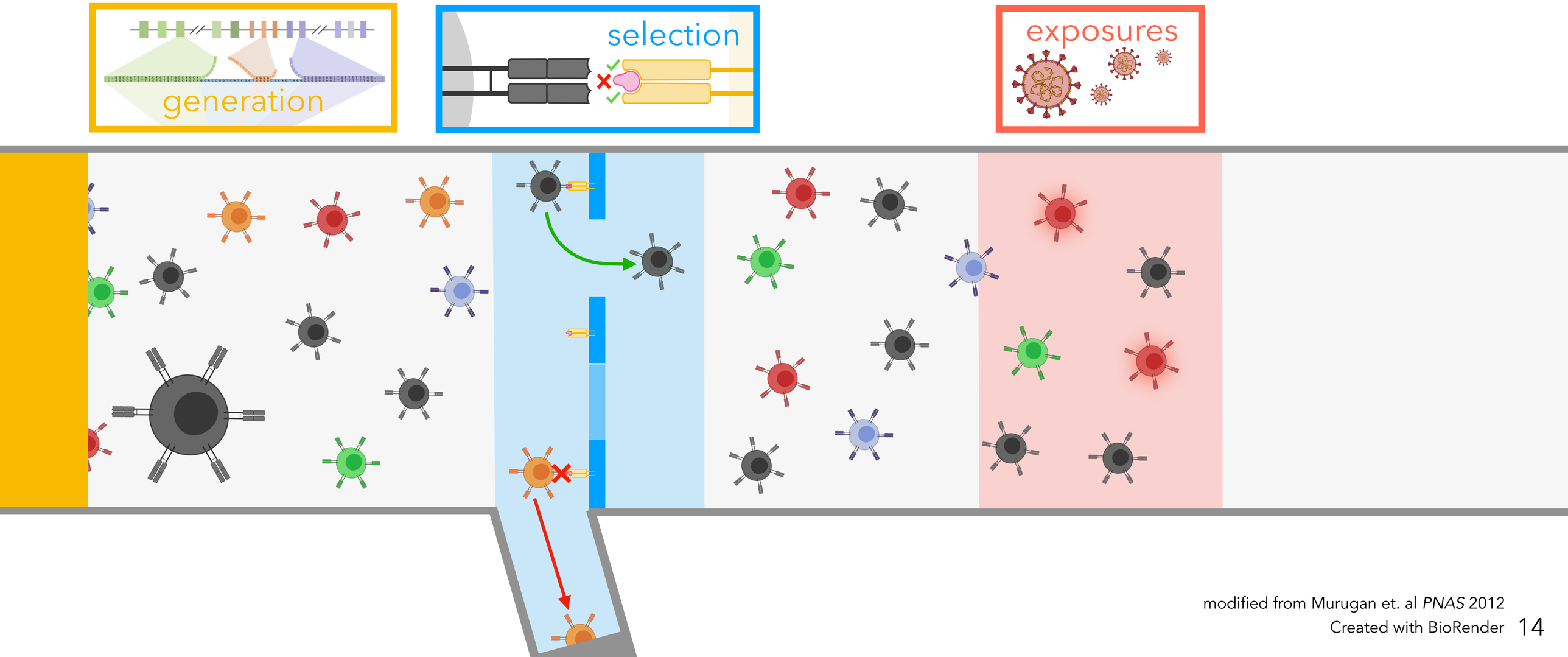




modified from Murugan et. al *PNAS* 2012

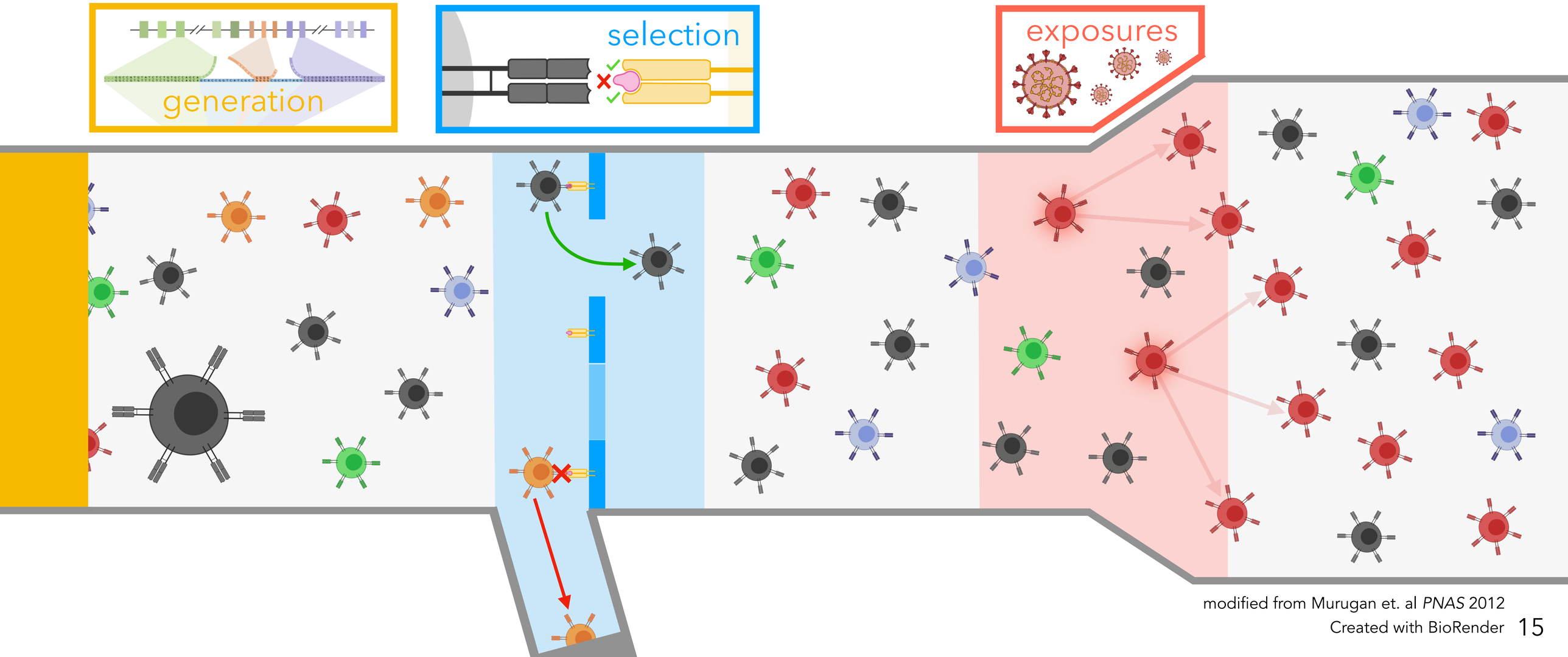
Created with BioRender

Repertoire composition is influenced by generation, selection, and **exposures**



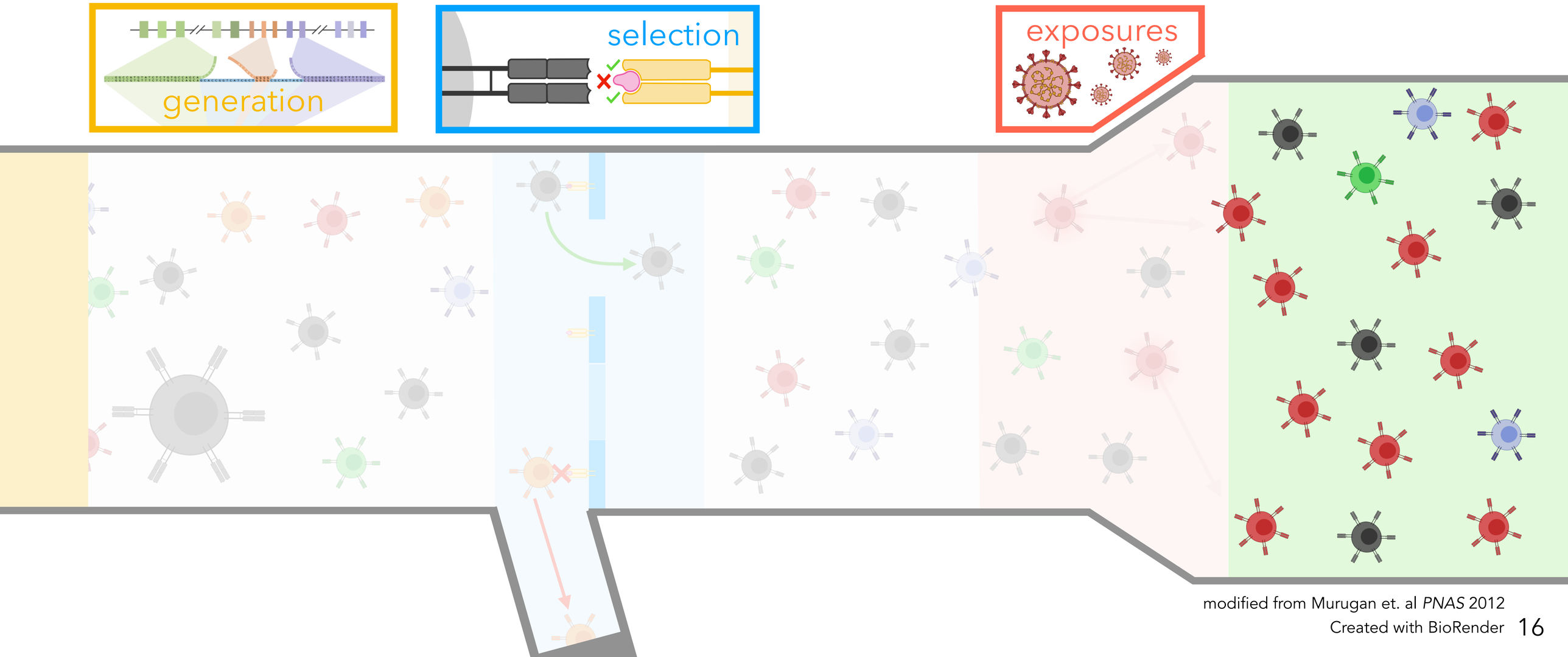
modified from Murugan et. al *PNAS* 2012
Created with BioRender

Repertoire composition is influenced by generation, selection, and **exposures**



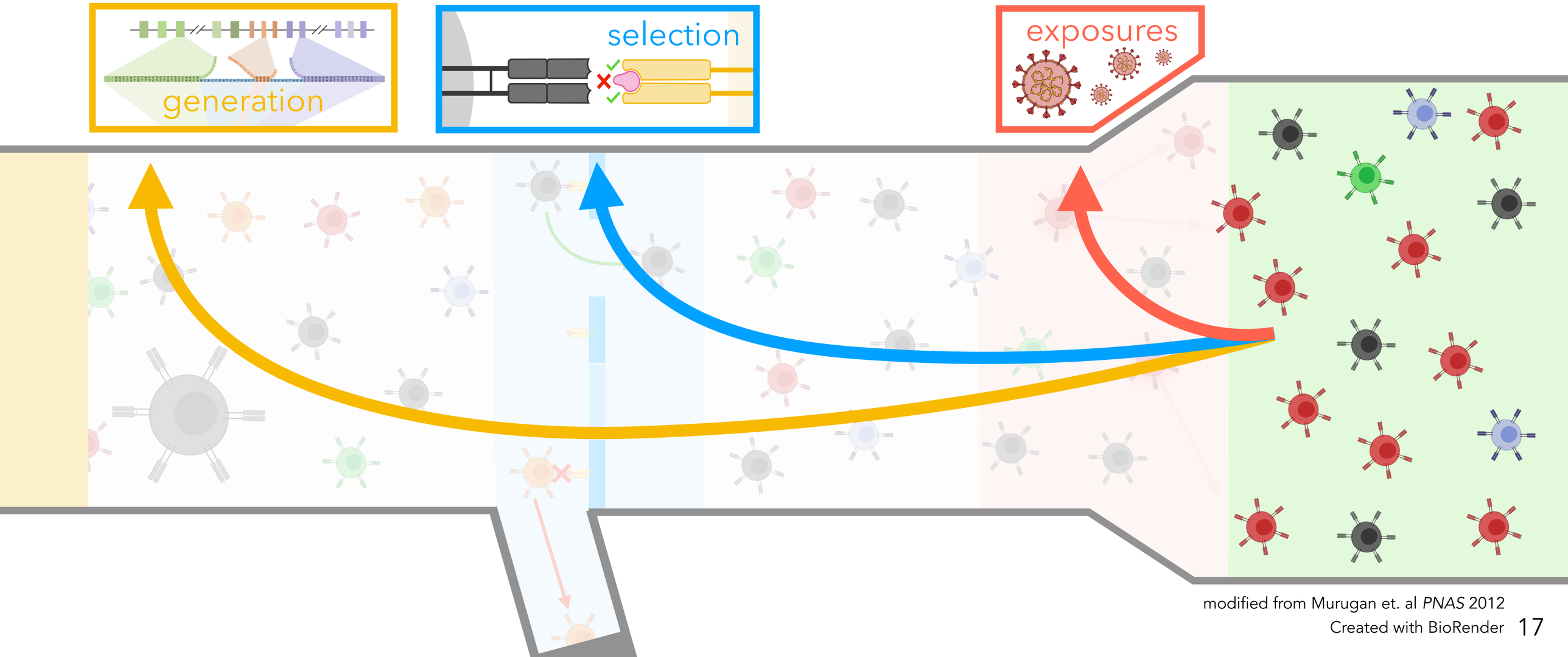
modified from Murugan et. al *PNAS* 2012
Created with BioRender

We can sample a **repertoire** using sequencing



modified from Murugan et. al *PNAS* 2012
Created with BioRender

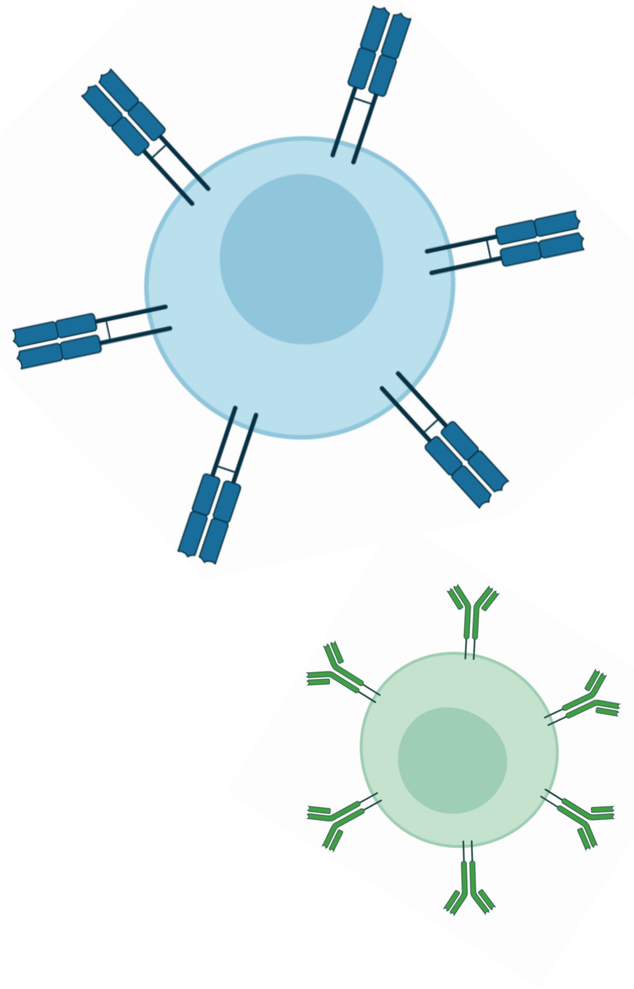
We can use **repertoire *statistics*** to infer previous dynamics



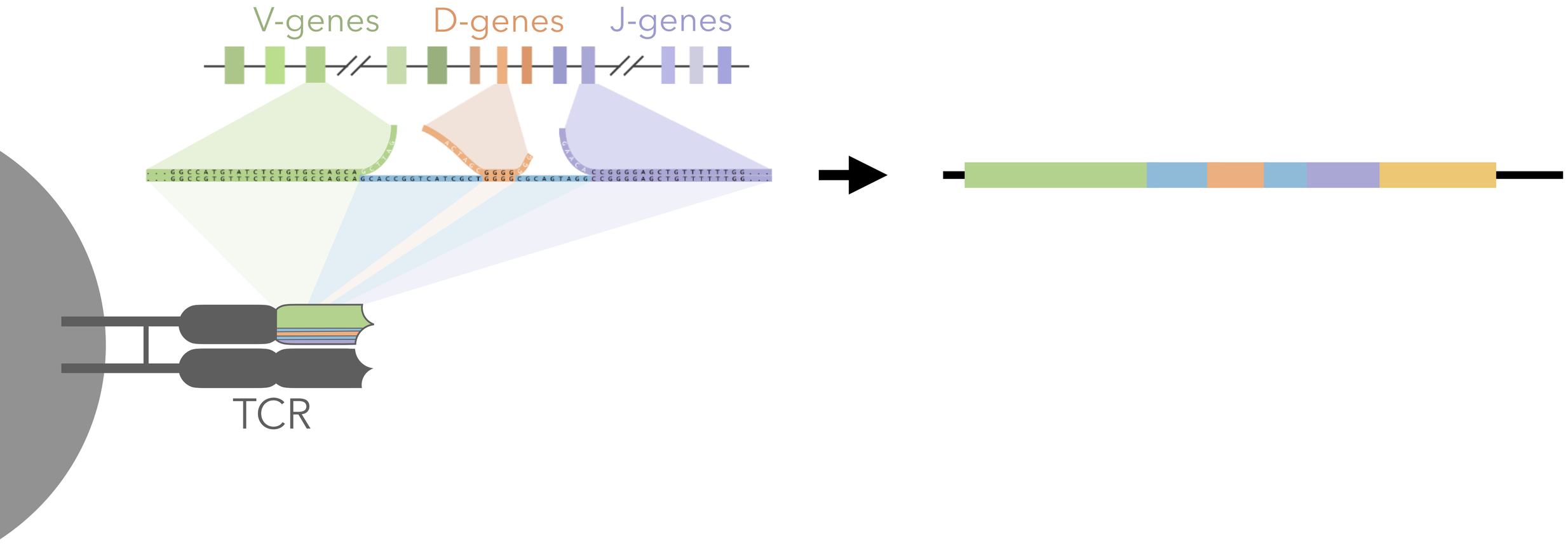
modified from Murugan et. al *PNAS* 2012
Created with BioRender

Lecture goals:

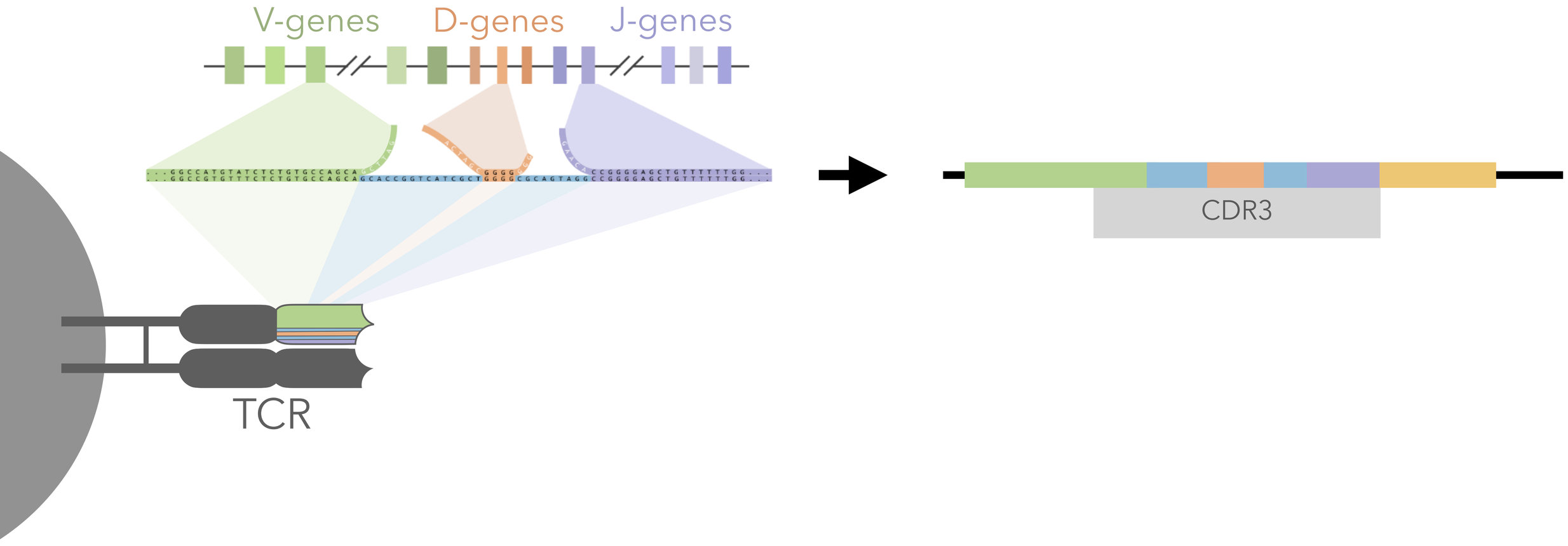
1. learn about immune repertoire sequencing
    - what are immune repertoires?
    - how are they formed?
    - **how are they sequenced?**



Sequencing is designed to capture the **most variable region**



Sequencing is designed to capture the **most variable region**



Both **bulk sequencing** and **single-cell sequencing** are possible

	Bulk	**Single-cell**
Sample size *(e.g. total # of sampled cells)*		
Repertoire coverage *(e.g. proportion of total repertoire sequenced)*		
Chain pairing *(e.g. each receptor consists of two protein chains)*		
Cost		

Both **bulk sequencing** and **single-cell sequencing** are possible

	Bulk	**Single-cell**
Sample size *(e.g. total # of sampled cells)*	**Large**	**Small**
Repertoire coverage *(e.g. proportion of total repertoire sequenced)*		
Chain pairing *(e.g. each receptor consists of two protein chains)*		
Cost		

Both **bulk sequencing** and **single-cell sequencing** are possible

	Bulk	**Single-cell**
Sample size *(e.g. total # of sampled cells)*	**Large**	**Small**
Repertoire coverage *(e.g. proportion of total repertoire sequenced)*	**High**	**Low**
Chain pairing *(e.g. each receptor consists of two protein chains)*		
Cost		

Both **bulk sequencing** and **single-cell sequencing** are possible

	Bulk	**Single-cell**
Sample size *(e.g. total # of sampled cells)*	**Large**	**Small**
Repertoire coverage *(e.g. proportion of total repertoire sequenced)*	**High**	**Low**
Chain pairing *(e.g. each receptor consists of two protein chains)*	**No**	**Yes**
Cost		

Both **bulk sequencing** and **single-cell sequencing** are possible

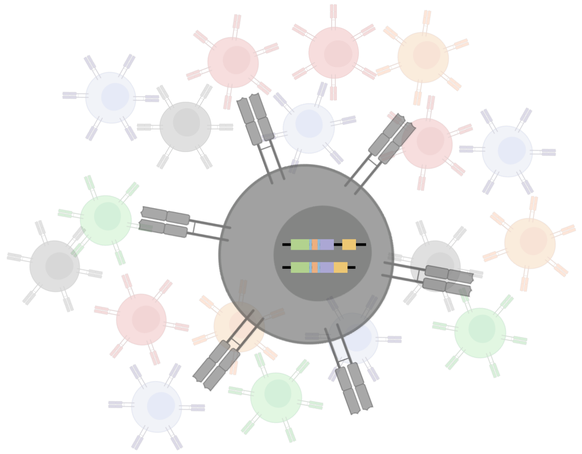
	Bulk	**Single-cell**
Sample size *(e.g. total # of sampled cells)*	**Large**	**Small**
Repertoire coverage *(e.g. proportion of total repertoire sequenced)*	**High**	**Low**
Chain pairing *(e.g. each receptor consists of two protein chains)*	**No**	**Yes**
Cost	**Low**	**High**

Both **bulk sequencing** and **single-cell sequencing** are possible

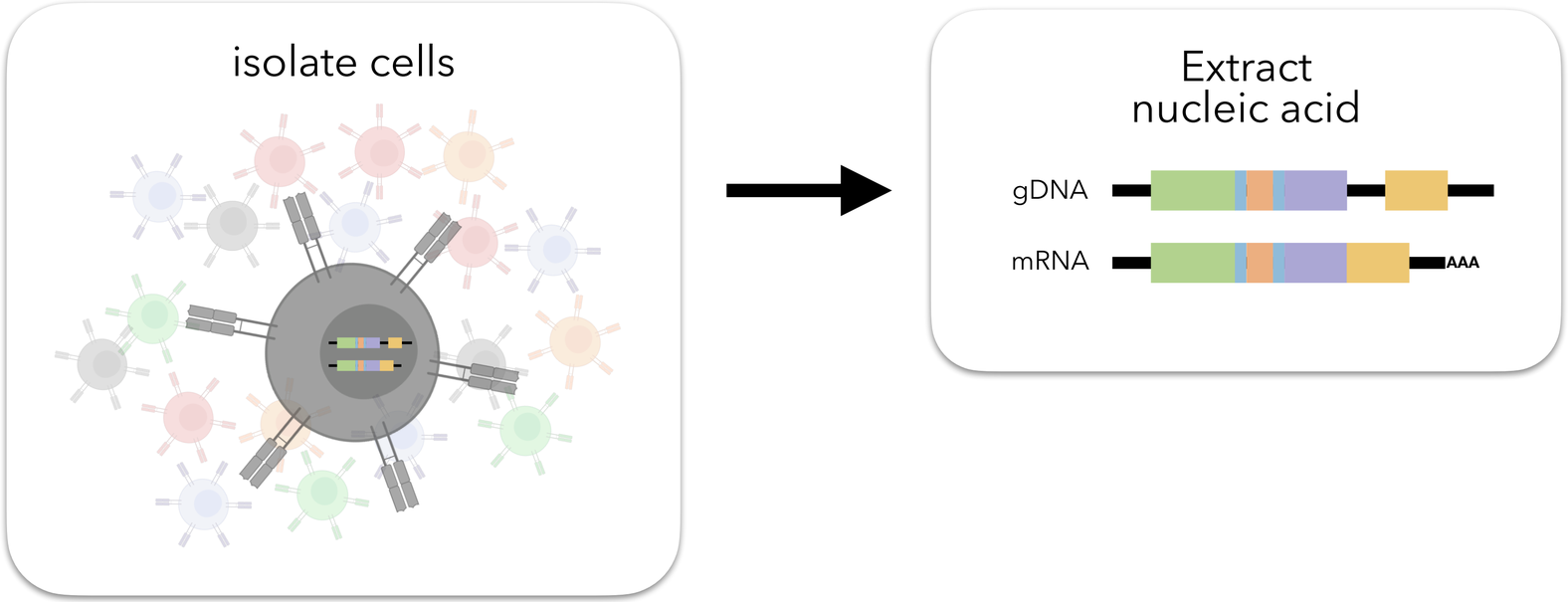
	Bulk	**Single-cell**
Sample size *(e.g. total # of sampled cells)*	**Large**	**Small**
Repertoire coverage *(e.g. proportion of total repertoire sequenced)*	**High**	**Low**
Chain pairing *(e.g. each receptor consists of two protein chains)*	**No**	**Yes**
Cost	**Low**	**High**

Bulk repertoire sequencing overview

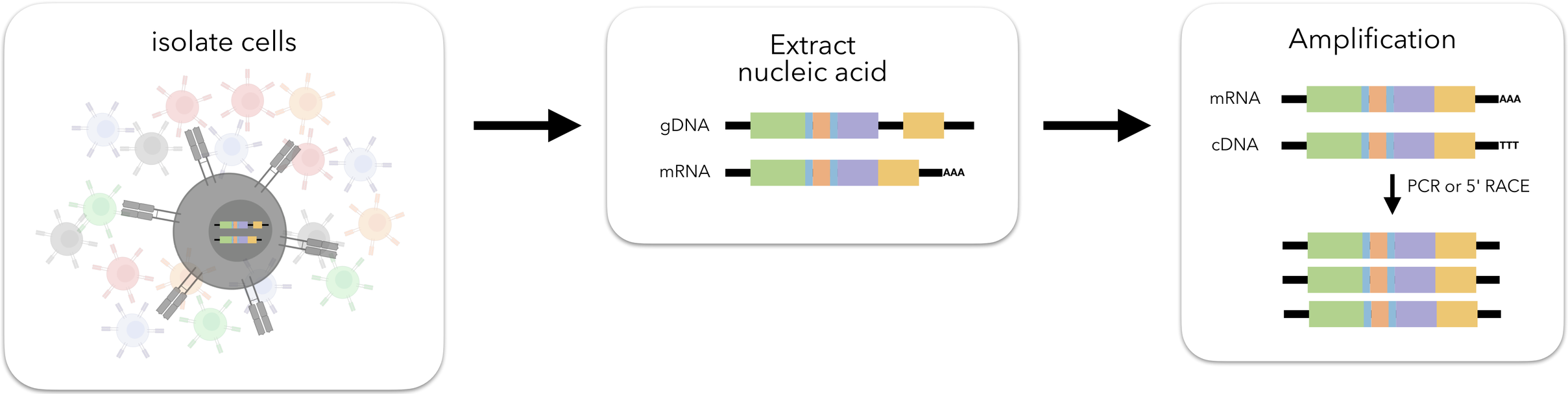
isolate cells



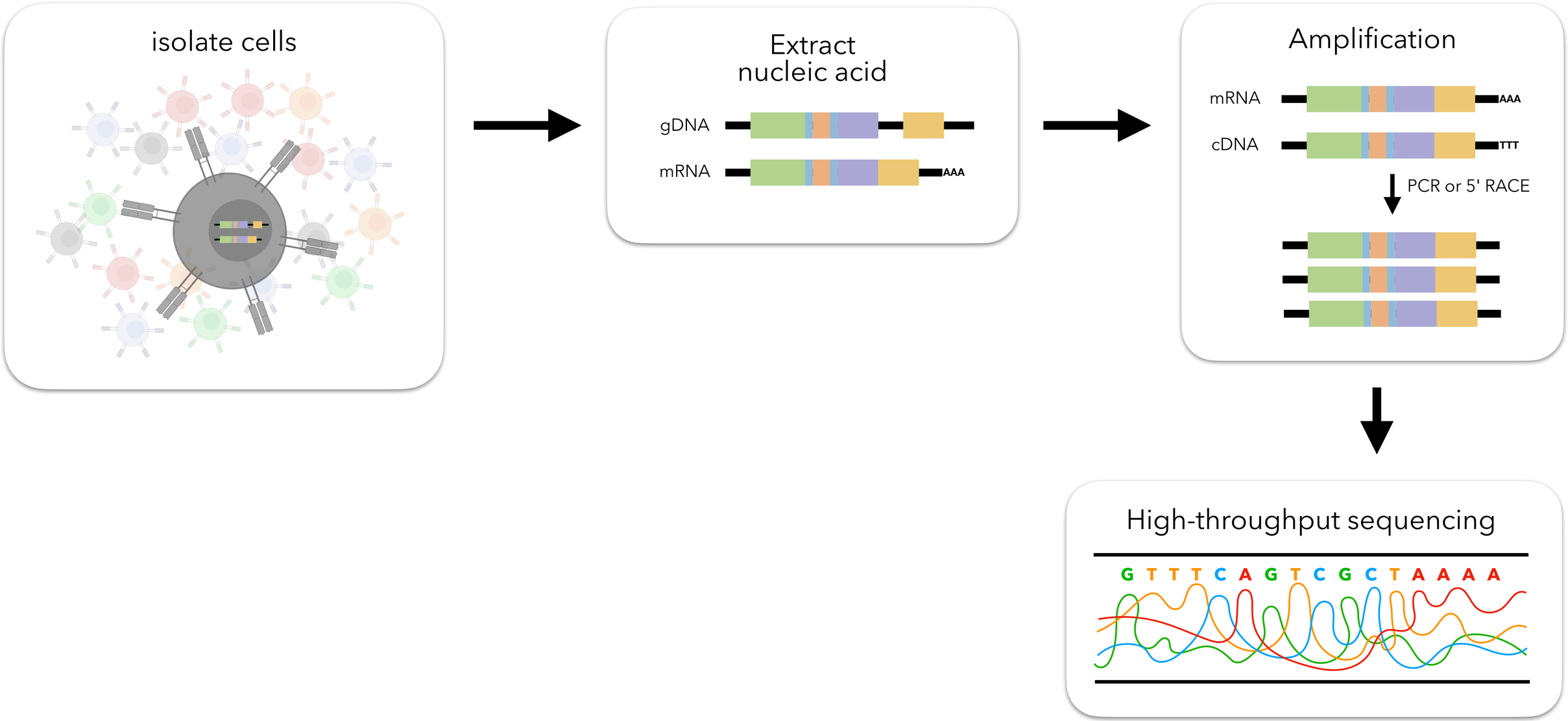
Bulk repertoire sequencing overview



Bulk repertoire sequencing overview

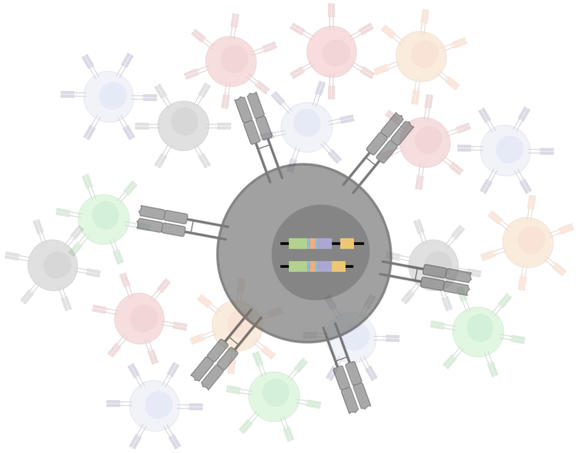


Bulk repertoire sequencing overview



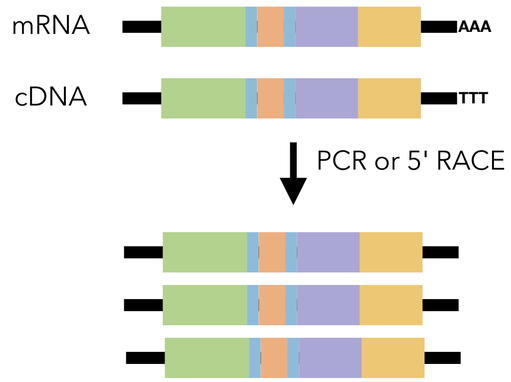
Bulk repertoire sequencing overview

isolate cells

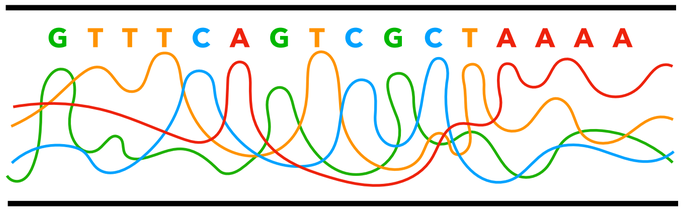


Extract nucleic acid

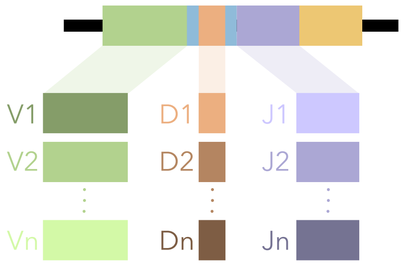
Amplification



High-throughput sequencing



Data preprocessing

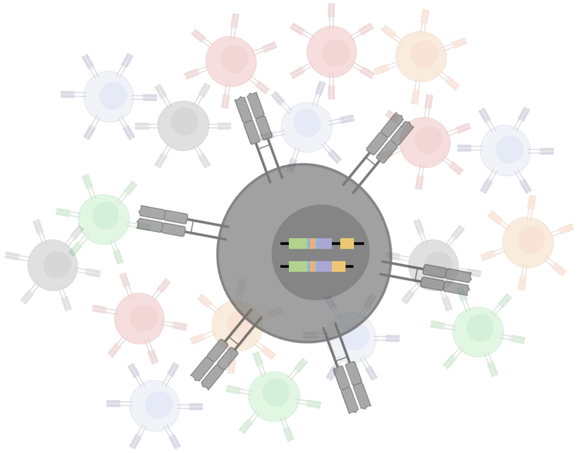


(filtering, alignment, annotation, etc.)

e.g. IGoR, MiXCR, IgBLAST, etc.

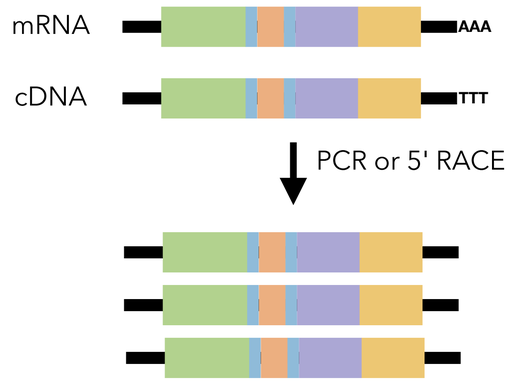
Bulk repertoire sequencing overview

isolate cells

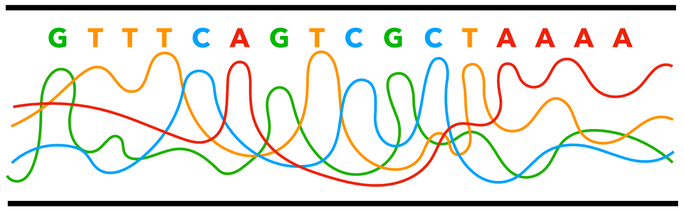


Extract
nucleic acid

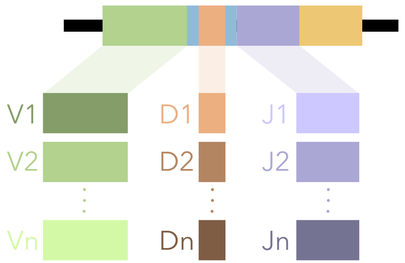
Amplification



High-throughput sequencing



Data preprocessing

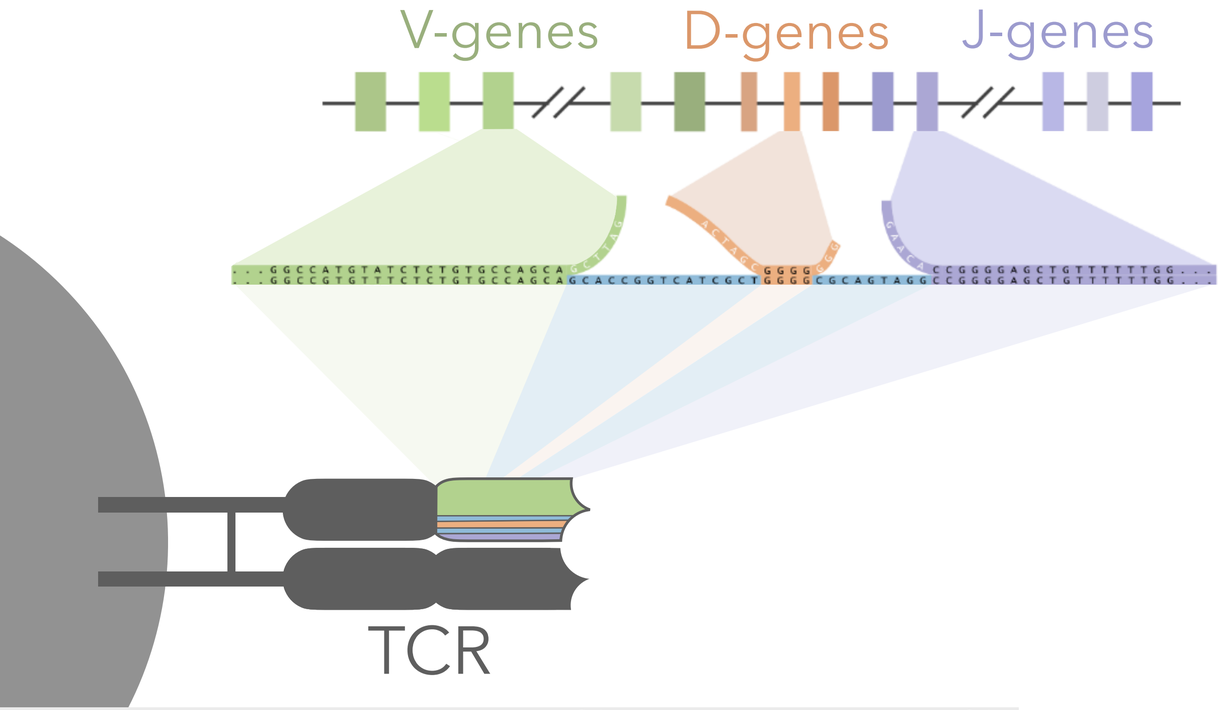


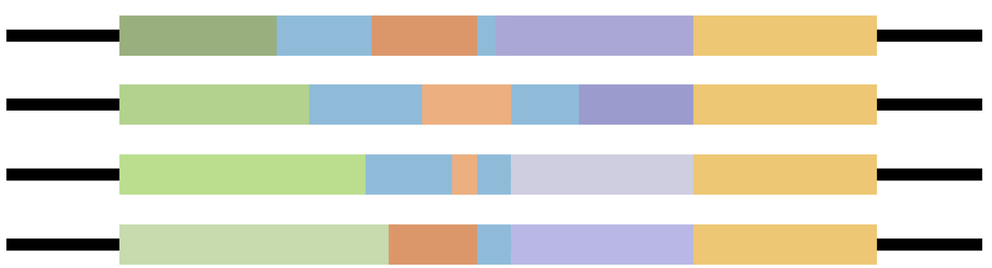
(filtering, alignment,
annotation, etc.)

e.g. IGoR, MiXCR, IgBLAST, etc.

Data analysis

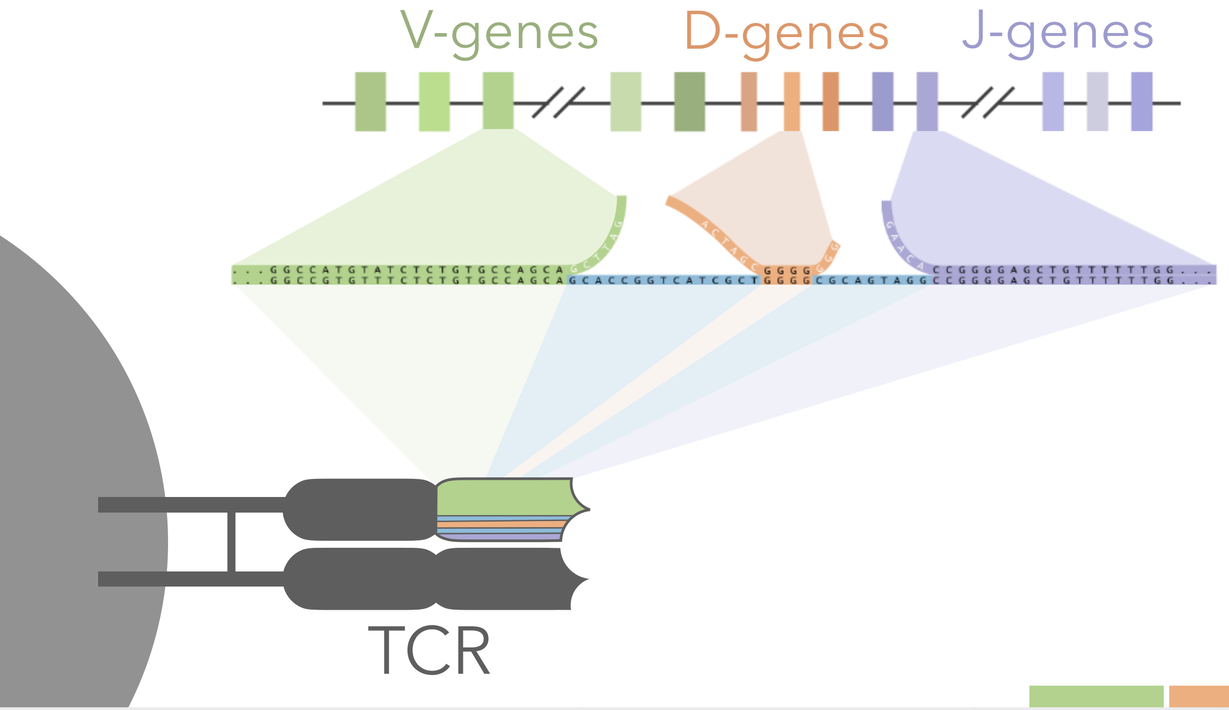
Processed **bulk repertoire sequencing** example output

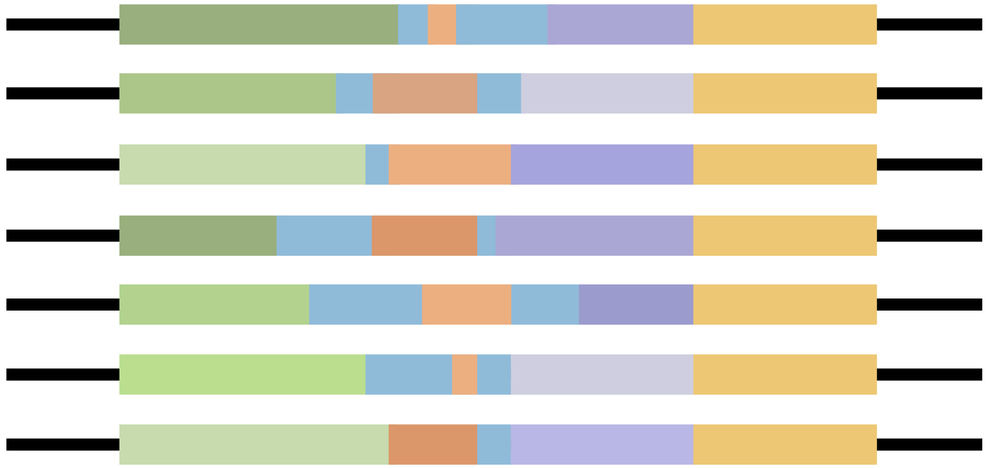




cdr3_nucseq	cdr3
<chr>	<chr>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF

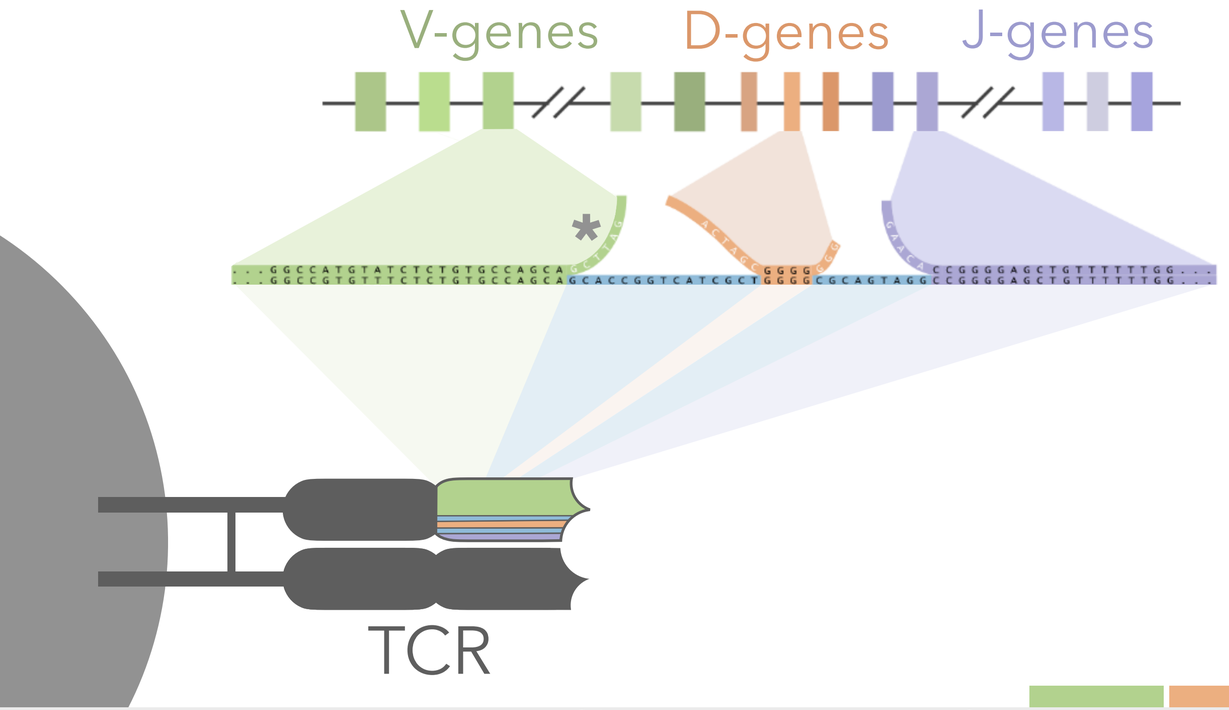
Processed **bulk repertoire sequencing** example output

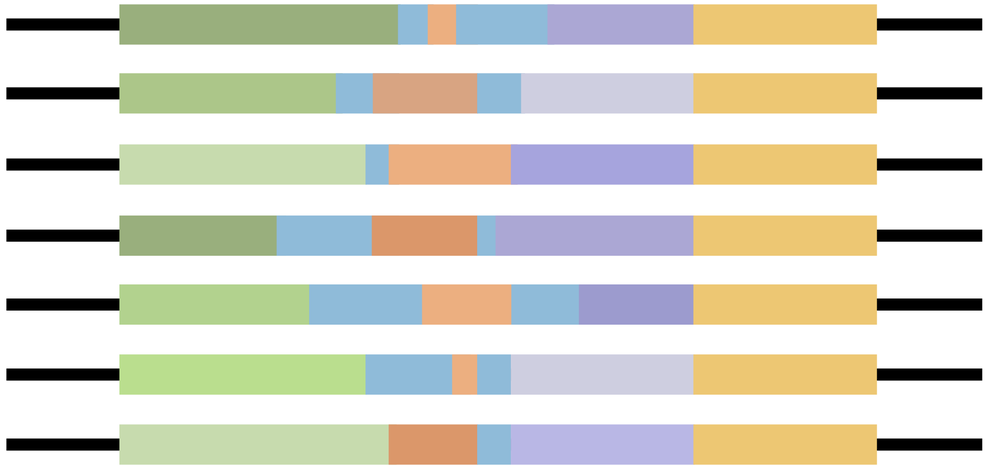




cdr3_nucseq	cdr3	v_gene	d_gene	j_gene
<chr>	<chr>	<chr>	<chr>	<chr>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF	TRBV5-6*01	TRBD2*02	TRBJ2-7*01
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF	TRBV5-1*01	TRBD1*01	TRBJ2-5*01
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF	TRBV5-1*01	TRBD2*01	TRBJ2-1*01
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF	TRBV7-2*01	TRBD2*02	TRBJ2-7*01
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF	TRBV30*01	TRBD1*01	TRBJ1-1*01
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF	TRBV2*03	TRBD1*01	TRBJ2-5*01
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF	TRBV28*01	TRBD2*02	TRBJ2-1*01
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF	TRBV30*01	TRBD2*01	TRBJ2-1*01
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF	TRBV6-6*02	TRBD1*01	TRBJ1-1*01
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF	TRBV5-5*01	TRBD2*02	TRBJ1-5*01
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF	TRBV6-1*01	TRBD1*01	TRBJ1-2*01

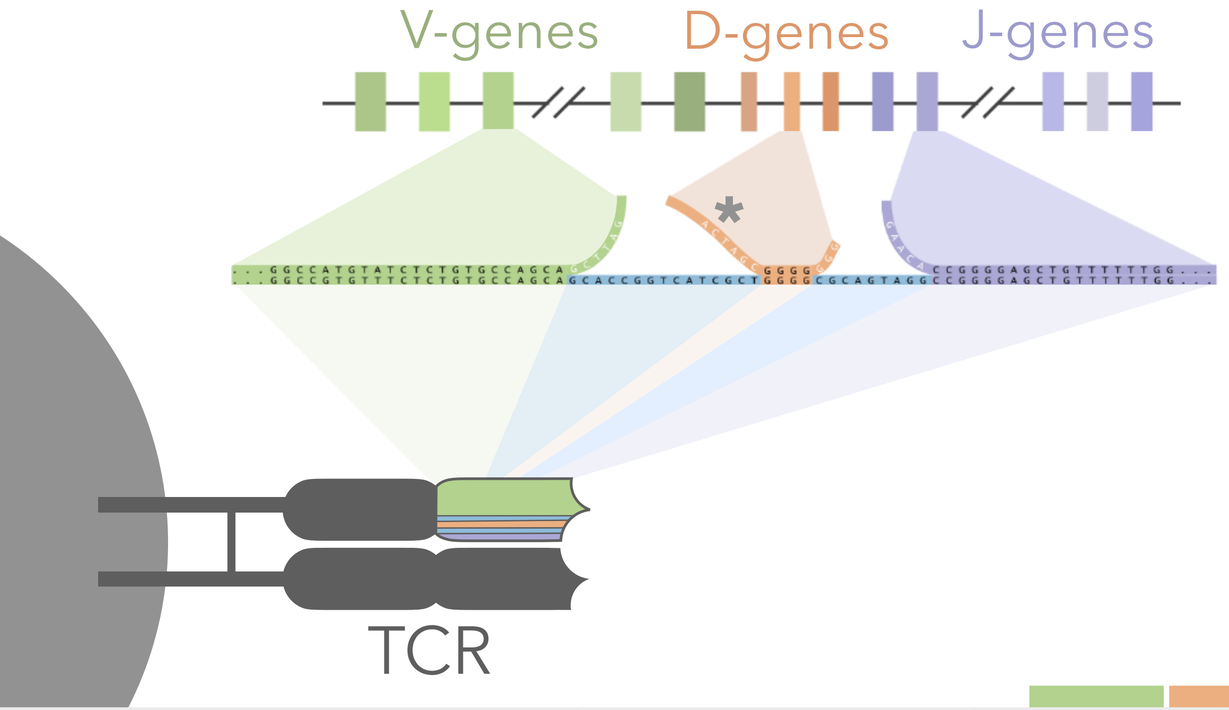
Processed **bulk repertoire sequencing** example output

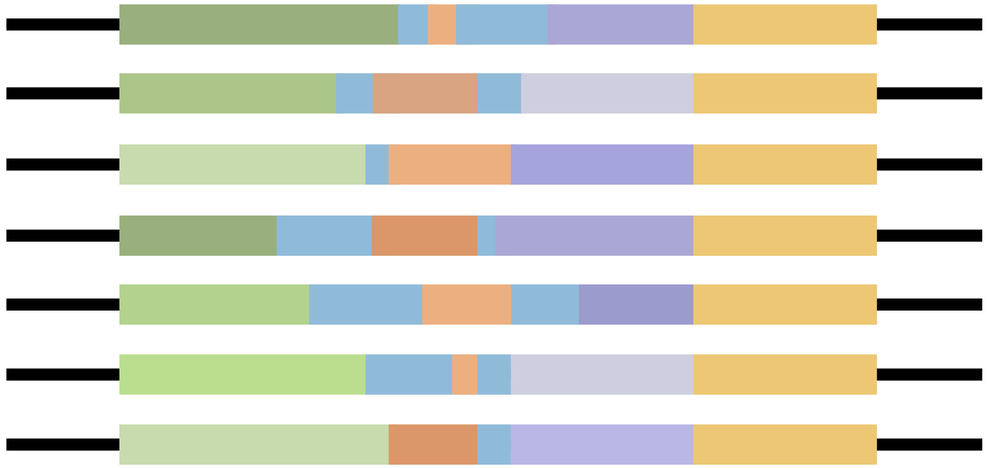




cdr3_nucseq	cdr3	v_gene	d_gene	j_gene	v_trim
<chr>	<chr>	<chr>	<chr>	<chr>	<int>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF	TRBV5-6*01	TRBD2*02	TRBJ2-7*01	1
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF	TRBV5-1*01	TRBD1*01	TRBJ2-5*01	2
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF	TRBV5-1*01	TRBD2*01	TRBJ2-1*01	9
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF	TRBV7-2*01	TRBD2*02	TRBJ2-7*01	2
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF	TRBV30*01	TRBD1*01	TRBJ1-1*01	5
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF	TRBV2*03	TRBD1*01	TRBJ2-5*01	10
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF	TRBV28*01	TRBD2*02	TRBJ2-1*01	7
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF	TRBV30*01	TRBD2*01	TRBJ2-1*01	3
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF	TRBV6-6*02	TRBD1*01	TRBJ1-1*01	4
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF	TRBV5-5*01	TRBD2*02	TRBJ1-5*01	7
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF	TRBV6-1*01	TRBD1*01	TRBJ1-2*01	9

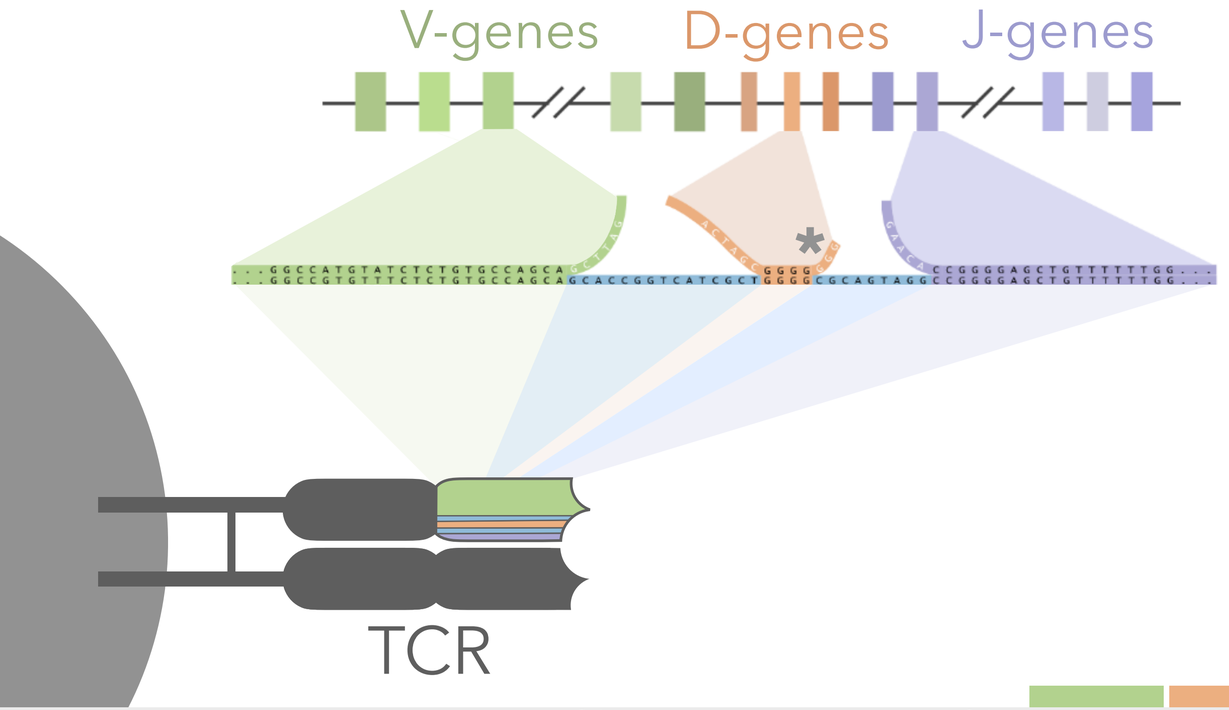
Processed **bulk repertoire sequencing** example output

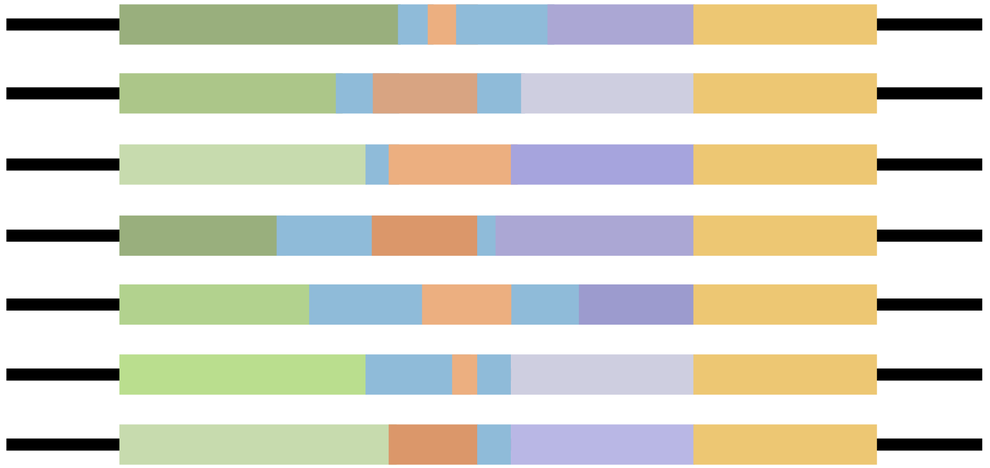




cdr3_nucseq	cdr3	v_gene	d_gene	j_gene	v_trim	d0_trim
<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF	TRBV5-6*01	TRBD2*02	TRBJ2-7*01	1	3
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF	TRBV5-1*01	TRBD1*01	TRBJ2-5*01	2	9
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF	TRBV5-1*01	TRBD2*01	TRBJ2-1*01	9	6
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF	TRBV7-2*01	TRBD2*02	TRBJ2-7*01	2	11
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF	TRBV30*01	TRBD1*01	TRBJ1-1*01	5	4
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF	TRBV2*03	TRBD1*01	TRBJ2-5*01	10	5
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF	TRBV28*01	TRBD2*02	TRBJ2-1*01	7	6
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF	TRBV30*01	TRBD2*01	TRBJ2-1*01	3	7
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF	TRBV6-6*02	TRBD1*01	TRBJ1-1*01	4	4
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF	TRBV5-5*01	TRBD2*02	TRBJ1-5*01	7	12
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF	TRBV6-1*01	TRBD1*01	TRBJ1-2*01	9	7

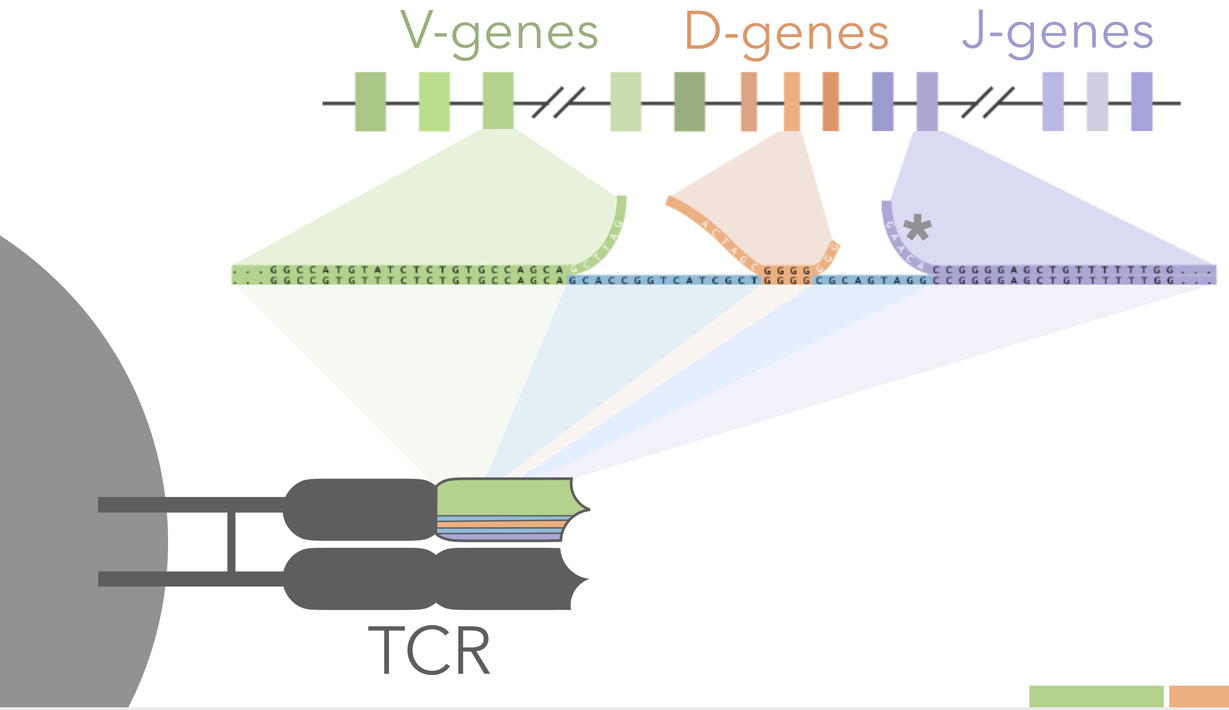
Processed **bulk repertoire sequencing** example output

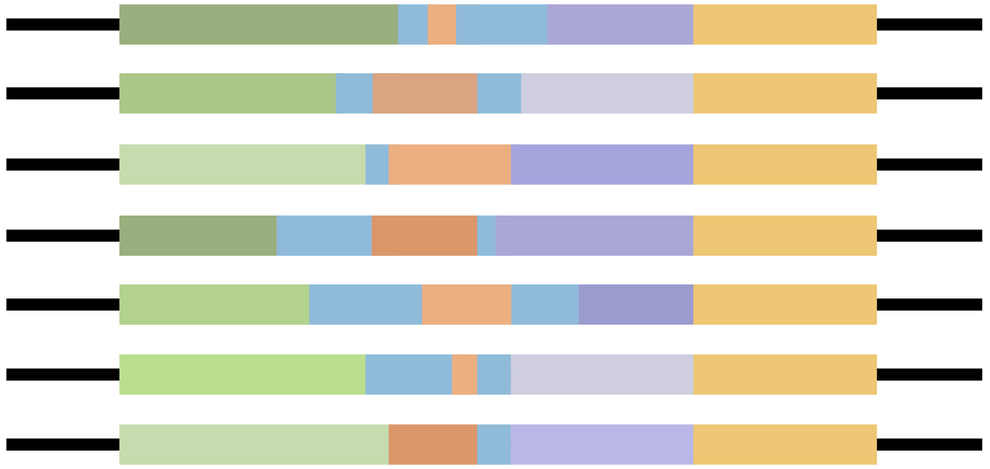




cdr3_nucseq	cdr3	v_gene	d_gene	j_gene	v_trim	d0_trim	d1_trim
<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>	<int>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF	TRBV5-6*01	TRBD2*02	TRBJ2-7*01	1	3	13
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF	TRBV5-1*01	TRBD1*01	TRBJ2-5*01	2	9	0
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF	TRBV5-1*01	TRBD2*01	TRBJ2-1*01	9	6	5
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF	TRBV7-2*01	TRBD2*02	TRBJ2-7*01	2	11	1
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF	TRBV30*01	TRBD1*01	TRBJ1-1*01	5	4	0
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF	TRBV2*03	TRBD1*01	TRBJ2-5*01	10	5	3
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF	TRBV28*01	TRBD2*02	TRBJ2-1*01	7	6	3
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF	TRBV30*01	TRBD2*01	TRBJ2-1*01	3	7	3
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF	TRBV6-6*02	TRBD1*01	TRBJ1-1*01	4	4	5
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF	TRBV5-5*01	TRBD2*02	TRBJ1-5*01	7	12	0
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF	TRBV6-1*01	TRBD1*01	TRBJ1-2*01	9	7	2

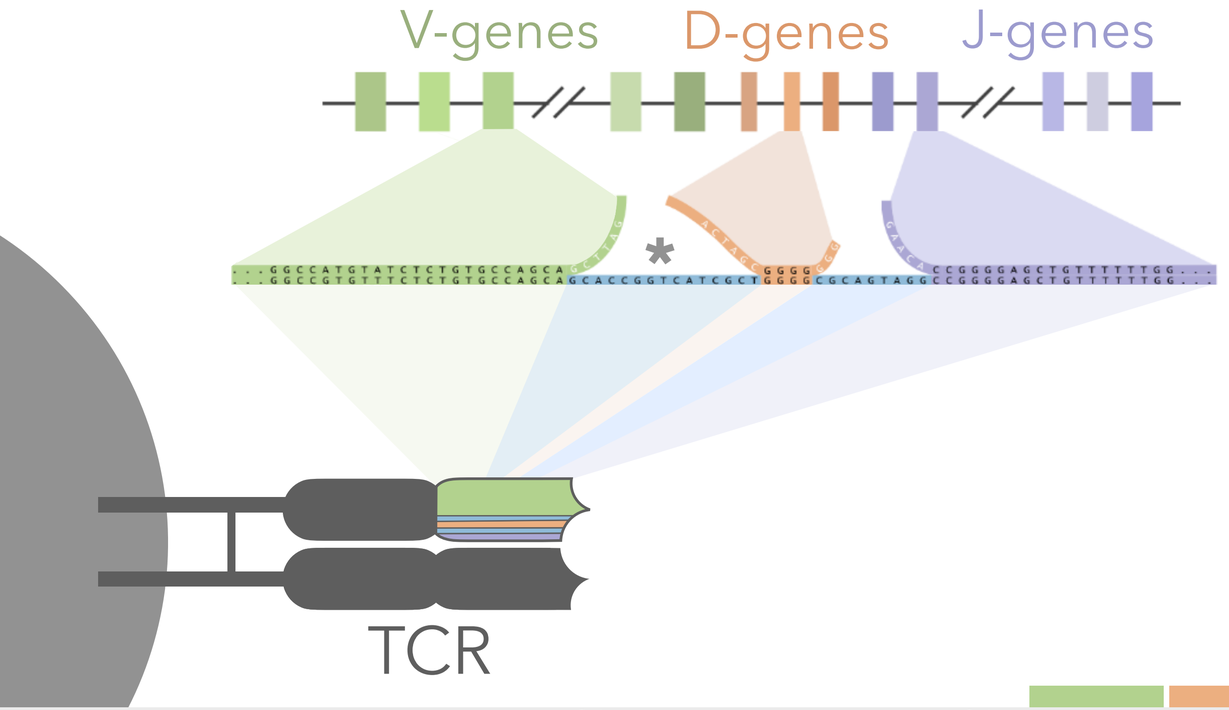
Processed **bulk repertoire sequencing** example output

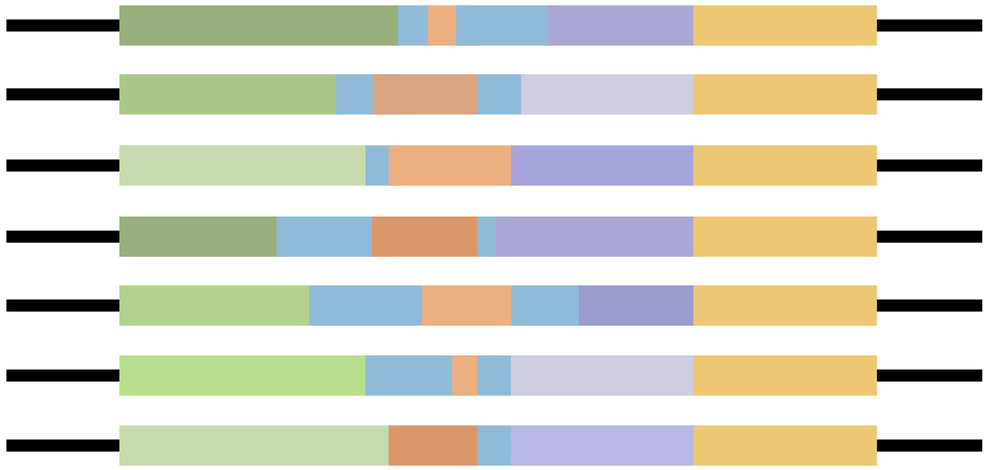




cdr3_nucseq	cdr3	v_gene	d_gene	j_gene	v_trim	d0_trim	d1_trim	j_trim
<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>	<int>	<int>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF	TRBV5-6*01	TRBD2*02	TRBJ2-7*01	1	3	13	5
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF	TRBV5-1*01	TRBD1*01	TRBJ2-5*01	2	9	0	4
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF	TRBV5-1*01	TRBD2*01	TRBJ2-1*01	9	6	5	0
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF	TRBV7-2*01	TRBD2*02	TRBJ2-7*01	2	11	1	4
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF	TRBV30*01	TRBD1*01	TRBJ1-1*01	5	4	0	2
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF	TRBV2*03	TRBD1*01	TRBJ2-5*01	10	5	3	1
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF	TRBV28*01	TRBD2*02	TRBJ2-1*01	7	6	3	9
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF	TRBV30*01	TRBD2*01	TRBJ2-1*01	3	7	3	5
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF	TRBV6-6*02	TRBD1*01	TRBJ1-1*01	4	4	5	4
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF	TRBV5-5*01	TRBD2*02	TRBJ1-5*01	7	12	0	3
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF	TRBV6-1*01	TRBD1*01	TRBJ1-2*01	9	7	2	4

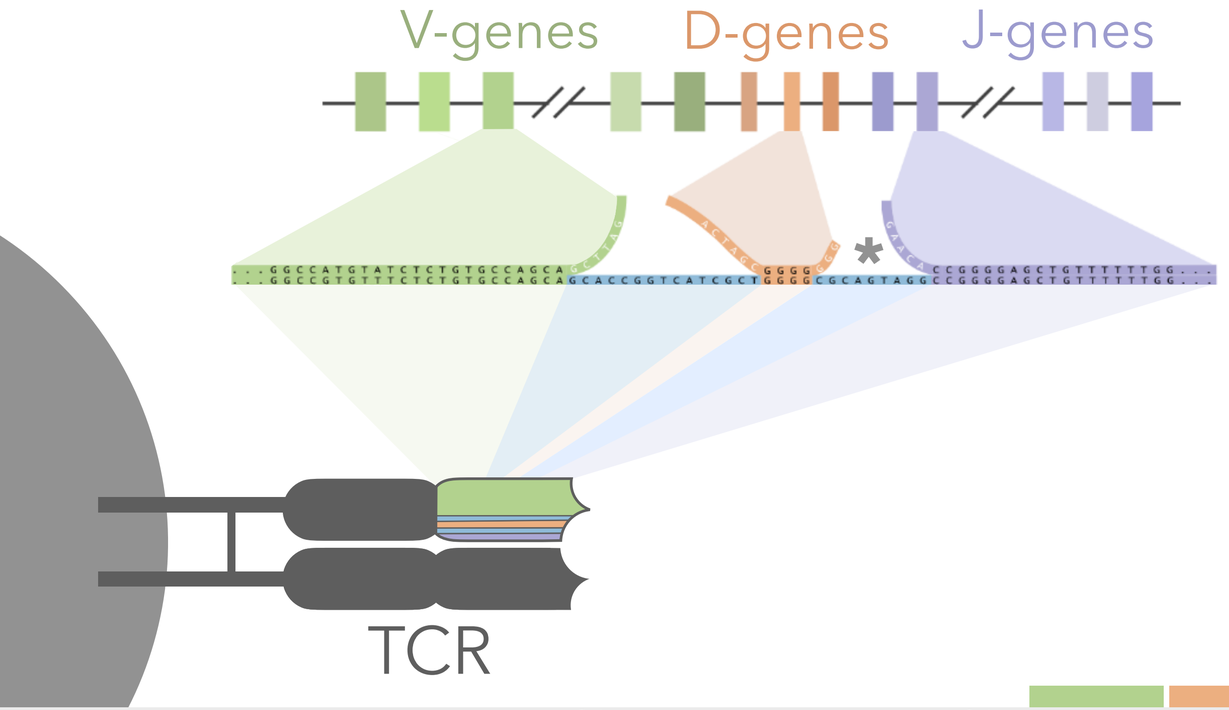
Processed **bulk repertoire sequencing** example output

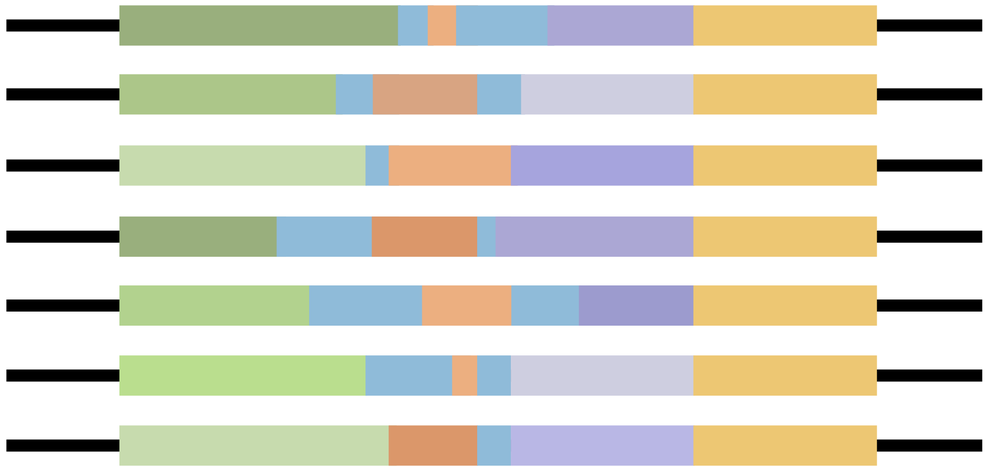




cdr3_nucseq	cdr3	v_gene	d_gene	j_gene	v_trim	d0_trim	d1_trim	j_trim	vd_insert	vd_insert_nucs
<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>	<int>	<int>	<int>	<chr>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF	TRBV5-6*01	TRBD2*02	TRBJ2-7*01	1	3	13	5	4	AATC
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF	TRBV5-1*01	TRBD1*01	TRBJ2-5*01	2	9	0	4	0	
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF	TRBV5-1*01	TRBD2*01	TRBJ2-1*01	9	6	5	0	4	GTCG
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF	TRBV7-2*01	TRBD2*02	TRBJ2-7*01	2	11	1	4	8	AATCTGGT
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF	TRBV30*01	TRBD1*01	TRBJ1-1*01	5	4	0	2	1	T
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF	TRBV2*03	TRBD1*01	TRBJ2-5*01	10	5	3	1	7	CCGAACG
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF	TRBV28*01	TRBD2*02	TRBJ2-1*01	7	6	3	9	1	T
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF	TRBV30*01	TRBD2*01	TRBJ2-1*01	3	7	3	5	11	CTCCCTCCCTG
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF	TRBV6-6*02	TRBD1*01	TRBJ1-1*01	4	4	5	4	2	AT
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF	TRBV5-5*01	TRBD2*02	TRBJ1-5*01	7	12	0	3	5	GGCCC
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF	TRBV6-1*01	TRBD1*01	TRBJ1-2*01	9	7	2	4	3	TGC

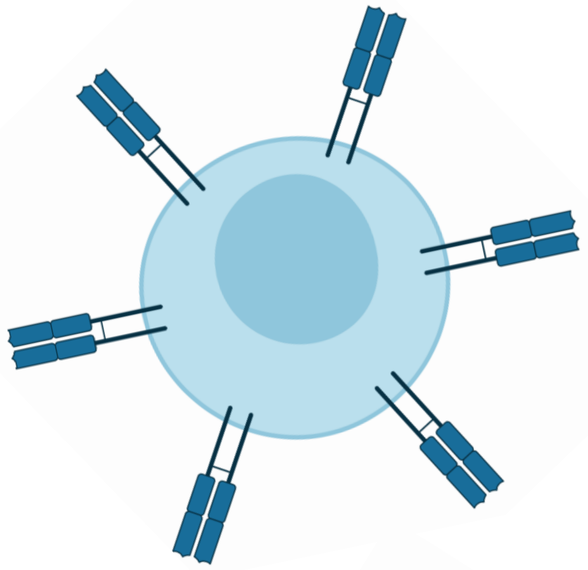
Processed **bulk repertoire sequencing** example output

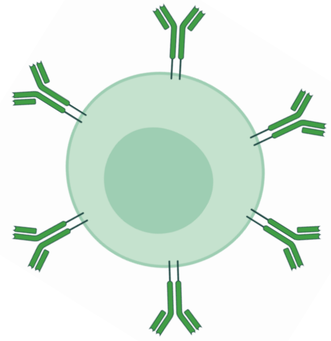




cdr3_nucseq	cdr3	v_gene	d_gene	j_gene	v_trim	d0_trim	d1_trim	j_trim	vd_insert	dj_insert	vd_insert_nucs	dj_insert_nucs
<chr>	<chr>	<chr>	<chr>	<chr>	<int>	<int>	<int>	<int>	<int>	<int>	<chr>	<chr>
TGTGCCAGCAGCTTGAATCACGAGCAGTACTTC	CASSLNHEQYF	TRBV5-6*01	TRBD2*02	TRBJ2-7*01	1	3	13	5	4	0	AATC	
TGCGCCAGCAGCTTGGCAGAGACCCAGTACTTC	CASSLAETQYF	TRBV5-1*01	TRBD1*01	TRBJ2-5*01	2	9	0	4	0	0		
TGCGCCAGTCGAGCGGCGAGCTCCTACAATGAGCAGTTCTTC	CASRAASSYNEQFF	TRBV5-1*01	TRBD2*01	TRBJ2-1*01	9	6	5	0	4	2	GTCG	GC
TGTGCCAGCAGCTTAAATCTGGTGAGGTACGAGCAGTACTTC	CASSLNLVRYEQYF	TRBV7-2*01	TRBD2*02	TRBJ2-7*01	2	11	1	4	8	0	AATCTGGT	
TGTGCCTGGTCAGGGGGCCCAAACACTGAAGCTTTCTTT	CAWSGGPNTEAFF	TRBV30*01	TRBD1*01	TRBJ1-1*01	5	4	0	2	1	3	T	ACC
TGTGCCACCGAACGAGGGCCCCAAGAGACCCAGTACTTC	CATERGPQETQYF	TRBV2*03	TRBD1*01	TRBJ2-5*01	10	5	3	1	7	2	CCGAACG	CC
TGTGCCAGCATAGCGGGAGGTGAGCAGTTCTTC	CASIAGGEQFF	TRBV28*01	TRBD2*02	TRBJ2-1*01	7	6	3	9	1	2	T	GG
TGTGCCTGGAGCTCCCTCCCTGGCGGGGAGAACAATGAGCAGTTCTTC	CAWSSLPGGENNEQFF	TRBV30*01	TRBD2*01	TRBJ2-1*01	3	7	3	5	11	3	CTCCCTCCCTG	AGA
TGTGCCAGCAGTTATCAGGTCACTGAAGCTTTCTTT	CASSYQVTEAFF	TRBV6-6*02	TRBD1*01	TRBJ1-1*01	4	4	5	4	2	2	AT	TG
TGTGCCAGCGGCCCAGGGCTCGGATACAATCAGCCCCAGCATTTT	CASGPGLGYNQPQHF	TRBV5-5*01	TRBD2*02	TRBJ1-5*01	7	12	0	3	5	8	GGCCC	ATAGGCTC
TGTGCCAGTGCGGGATTCTATGGCTACACCTTC	CASAGFYGYTF	TRBV6-1*01	TRBD1*01	TRBJ1-2*01	9	7	2	4	3	3	TGC	TTA

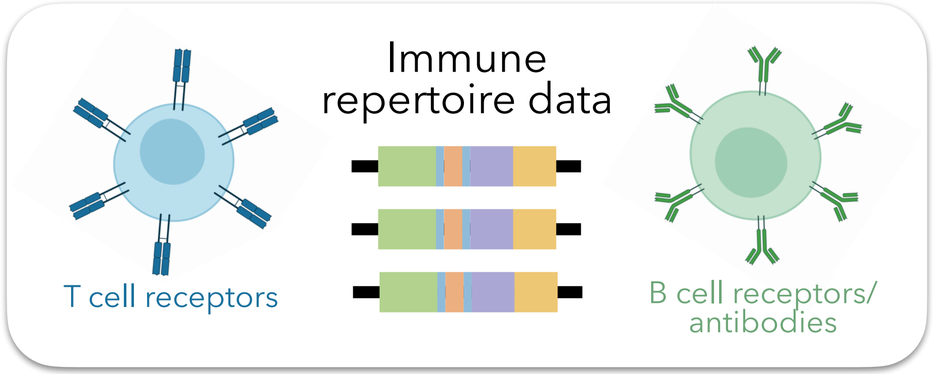
Lecture goals:





1. learn about immune repertoire sequencing
    - what are immune repertoires?
    - how are they formed?
    - how are they sequenced?
    - **what are some common areas of repertoire analysis?**

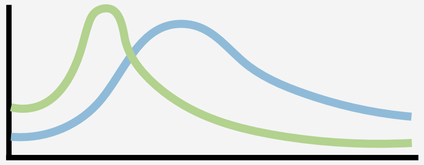
Immune repertoire analyses often focus on diversity, architecture, evolution, or convergence



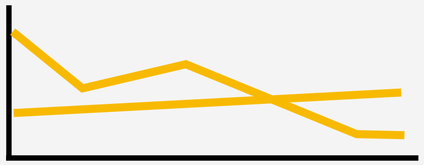
Immune repertoire analyses often focus on **diversity**, architecture, evolution, or convergence

Underlying mechanisms of diversity generation

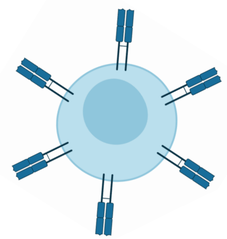
probabilistic sequence annotation



recombination statistics to learn about generation and selection

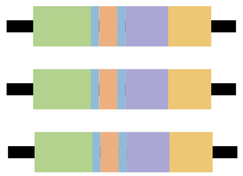


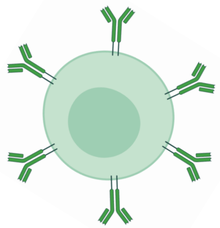
clonotype diversity dynamics



T cell receptors

Immune repertoire data

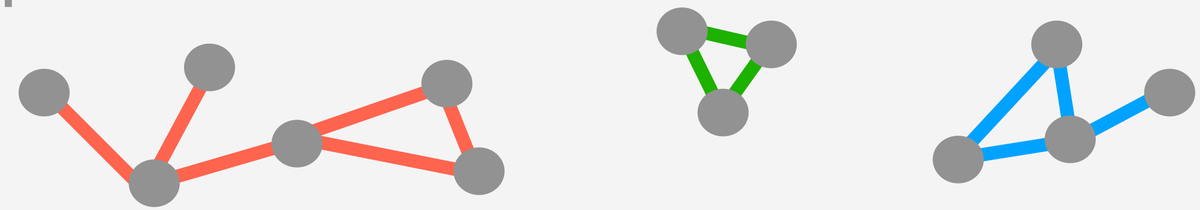




B cell receptors/ antibodies

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

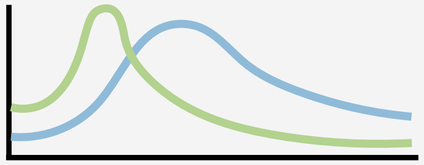
Repertoire architecture



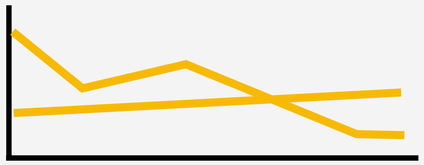
defining antigen recognition breadth using network analysis

Underlying mechanisms of diversity generation

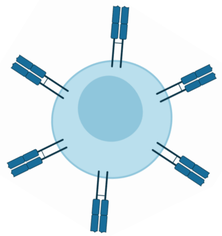
probabilistic sequence annotation



recombination statistics to learn about generation and selection



clonotype diversity dynamics



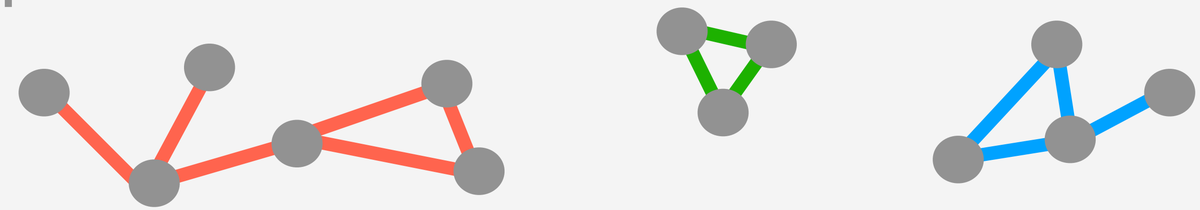
T cell receptors

Immune repertoire data

B cell receptors/ antibodies

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

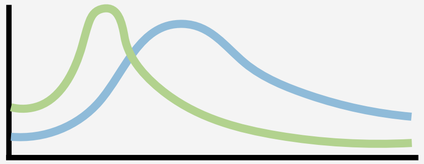
Repertoire architecture



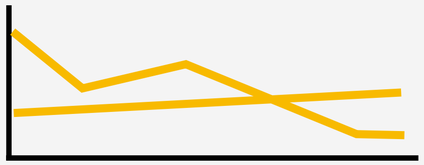
defining antigen recognition breadth using network analysis

Underlying mechanisms of diversity generation

probabilistic sequence annotation

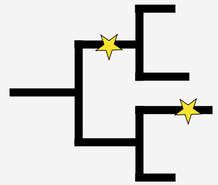


recombination statistics to learn about generation and selection

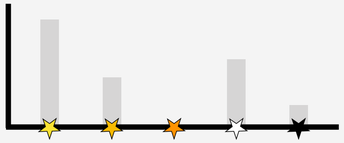


clonotype diversity dynamics

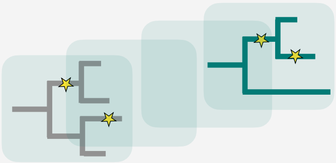
Antibody evolution



reconstruction of phylogenetic trees



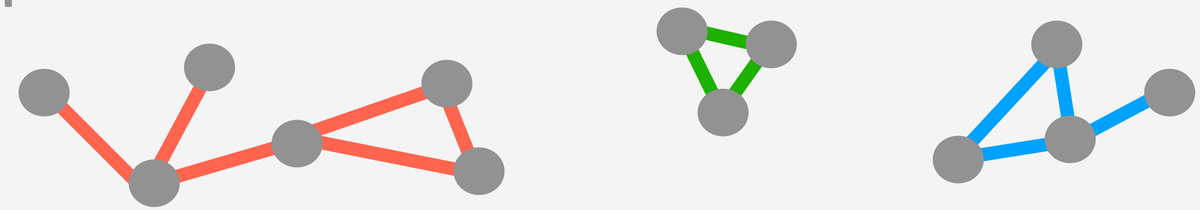
inferring mutation statistics



simulating antibody repertoire evolution

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

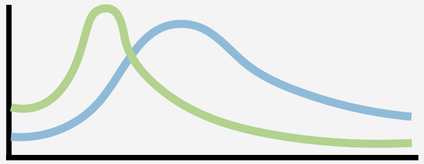
Repertoire architecture



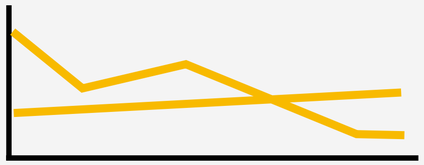
defining antigen recognition breadth using network analysis

Underlying mechanisms of diversity generation

probabilistic sequence annotation

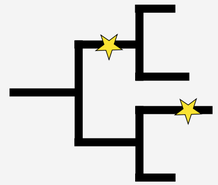


recombination statistics to learn about generation and selection

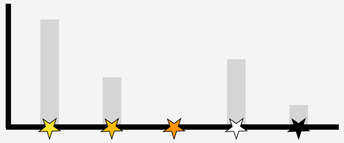


clonotype diversity dynamics

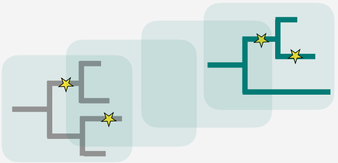
Antibody evolution



reconstruction of phylogenetic trees

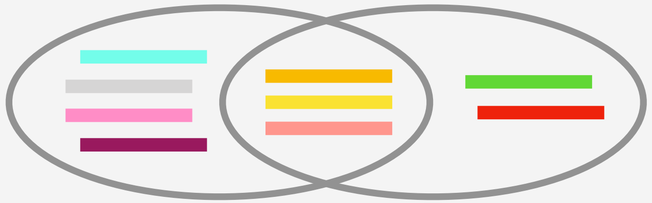


inferring mutation statistics



simulating antibody repertoire evolution

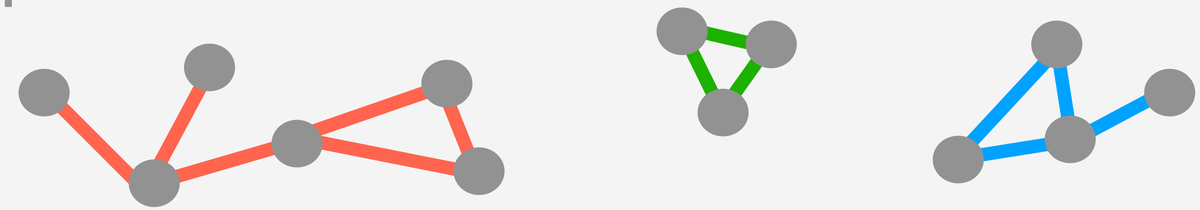
Molecular convergence



exploring cross-individual sequence similarity and convergence

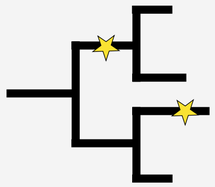
Immune repertoire analyses often focus on diversity, architecture, evolution, or convergence

Repertoire architecture

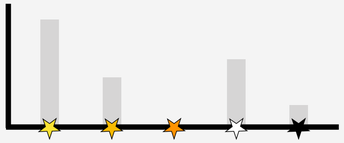


defining antigen recognition breadth using network analysis

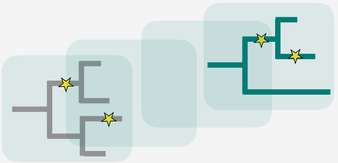
Antibody evolution



reconstruction of phylogenetic trees



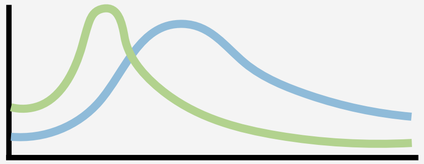
inferring mutation statistics



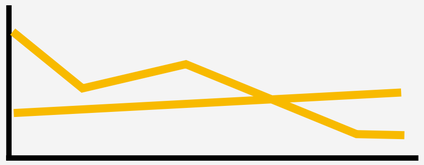
simulating antibody repertoire evolution

Underlying mechanisms of diversity generation

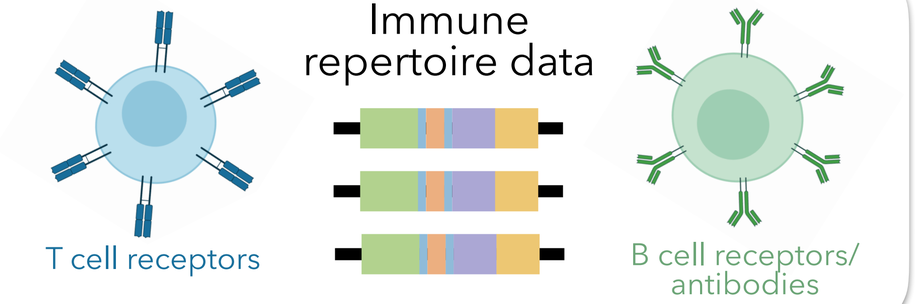
probabilistic sequence annotation



recombination statistics to learn about generation and selection



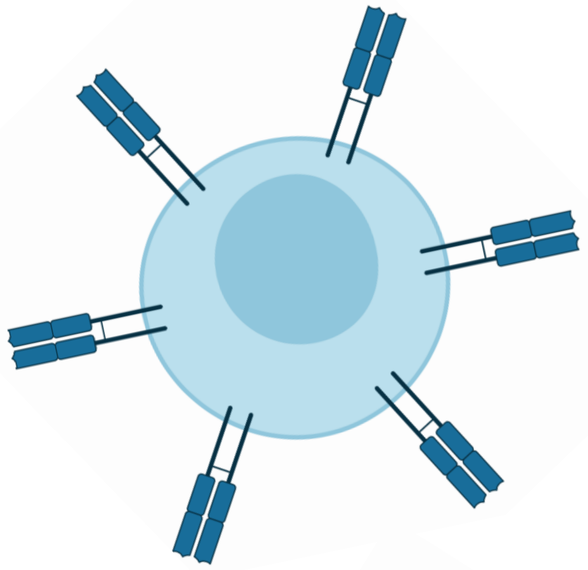
clonotype diversity dynamics

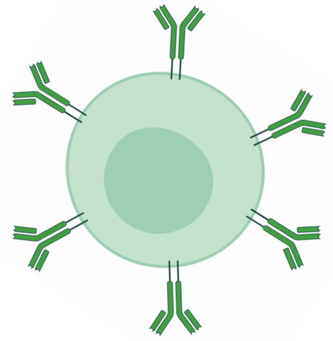


Molecular convergence

exploring cross-individual sequence similarity and convergence

Lecture goals:





1. learn about immune repertoire sequencing
    - what are immune repertoires?
    - how are they formed?
    - how are they sequenced?
    - what are some common areas of repertoire analysis?
2. familiarize with immune repertoire data
3. work through an example analysis