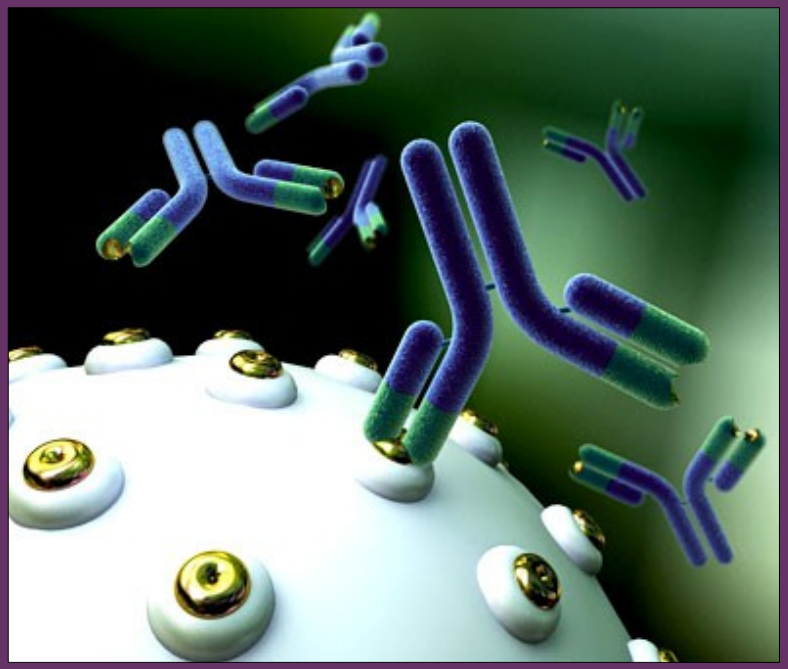


The B-side

An insight into the immune response

A DSE lapse of time

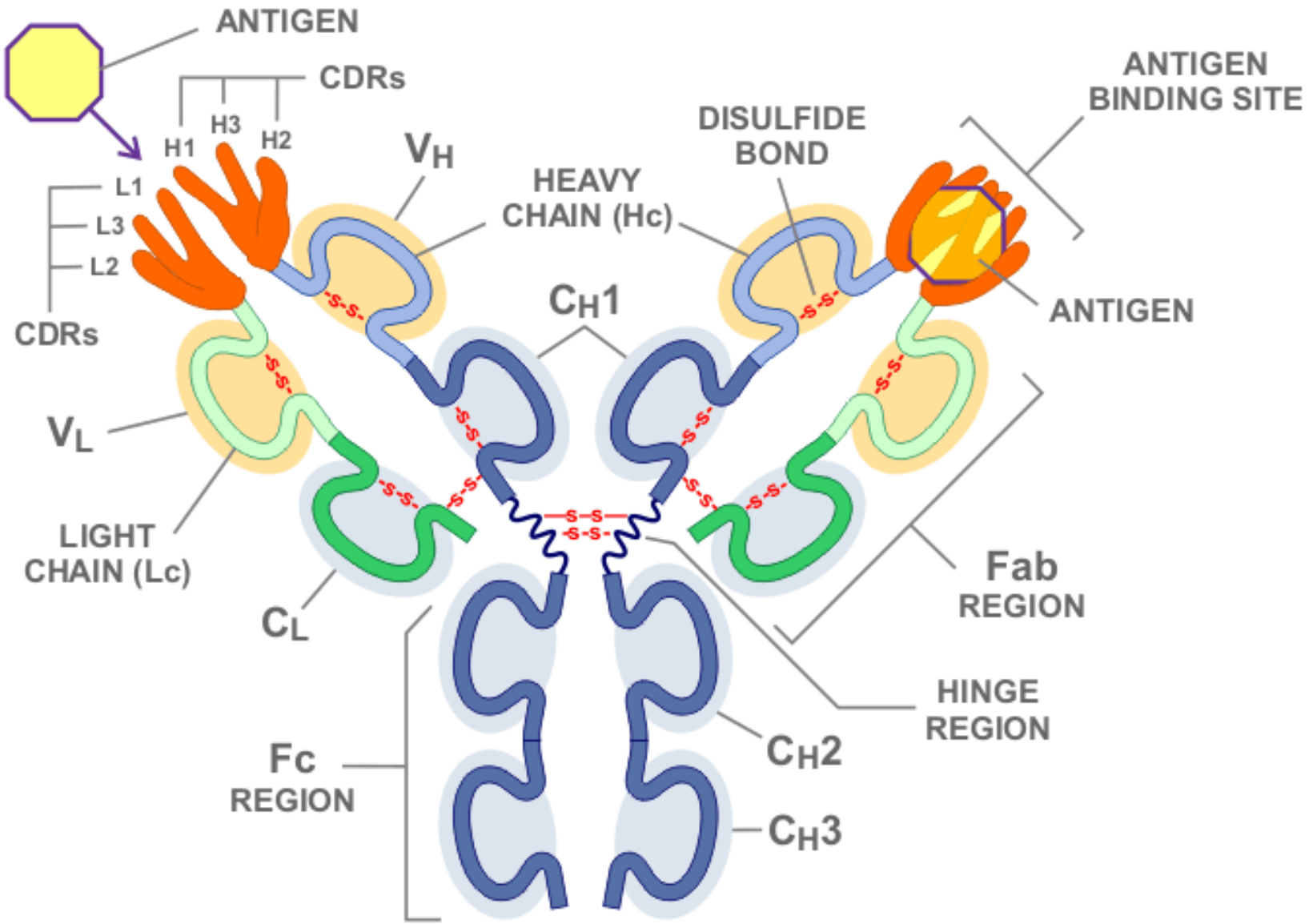
The problem of the immune system

- How to respond to an immense diversity of foreign molecular structures
- **Solution**: diverse antibodies specifically induced against any molecule (i.e. antigen)

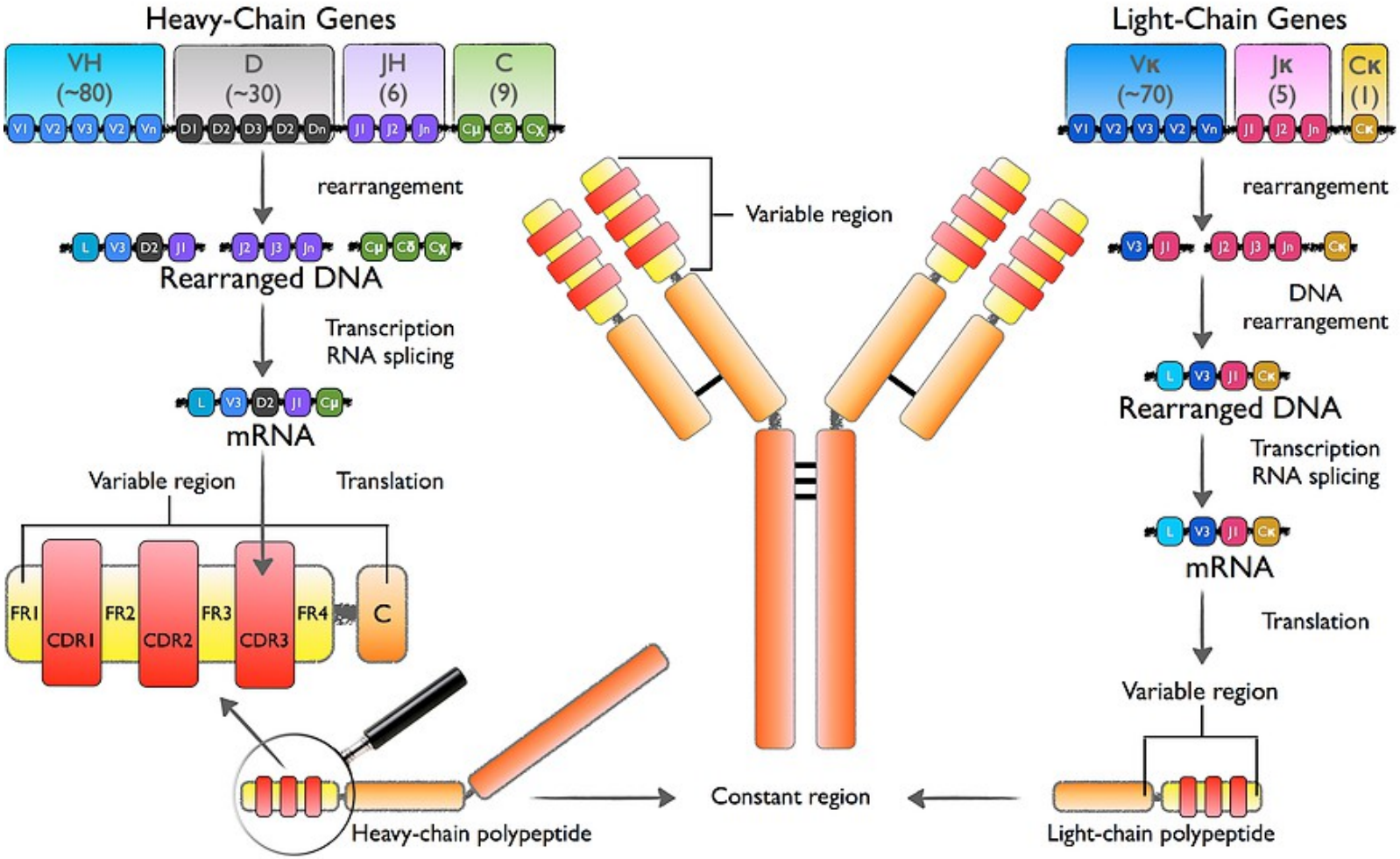
Background info

- Antibodies are large Y-shaped molecules, consisting of paired **heavy** and **light** polypeptide chains
- These form a variable (**V**) antigen-binding region (known as Heavy (**VH)** and Light (**VL)** chains), as well as a constant (**C**) region
- Variability is concentrated in the 3 complementarity-determining regions (**CDR1-3**), which form a variety of tertiary structures to bind different antigens
- The sum of all circulating antibodies (~106) is known as **the antibody repertoire.** It is not possible to sequence every single B cell in humans, so representative samples, generally derived from peripheral blood, are taken

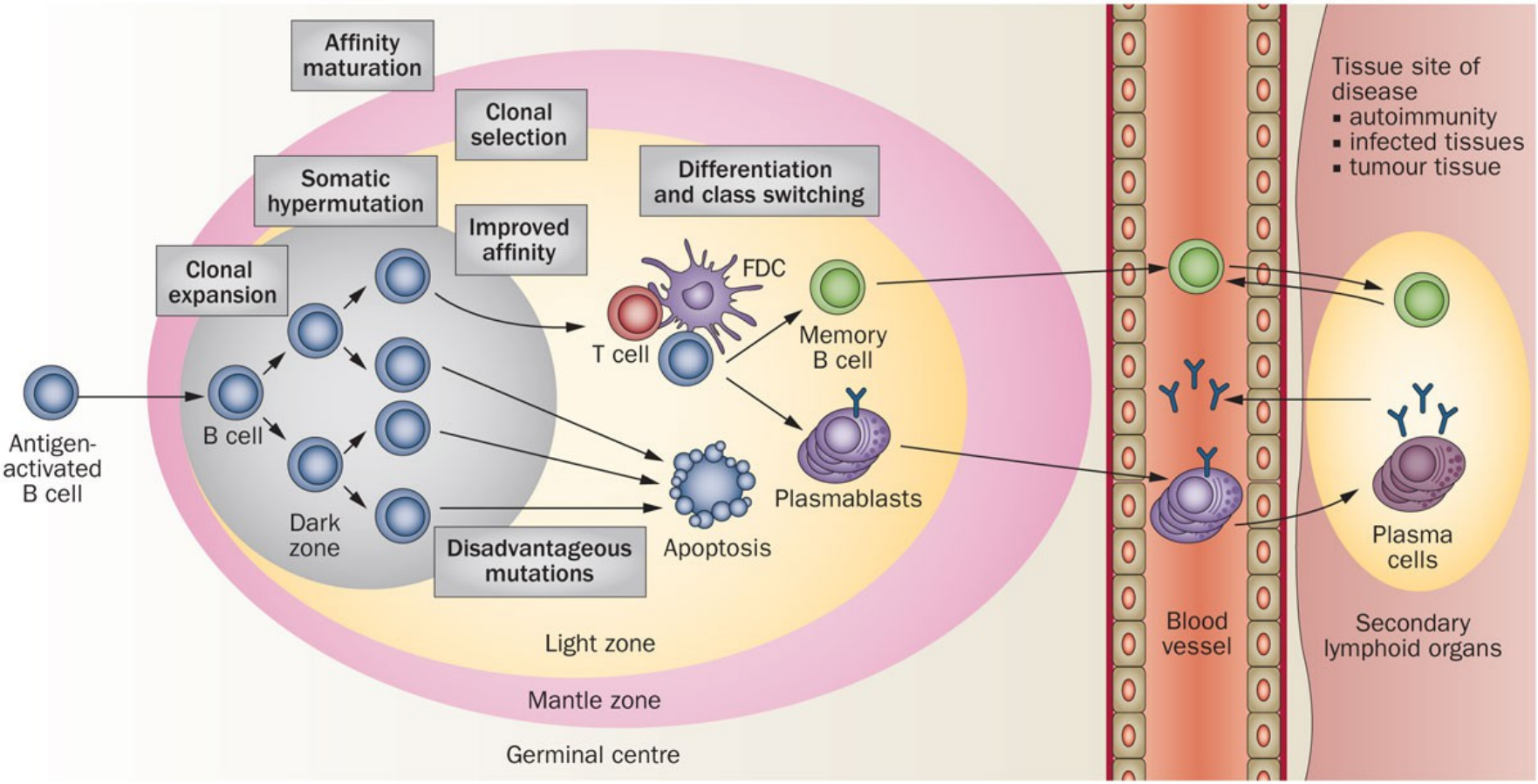
Diagram of an Antibody



The **V**ariable region is responsible for antigen binding and specificity



The steps of B-cell differentiation and diversification of the antibody repertoire

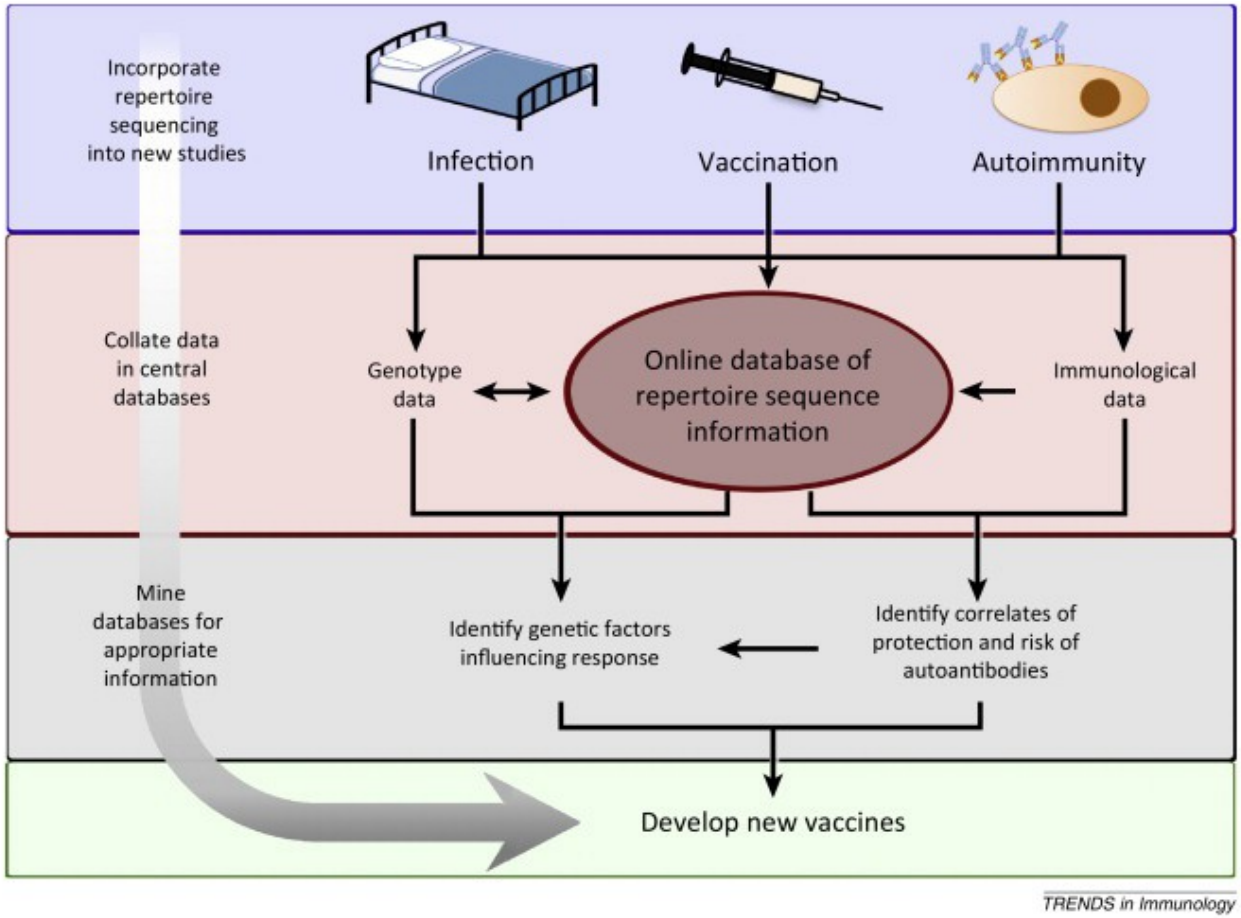


Robinson, W. H. (2014) Sequencing the functional antibody repertoire—diagnostic and therapeutic discovery. *Nat. Rev. Rheumatol.*

The Repertoire from a vaccines' point of view

- Nearly all licensed vaccines have been developed to confer protection by stimulating the production of antibodies by B-cells
- Sequencing the antibody repertoire after vaccination can provide a detailed **dissection of the vaccine-induced B-cells** underlying the antibody response and a potential application in **deriving novel correlates of protection**
- A critical first step is to **distinguish the vaccine-specific repertoire** from the total repertoire
- Recent methods allowing high-throughput identification of serum antibodies will also help to show which antigen-specific sequences give rise to long-lived plasma cells and **mediate long-term protection**

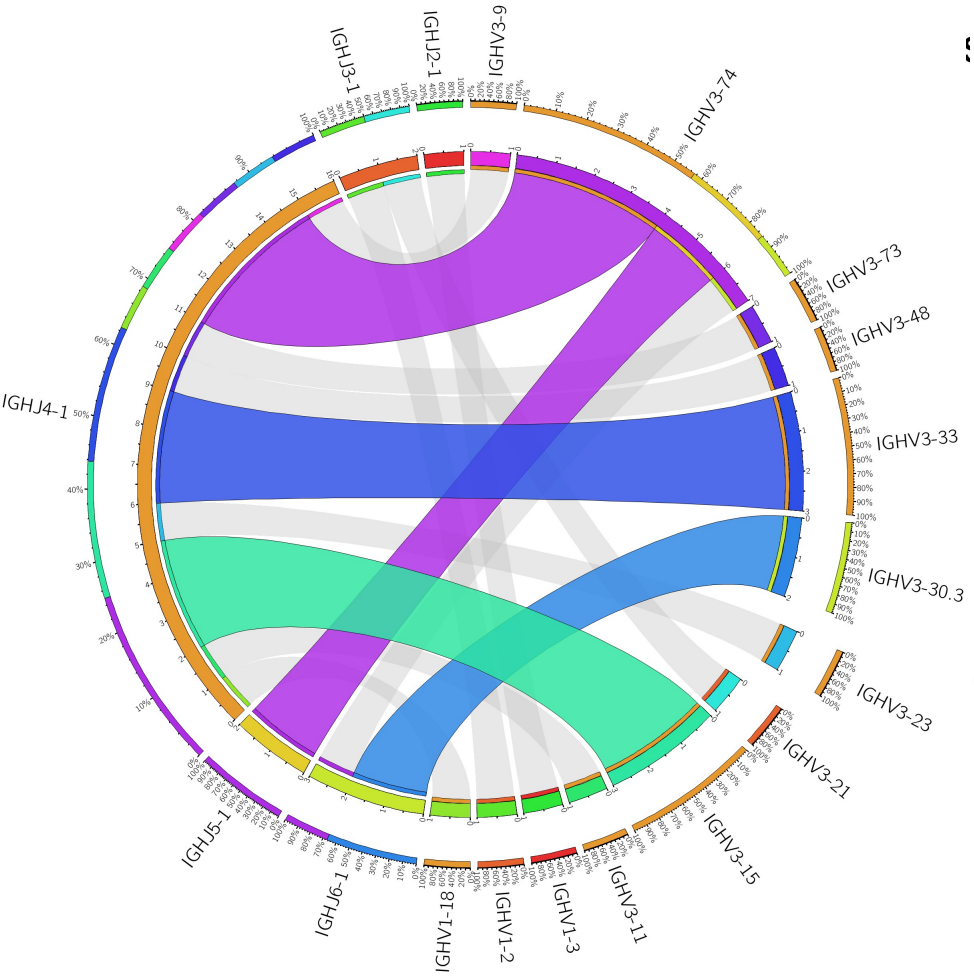
A integrated flow of information for repertoire sequencing projects

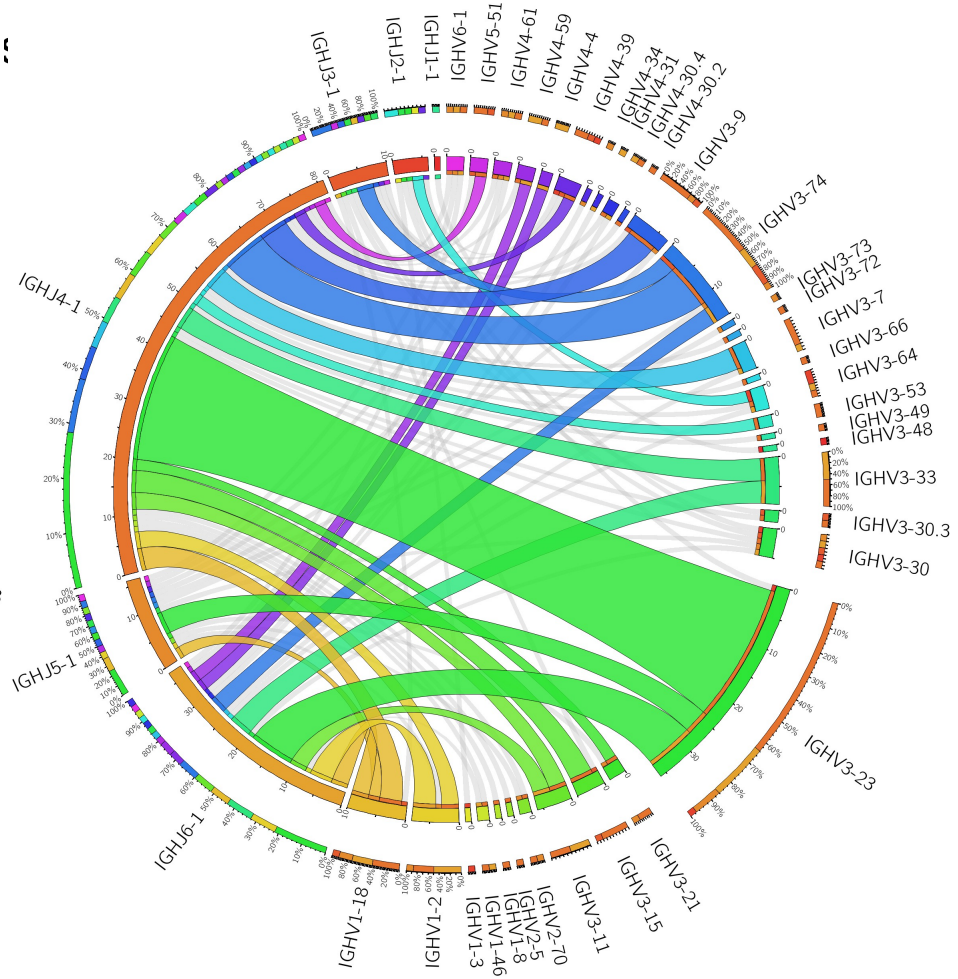


- **Natural infection**:
  - Influenza
  - Rotavirus
  - HIV
  - Hepatitis C virus
  - Cytomegalovirus (CMV), Epstein–Barr virus (EBV)
  - *Staphylococcus aureus*
  - Dengue virus
- **After vaccination**:
  - Influenza
  - Tetanus
  - *Haemophilus influenzae*

Jacob D. Galson , Andrew J. Pollard , Johannes Trück , Dominic F. Kelly - *Studying the antibody repertoire after vaccination: practical applications*. Trends in Immunology, Volume 35, Issue 7, 2014, 319 - 331

Different ways to look at the Ig repertoire



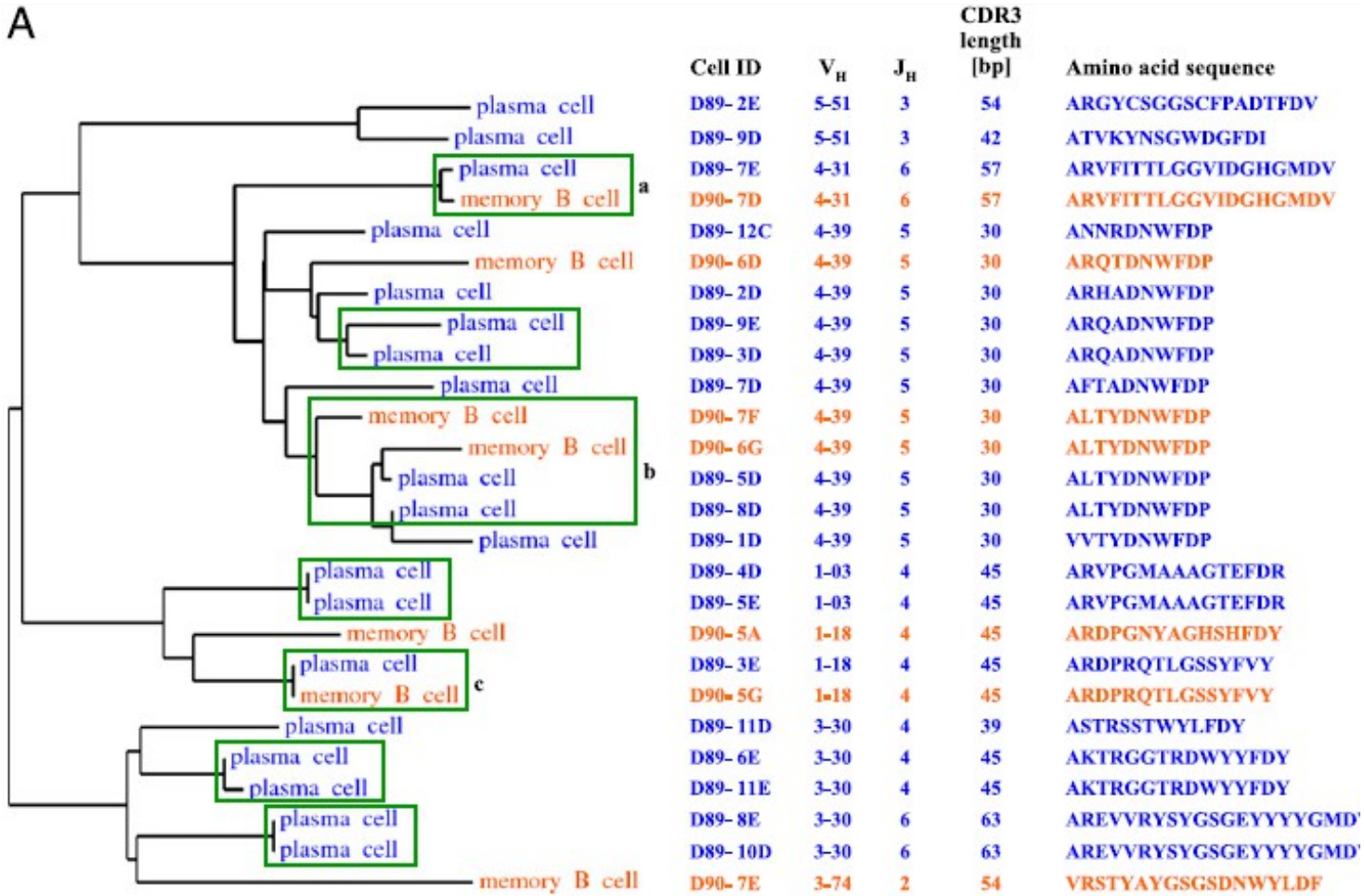


Pre-immune

Post-vaccination

Different ways to look at the Ig repertoire

Phylogenetic analysis



Cell ID	V_H	J_H	CDR3 length [bp]	Amino acid sequence
D89-2E	5-51	3	54	ARGYCSGGSCFPADTFDV
D89-9D	5-51	3	42	ATVKYNSGWDGFDI
D89-7E	4-31	6	57	ARVFITTLGGVIDGHGMDV
D90-7D	4-31	6	57	ARVFITTLGGVIDGHGMDV
D89-12C	4-39	5	30	ANNRDNWFDP
D90-6D	4-39	5	30	ARQTDNWFDP
D89-2D	4-39	5	30	ARHADNWFDP
D89-9E	4-39	5	30	ARQADNWFDP
D89-3D	4-39	5	30	ARQADNWFDP
D89-7D	4-39	5	30	AFTADNWFDP
D90-7F	4-39	5	30	ALTYDNWFDP
D90-6G	4-39	5	30	ALTYDNWFDP
D89-5D	4-39	5	30	ALTYDNWFDP
D89-8D	4-39	5	30	ALTYDNWFDP
D89-1D	4-39	5	30	VVTYDNWFDP
D89-4D	1-03	4	45	ARVPGMAAAGTEFDR
D89-5E	1-03	4	45	ARVPGMAAAGTEFDR
D90-5A	1-18	4	45	ARDPGNYAGHSHFDY
D89-3E	1-18	4	45	ARDPRQTLGSSYFVY
D90-5G	1-18	4	45	ARDPRQTLGSSYFVY
D89-11D	3-30	4	39	ASTRSSTWYLFDY
D89-6E	3-30	4	45	AKTRGGTRDWYYFDY
D89-11E	3-30	4	45	AKTRGGTRDWYYFDY
D89-8E	3-30	6	63	AREVVRYSYGSGEYYYYGMD
D89-10D	3-30	6	63	AREVVRYSYGSGEYYYYGMD
D90-7E	3-74	2	54	VRSTYAYGSGSDNWYLDF

Frolich et al JI, 2010

Different sequencing methods for studying the antibody repertoire

	Sanger sequencing	Deep sequencing
Strategy	Separate reaction for the sequencing of each single B-cell	One single reactions for the simultaneous analysis of the whole repertoire
Use	Study a small number of antigen specific B-cells	Characterizing whole repertoires and their perturbations in time
Pro	High precision	Highly cost-effective and efficient
Cons	Expensive and time-consuming due to limited automation and the high number of different reactions	Interpretation of the abundance of data is challenging

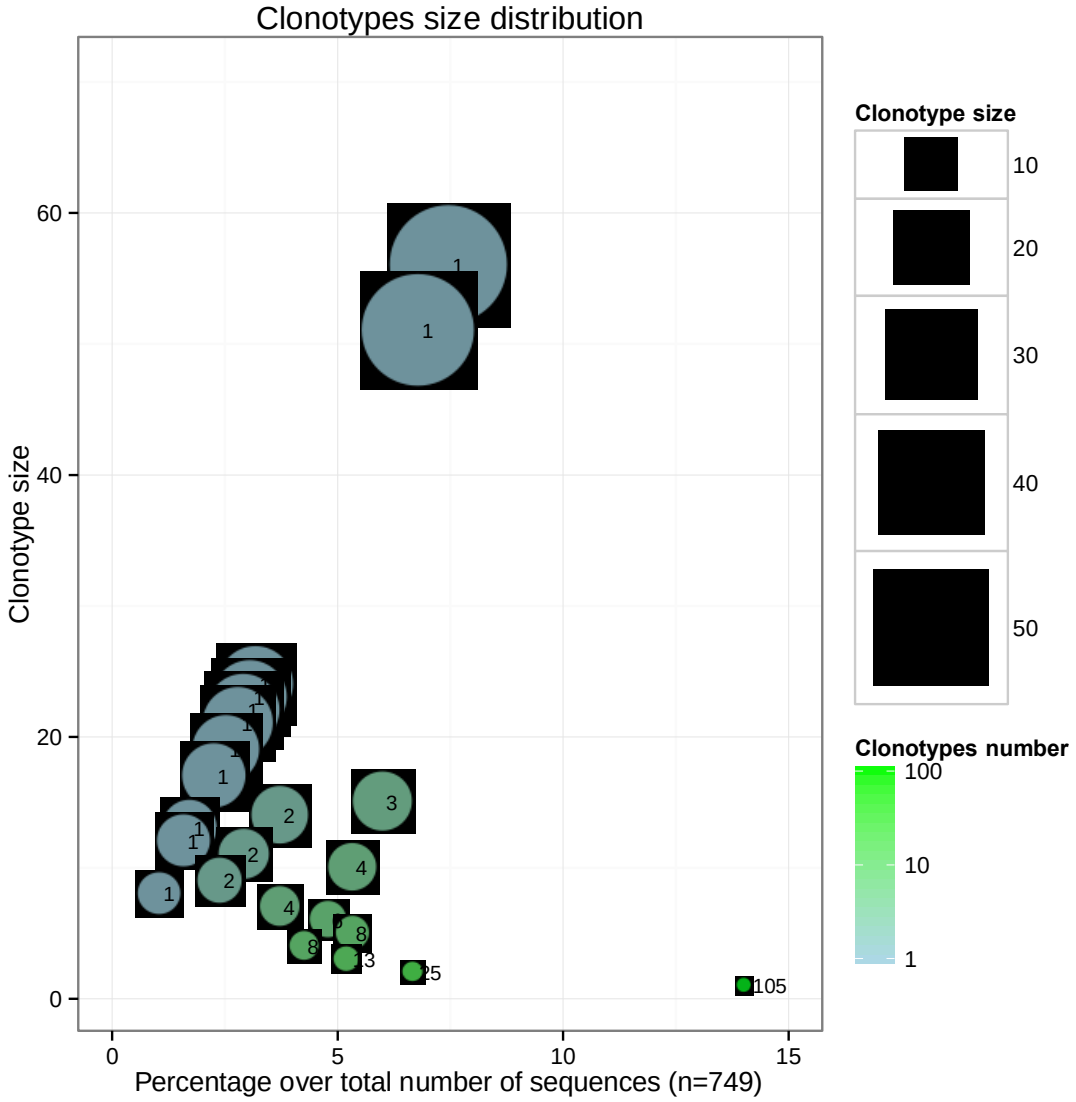
Low-resolution methods for studying the antibody repertoire

- Isoelectric focusing of antibodies in polyacrylamide gels (Eder, J. (1972). *Journal of immunological methods*)
- CDR3 lengths **spectratyping** analysis through a PCR method
- **Sanger sequencing**:
  - Lymphocytes can be isolated, followed by production of immortalized cell lines. **Rearranged immunoglobulin mRNA can then be amplified** from these cell lines, and sequenced
  - **FACS** sorting allows more precise definition of the B cell subsets
  - So far, used to **investigate small numbers of antibodies** and antigen-specific sequences used in response to vaccines against influenza, tetanus, Hib, and some serotypes of *S. pneumoniae*

Lessons learnt through low-resolution methods

- Studies have shown that the B cell **response to an antigen with a simple biochemical structure**, such as Hib polysaccharide:
  - appears markedly **oligoclonal**
  - reveals **similar clones** (predominantly **VH3–23**, with a conserved 'GYGMD' CDR3 amino acid motif) **dominating the repertoire in different individuals**
- Repertoire diversity is also restricted for protein antigens [e.g., tetanus toxoid (TT), and influenza haemagglutinin (HA)], **but the expanded clones differ more between individuals**
- Even in a single individual, there appears to be little similarity in the response after repeated TT vaccination, with **only one-**

Repertoire diversity for tetanus toxoid (TT)



NGS enabled the sequencing of antibody genes from millions of cells simultaneously

- Sanger sequencing for characterization of single antibody genes from small numbers of cells is highly robust, as there is a **high signal-to-noise ratio, low error rates, and a long read length**
- For characterizing whole repertoires of antibody genes from large numbers of cells, **Sanger sequencing is too labor intensive**, so NGS has been used instead
- High-throughput antibody repertoire sequencing was initially conducted using Roche 454, because this was the only platform with a sufficient read length to cover the entire VH region
- Recent advances in the Illumina chemistry now allow 300bp paired-end sequencing on their MiSeq platform

Primary targets of NGS studies

- The first NGS human studies sought:
  - To estimate the **total size** of the antibody repertoire
  - To calculate **frequencies** of V(D)J segment usage
  - To evaluate the **similarity** in V(D)J segment usage between individuals
  - To spot the **CDR3** sequences between individuals
- Size estimates are difficult because samples taken are a very small representation of the entire repertoire
- Sequencing independent replicate libraries in a capture–recapture analysis indicates a minimum bound of 106 unique VH rearrangements

Comparing **different individuals'** repertoires

- V(D)J segments were used in unequal frequencies
- Their **pattern of usage** was **similar** between individuals:
  - inherent biases in the V(D)J recombination process
  - preferential use of core genes
- Despite similar V(D)J usage, there was **limited overlap in the CDR3 repertoires,** hence clonotypes

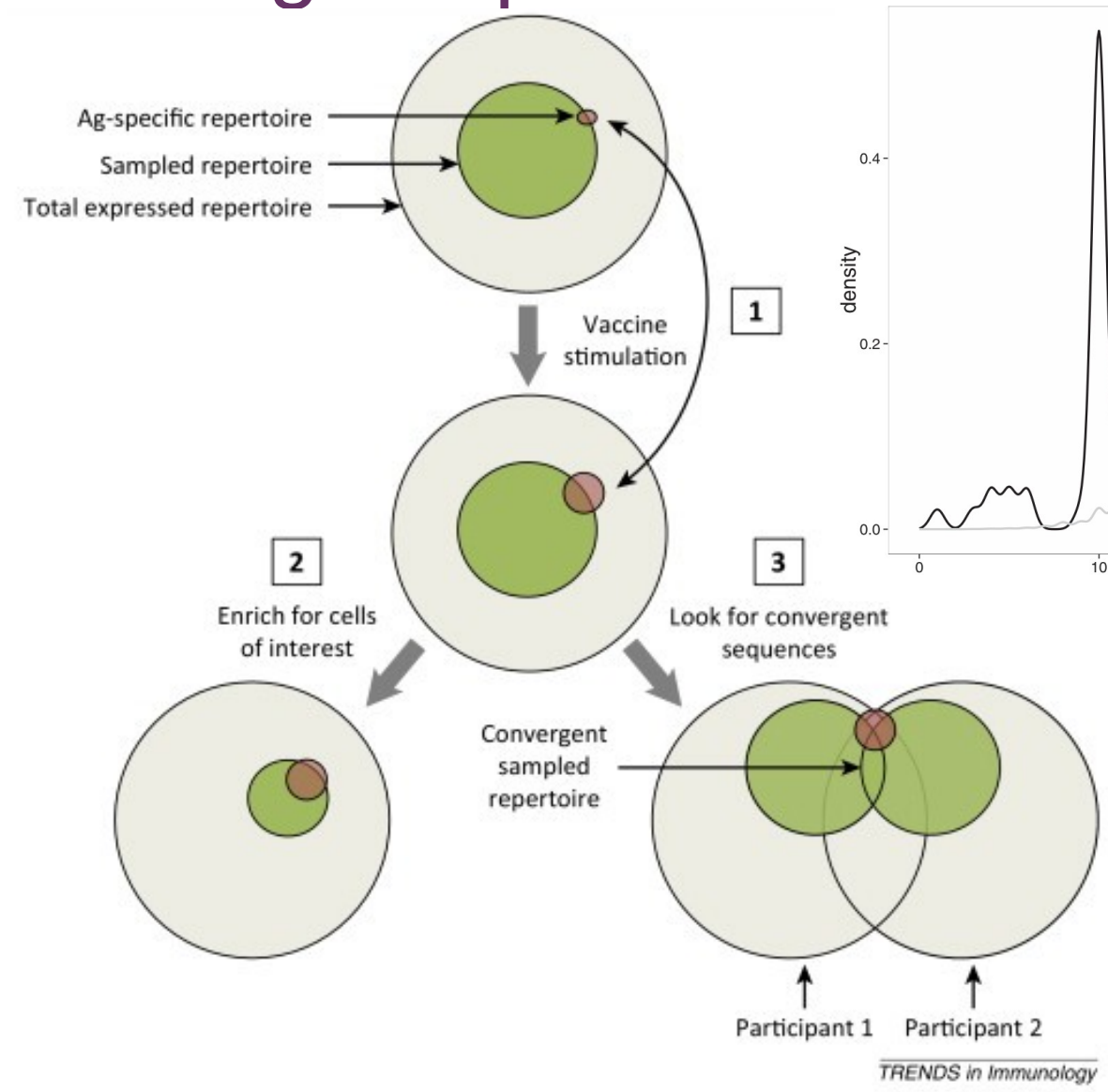
Identification of the antigen-specific repertoire

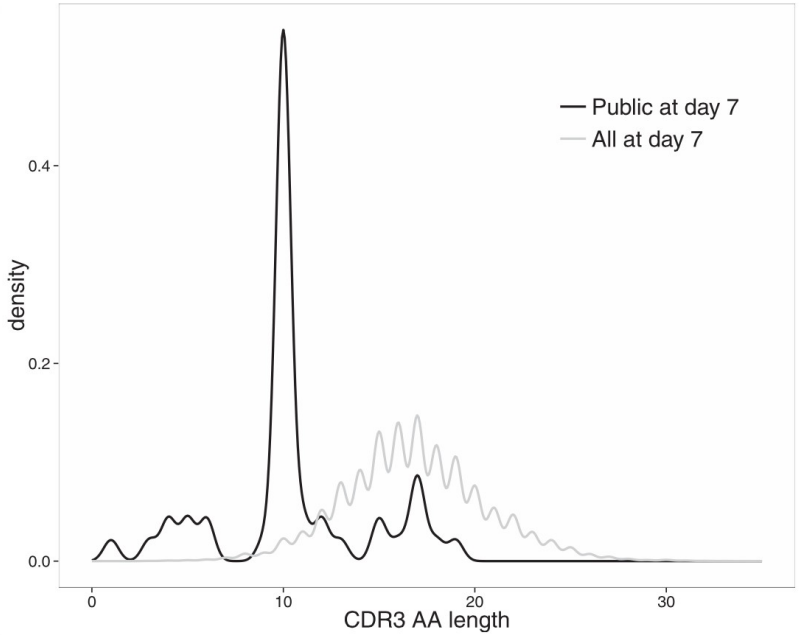
- Although **NGS methods are well suited to studying perturbations in the total sampled repertoire**, to date, the low-resolution methods have provided the most insight into the antigen-specific repertoire after vaccination

- Antigen-specific sequences can be *a priori* identified by spotting the antibody sequences that are shared between multiple individuals following recent exposure (through infection, immunization or autoimmunity) to a common antigen – this is termed the '**convergent repertoire**'

- A study comparing the antibody repertoire in 60 dengue patients during acute disease, and at convalescence, indicated the presence of **convergent CDR3 signatures** (predominantly 'ARLDYYYYYGMDL') between individuals that were **enriched during acute disease** compared to at

Using vaccines to investigate the antigen-specific antibody repertoire





Potential confounding effects: the **genotype**

- By studying the antibody repertoire in two pairs of monozygotic twins, Glanville *at al*. investigated the relative influence of genetic and environmental factors in structuring the naïve repertoire:
  - They found that certain **VH and DH segments were used at significantly different frequencies between the different twins compared to within twin pairs**, indicating that individual genetic differences should be taken into account when studying changes in V(D)J segment usage after vaccination
  - There was limited overlap in the CDR3 repertoires both between and within twins

The potential confounding effects: the **antigen exposure history**

- In addition to previous infections affecting how the repertoire responds to vaccination, the effect of chronic infection needs to be considered
- Wang *et al*. documented chronic CMV and EBV infection status in a repertoire study of 27 individuals over 2 years
  - **CMV infection resulted in increased VH mutation** in the IgG and IgM repertoire
  - **EBV** infection resulted in an increased number of **persistent clonal groups**
  - Neither infection resulted in altered VDJ gene segment usage in the repertoire
- Segment usage does appear to be different also between healthy donors, patients with chronic hepatitis C infection, and

The potential confounding effects: the **age**

- Jiang *et al*. studied the antibody repertoire before and after influenza vaccination in children (8– 17 years), young adults (18–30 years), and elderly individuals (70–100 years):
  - This study did show that **in two of the elderly individuals, the repertoire was highly clonal and had a greater mutational load compared to the younger ones**
- In another study, the most striking differences were in the IgA and IgM repertoires, which displayed **slower clonal expansion as well as less mutation, and longer CDR3 regions in the elderly group**
- Even **with no vaccine stimulation**, there appear to **be age-related differences in the naÏve repertoire**, with **elderly** individuals having **different V(D)J recombination frequencies**, **longer CDR3** regions,

Determining whether sequences derived from the same B-cell clone

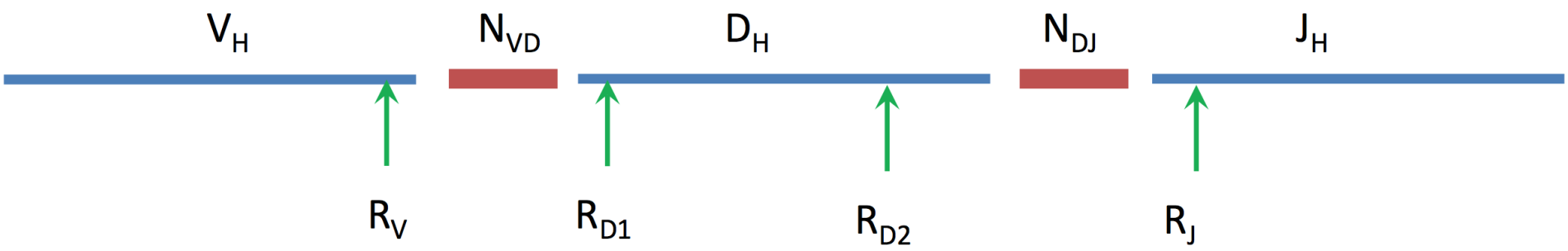
- In different samples of a longitudinal time course
- In different tissue sites, sample types
- B-cell subsets

- Initially, a **clonotype definition** requiring that the **V and J segments be the same** and that **the CDR3 sequence be the same length and match at 80%** of the nucleotides was used

- Quantitation of the contribution of clonally expanded B-cell populations to an observed repertoire can be performed from multiple-replicate library data using a modified form of the Gini-Simpson index adapted for NGS data

Rearrangement Model's Parameters

- Choose Gene Segments: VH, DH, JH
- Choose Recombination points: RV, RD1, RD2, RJ
- Choose N-nucleotide sequences: NVD, NDJ



Take home messages

- Need for standardization:
  - Lack of a standardized laboratory protocol
  - Linking VH and VL is a fundamental step
  - Lack of a standardized analysis pipeline
  - Linking sequence data to functional meaning
- The extent to which an individual's B cell repertoire for an antigen is shared with other individuals remains uncertain and is a key question to answer in order to determine the utility of convergent repertoire analysis
- Vaccine studies are an ideal tool for investigating the degree to which convergent sequences are likely to be antigen-specific

Backup matter

Please find here what doesn’t fit

Individual Variation in the Germline Ig Gene Repertoire Inferred from Variable Region Gene Rearrangements

(Scott D. Boyd et al., J Immunol 2010; 184:6986-6992)

Mean frequency and the range of frequencies of rearrangements of core IGHV genes, n=11