PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Review

Cite this article: Cobey S, Wilson P, Matsen IV FA. 2015 The evolution within us. *Phil. Trans. R. Soc. B* **370**: 20140235.
http://dx.doi.org/10.1098/rstb.2014.0235

Accepted: 18 May 2015

One contribution of 13 to a theme issue 'The dynamics of antibody repertoires'.

Subject Areas:
evolution, computational biology, immunology, ecology

Keywords:
antibody repertoires, B-cell sequencing, vaccine design

Author for correspondence:
Sarah Cobey
e-mail: cobey@uchicago.edu

The evolution within us

Sarah Cobey[1], Patrick Wilson[2,3,4] and Frederick A. Matsen IV[5]

[1]Department of Ecology and Evolution, [2]Department of Medicine, [3]Committee on Immunology, and [4]Knapp Center for Lupus and Immunology Research, University of Chicago, Chicago, IL 60637, USA
[5]Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

FAM, 0000-0003-0607-6025

The B-cell immune response is a remarkable evolutionary system found in jawed vertebrates. B-cell receptors, the membrane-bound form of antibodies, are capable of evolving high affinity to almost any foreign protein. High germline diversity and rapid evolution upon encounter with antigen explain the general adaptability of B-cell populations, but the dynamics of repertoires are less well understood. These dynamics are scientifically and clinically important. After highlighting the remarkable characteristics of naive and experienced B-cell repertoires, especially biased usage of genes encoding the B-cell receptors, we contrast methods of sequence analysis and their attempts to explain patterns of B-cell evolution. These phylogenetic approaches are currently unlinked to explicit models of B-cell competition, which analyse repertoire evolution at the level of phenotype, the affinities and specificities to particular antigenic sites. The models, in turn, suggest how chance, infection history and other factors contribute to different patterns of immunodominance and protection between people. Challenges in rational vaccine design, specifically vaccines to induce broadly neutralizing antibodies to HIV, underscore critical gaps in our understanding of B cells' evolutionary and ecological dynamics.

1. Introduction

By the time many of us give serious thought to passing on our genes, the pathogens circulating on the day of our birth have evolved for thousands of generations—and yet most people are fortunate not to greet parenthood at death's door. We owe this good fortune to the adaptive immune system, and especially B cells. B cells evolve on the time scale of pathogen populations and secrete pathogen-specific antibodies that protect us from infection. But despite their importance to our survival and the unique circumstances of their evolution, the dynamics of B-cell populations remain largely unexplored.

Vaccines and antibody-based therapies saved millions of lives, even before their mechanisms of action were known, and continue to be an essential part of medicine. Edward Jenner is often credited with developing the first vaccine in 1796, when he demonstrated that inoculation with cowpox protected against smallpox infection [1]. Centuries earlier, a similar practice known as 'variolation'—the induction of immunity by a deliberate, attenuated smallpox infection—had spread throughout Central and East Asia [2]. Smallpox vaccination became compulsory in parts of Europe and the USA by the mid-nineteenth century. Serum therapy against diphtheria and tetanus was first demonstrated by Emil von Behring and Shibasaburo Kitasato in 1890 [3]. Several years later, Paul Ehrlich conceived of the side-chain theory, proposing that chemical structures in circulating cells could protect against pathogens [4]. These discoveries led to the rapid adoption of serum therapy against diphtheria in the 1890s and earned Behring, Kitasato and Ehrlich Nobel Prizes [5]. More recently, transfer of specific antibodies to patients infected with respiratory syncytial virus has demonstrated the feasibility of monoclonal antibody therapy during infection [6,7]. Monoclonal antibody therapy has also shown high efficacy against tumour-specific antigens in some cancers [8–12].

The success of antibodies against infections and cancers is offset by a persistent and poorly understood risk of antibody-mediated autoimmunity. Autoreactive

THE ROYAL SOCIETY PUBLISHING

antibodies, or 'horror autotoxicus' as termed by Paul Ehrlich over a hundred years ago [13], are involved in chronic autoimmune disorders that affect 3–4% of the population [14]. The autoimmune pathology is determined by which self-antigen is targeted. Autoantibodies to thyroid-stimulating hormone receptor cause hyperthyroidism in Graves' disease, whereas autoantibodies to thyroid peroxidase or thyroglobulin cause hypothyroidism in Hashimoto's thyroiditis [15]. Ribonucleoprotein and other nuclear antigens are frequent targets in primary Sjögren's syndrome and systemic lupus erythematosus [16,17], conditions associated with widespread inflammation. A frustrating feature of autoimmune disorders involving B cells is that autoantibodies appear without warning, often after infections [18,19]. Although there are predisposing genetic factors [14,18,20], why some people develop autoreactive antibodies while others do not is still mostly a mystery.

Other unexplained differences between individuals' antibody repertoires can have clinical consequences. For instance, in a small fraction of people infected with the human immunodeficiency virus (HIV) [21–23] or vaccinated to influenza H1N1 [24–26], broadly neutralizing antibodies (antibodies effective against a wide diversity of viral strains) dominate the immune response. In HIV, these antibodies can protect animals against challenge [27,28], but their densities are uncorrelated with viral load in humans already infected with the virus [29]. Broadly neutralizing antibodies to influenza may confer protection to potentially pandemic strains [24,30–32]. It is not yet clear whether such antibodies are inducible in everyone. It is also not obvious if broadly neutralizing antibodies will persist over time, or if they will be outcompeted by other responses. Antibody repertoires show influences of infection history [33–35], host age [36–38], host genotype [39,40], and genotypic or phenotypic convergence [41–45], but the relative importance of each is not yet clear.

The diversity of antibody repertoires tracks the ongoing dynamics of B-cell populations in each individual. Clonal selection theory, the idea that cell populations with unique receptors undergo selective expansion upon encounter with antigen, was proposed independently by David Talmage [46] and MacFarlane Burnet [47] in the late 1950s. These hypotheses predate the first observations of affinity maturation [48], or even the observation that B cells are a distinct lineage of lymphocytes [49]. Yet clonal selection is key to the adaptability and survival of vertebrate hosts that frequently encounter new and familiar pathogens.

This extraordinary adaptability arises from B cells' large population sizes, high mutation rates, competition in structured metapopulations, and evolutionary 'checkpoints' [50,51]. It is also enhanced by many generations of natural selection on naive B-cell repertoires. Naive B cells, which have not encountered antigen, show extensive diversity arising from several sources: polymorphism in the receptor-coding genes, recombination between these genes, insertions and deletions [52,53]. The receptor genes themselves show selection for mutational hotspots and coldspots in structurally beneficial areas [54,55]. Precursors of naive B cells are further selected for non-autoreactivity [56]. In sum, B cells are an extraordinary case of evolved evolvability, and thus they pose challenging questions for evolutionary biology.

Recent technical and conceptual advances have enabled detailed investigation of B-cell dynamics. The wide variety of high-throughput sequencing platforms provide different options trading off depth, read length, and accuracy; new methods exist to pair the heavy and light chain sequences that comprise receptors and assess antigen binding properties [57–59]. New methods of B-cell sorting and cloning allow multiple functional assays to be performed on individual cells [60]. Neutralization assays measure the effectiveness of antibodies at prohibiting infections in cell culture, and cross-competition assays reveal the targets of and steric interference with other antibodies. The movements of individual B cells are now directly observable through fluorescent microscopy [61]. As reviewed in detail below, molecular evolutionary models of the evolution of B-cell populations, and ecological models of their competition, are being tested with these new genetic and phenotypic observations.

This review describes what is known about the dynamics of antibody repertoires, and the uncertainties that stand in the way of predictive models and more effective clinical interventions. We first review the antibody response to primary infections and then jump scales, summarizing the patterns reported in natural repertoires. We next survey methods to address the difficult problem of inferring the evolutionary histories of B-cell populations. We also describe studies of competing B-cell populations that take a different approach and offer preliminary explanations for the dynamics of responses over multiple exposures. These concepts converge in the challenge of vaccine design, an area that we predict will benefit from a deeper understanding of B-cell repertoire dynamics.

2. Brief overview of B cells and their evolution

B cells evolve in each individual through their receptors, which are secreted in soluble form as antibodies by some classes of B cells. These receptors, also known as *immunoglobulins*, are Y-shaped proteins composed of four polypeptides: two identical *light chains* and two identical *heavy chains*. The distal parts of these chains (the arms of the Ys) are composed of *variable* and *constant* regions, with the variable regions of each heavy and light chain pair directly binding to antibody. The site at which binding occurs on the antibody or B-cell receptor is the *paratope*; the site bound on the antigen is the *epitope*. The variable region of the heavy chain is approximately 100 amino acids in length. Paratopes are thought to be much smaller, roughly one to two dozen amino acids, and can be located in a variety of locations in the variable region. The tail end of an antibody is called the *Fc region*: this portion does not bind to antigens but rather interacts with the cellular and complement portions of the immune system. Important discoveries have recently been made concerning the Fc region, including its potentially important role in HIV antibodies that were previously thought to operate by blocking viral entry [62,63]. In general, antibody efficacy is not determined solely by the identity of the variable region, although the variable region will be the focus of this issue.

Naive B-cell receptors have extraordinary diversity because of a rearrangement process that incorporates recombination, insertions and deletions. Susumu Tonegawa was awarded the Nobel Prize in 1987 for deciphering the molecular mechanism driving B-cell repertoire diversity [64]. Genes from two loci, the immunoglobulin loci *kappa* and *lambda*, encode light chains. Only one of the two light chain loci is used in a single B cell, whereas heavy chains are encoded by a single locus. The variable regions of both chains arise

from recombination at their respective loci. In humans, each locus (kappa, lambda and heavy) contains several dozen variable (V) genes and four to six joining (J) genes. The heavy chain locus also contains roughly two dozen diversity (D) genes. B-cell receptors are constructed by rearrangement of randomly chosen V and J genes in the light chain and V, D and J genes in the heavy chain, yielding approximately 2×10^6 unique potential VDJ combinations. Variable addition and subtraction of nucleotides at the junctions between gene segments increases the number of unique potential antibodies to 10^{11}, although the number of unique circulating naive B cells at any time is closer to 10^6–10^7 [52,65]. Many immature B cells do not develop to naive cells because they do not generate functional antibodies, or they fail checkpoints for self-reactivity. These cells either further rearrange their receptors, or they apoptose. Germline receptor diversity may arise from slightly different processes, including gene conversion, in some animals [66,67].

B cells further evolve upon exposure to antigen [48,68,69]. This process, known as affinity maturation, involves strong competition and selection for B-cell receptor binding to antigen [70]. Thus, in contrast to the diversity of the naive repertoire, which arises primarily by recombination, B cells exposed to antigen evolve primarily by point mutations (although insertion–deletion mutations occur too [71]). These so-called *somatic hypermutations* are induced by activation-induced cytidine deaminase (AID) [72,73], which preferentially mutates cytosines to uracils [74]. These mutations induce repair activities by DNA polymerases, which may introduce additional mutations while repairing the initial error. The fact that AID recognizes cytosine, particularly in certain motifs, led to the discovery of hotspots and coldspots of mutational activity in variable regions [75]. Hotspots tend to lie in complementarity-determining regions (CDRs), which are involved in binding, and coldspots in framework regions (FWRs), which are thought to be structurally conserved [76]. Antibody variable regions also have biased codon usage, with cytosines favoured in silent sites in CDRs and potential terminal codons in FWRs [76,77]. The B cells descending via somatic hypermutation from a single naive B cell are said to form a set of expanded clones, although different definitions exist [78].

Most B cells undergo affinity maturation in germinal centres, which are aggregations of immune cells that form several days after immunization or the start of an infection. The number of germinal centres that forms after infection is unknown but appears to be highly variable [79]. High-affinity naive B cells enter each germinal centre and compete for antigens presented on the surface of follicular dendritic cells. After binding and removing antigen, B cells present digested antigen peptides to a class of helper T cells known as follicular helper T cells. B cells that are activated by follicular helper T cells undergo replication under the influence of AID. On average, B-cell receptors accumulate one mutation per 10^3 bases, or roughly one mutation per round of replication [80]. These cells with mutated receptors then compete for antigens, and cells that fail to bind to antigens or to receive T cell help apoptose. Each generation takes approximately 8–12 h, and germinal centres probably persist for several weeks (reviewed in [81]).

Responses to primary infection can be complicated by feedback from antibodies [82], competition between different lineages of B cells in different germinal centres, competition between memory and naive responses [83], and B-cell activation and affinity maturation outside germinal centres [84]. Mature B cells can also differentiate or *class switch* to classes that are distinguished by the isotype of antibody produced: B cells secreting the IgG and IgA forms of antibodies, for instance, maintain the affinity of the original B-cell receptor but differ in their constant regions and avidity for antigen (i.e. overall binding rate). Before considering the potential effects of these dynamics, it is worth looking at the final result: naturally occurring antibody repertoires.

3. Observing the natural repertoire

The naive B-cell repertoire deviates significantly from what one would find given a uniform distribution on all of the formational probabilistic events, which is called bias in the literature ([39], reviewed in [44]). For example, individual heavy chain variable genes range in usage frequency from 0.1% to 10% of rearrangements in a repertoire [39,85]. The usage distributions of heavy chain D and J genes and light chain V and J genes are likewise skewed; there is also preferential usage of particular alleles within each gene [65]. These patterns are roughly conserved across individuals and may be intrinsic to rearrangement, although Collins *et al.* [86] show in this issue that VDJ usage may vary within different strains of a species. Some of this bias is explained by differences in recombination signal sequences [87,88] and variation in the number of gene copies between individuals [89,90]. For example, among 35 individuals, copies of particular IGHV1-69 variants ranged from 0 to 4, while the variants' frequencies in repertoires varied correspondingly from 0 to 11% [91]. Another striking and unexplained feature is bias in the DJ pairings on heavy chains [65,92,93]. There is less evidence of intrinsic bias in the pairing of heavy and light chains [44], although negative selection of naive B cells for autoreactivity suggests a mechanism by which such bias might arise [56].

Identifying the sources of these patterns is challenging because of the large number of functional genes, closely related allelic variants and pseudogenes at receptor loci. Although individual heavy chain variable loci contain only a few dozen genes, researchers have identified over 100 putative heavy chain variable genes, pseudogenes, and open reading frames so far, and together they have over 200 alleles [65,89,90]. Bulk high-throughput sequencing provides a useful perspective on natural repertoires. As described in this issue, other techniques, such as hybridomas and B-cell cloning, allow direct experimentation on B cells with a receptor of interest [78].

Antigen-experienced B-cell populations are similarly not a uniform draw from the naive pool. This pattern holds for the long-lived memory B cells produced during affinity maturation and *plasmablasts*, the precursors to antibody-secreting plasma cells showing the active response to infection. Adult recipients of the influenza vaccine have memory responses dominated by a few clonal lineages [37]. Some HIV-infected patients show very old lineages, over 20% diverged from their naive B-cell ancestors, that appear to have coevolved with the founding virus [94,95]. Memory B-cell clones to influenza also show differences in size and divergence from the germline, suggesting that they re-enter germinal centres and undergo further affinity maturation upon exposure to related viral strains [96]. These observations suggest that chronic and

repeat infections might amplify a few clones that were randomly activated early on, leading to uneven distributions of clone sizes.

Three articles in this issue provide further detail on B-cell repertoire development. Although it has been observed that particular genes, like IGHV1–69, frequently appear in broadly neutralizing antibodies to HIV and influenza, Dunand & Wilson [97] argue that responses to many different antigens may be 'canonical.' Antibodies specific to HIV, influenza, *Streptococcus pneumoniae*, rotavirus, self-antigens and even B-cell tumours with unknown specificities are strongly associated with particular receptor genes, implying that particular clonal expansions may be more expected than others. Also in this issue, Martin *et al.* [98] present evidence that gene frequencies change with age. Intriguingly, although age-related changes in antigen exposure and other selective pressures could explain these differences in experienced cells, naive repertoires also have shifts in bias. Disease and other abnormalities might partially explain these changes [98].

While high-throughput sequencing and other technologies are broadening our view of B-cell populations, the resulting data have simultaneously underscored our ignorance. Why are some lineages more frequently seen than others? How can we recognize the members of a clonal family? What selective pressures operate on different B-cell types? Increasingly sophisticated methods of sequence analysis have begun to answer these kinds of questions.

4. Inference under statistical models of B-cell sequence evolution

Statistical models of B-cell evolution use sequences to reconstruct the unobserved evolutionary trajectories of populations. The complexity of B-cell evolution, however, limits the utility of simple methods for two reasons. First, inferring rearrangements in heavy and light chain genes is difficult. The number of heavy chain variable genes on each chromosome is unknown, and some of these genes may be duplicates or previously unreported alleles. Junctional diversity is high from random insertions and deletions. Second, somatic hypermutation during affinity maturation is biased by the presence of motifs that direct mutations to particular sites in different genes [75]. Moreover, measuring the precise effects of these motifs is difficult because of mutational saturation and epistasis [55]. High-throughput B-cell sequencing further requires methods that tolerate large collections of sequences and unpaired heavy and light chain sequences.

Identifying the VDJ rearrangement that formed a naive B-cell receptor is important for measuring recombinatorial biases and total allelic diversity, but this task is complicated by the potential for mutation, deletion and insertion to give identical results, as well as by the similarity of some germline genes. Although general-purpose tools can be used productively and efficiently to solve this problem [99,100], others, including iHMMune-Align [101], SoDA2 [102], and two in this issue [103,104], evaluate B-cell rearrangement events statistically. Frost *et al.* [103] present Immunoglobulin Subtype Classification Using Evolutionary Algorithms (IgSCUEAL), which uses maximum likelihood to assign heavy chain V and J regions to germline alleles, taking into account the phylogenies of V and J gene families and uncertainty from undetected alleles. This method tolerates B cells that have undergone additional somatic mutation [105,106]. In a complementary vein, Elhanati *et al.* [104] use a large dataset of non-productive naive B-cell sequences to infer the probabilities of specific VDJ rearrangements and segment-specific rates of insertions and deletions. These detailed models could in principle be used for the next generation of VDJ inferential procedures.

A related challenge is how to infer *clonal families*, i.e. the members of an expanded clone. As a starting point for clonal family inference, one could use the best nucleotide-by-nucleotide annotations inferred by methods in the previous paragraph. However, the resulting estimates are uncertain because there are many ways to make a given B-cell receptor sequence. Current practice is to use single-linkage clustering [37,107,108] or a distance-based cut-off to an inferred ancestral sequence [109]. The development of statistical models to identify clonal families, with corresponding uncertainty estimates, would help to quantify the structure of B-cell populations and test hypotheses about their evolution.

The peculiar features of somatic hypermutation and B-cell sampling make inference of evolutionary characteristics of B-cell populations a non-trivial task. McCoy *et al.* [110] developed a new method for inferring per-residue selection pressure in the presence of context-sensitive mutation and applied it to a dataset of 15 million B-cell receptor sequences. Like Elhanati *et al.* [104], they used non-productive sequences as a neutral model to infer rates of mutation without the confounding effects of selection for affinity maturation. Elhanati *et al.* [104] also find support for an evolutionary model with heterogeneous substitution rates across sites that takes into account context, such as the presence of certain sequence motifs. This context appears vital to explain patterns of substitutions across variable regions [55,75,76].

These analyses have unveiled surprising features of B-cell evolution. In particular, the studies of both [104,110] find that evolutionary parameters are highly non-uniform and consistent between individuals: in the former paper for rearrangement (e.g. gene deletion lengths) and in the latter for per-residue natural selection estimates. Remarkably, there is a high correlation between the probability that a sequence is generated by recombination and its probability, once generated, of being selected for the naive repertoire. This result suggests a mechanism by which evolution has increased the chance that useful B-cell receptor sequences will be generated in the first place.

We predict that the development of statistical and molecular evolutionary methods will clarify the dynamics of affinity maturation. These methods could in principle be used to estimate the number of unique clones, which is currently estimated with heuristics rather than statistics [78]; the ages of lineages; and the neutral and selective forces shaping the evolution of different B-cell types. For instance, examining clonal phylogenies, Yaari *et al.* [111] propose that the phylogenies of B-cell families reflect multiple time scales of positive selection. Following trends in infectious disease research [112,113], models could eventually integrate phenotypic information, such as the affinities and specificities of receptors for different epitopes. Another possible direction is to use structured coalescent approaches to infer rates of idiotype switching [114]. For now, there is a gap between repertoire analyses that are sequence-based and others that model the competitive dynamics and phenotypic evolution of clonal populations.

5. Growth and competition among the clones

Why are some B-cell lineages more abundant than others, and why are some epitopes targeted more often? This is known as the problem of immunodominance—the preferential targeting by the immune system of certain antigenic sites over others (ch. 6 of [115]), and its importance was recognized long before the development of modern methods of repertoire profiling (e.g. [116–118]). Differences in immunodominance may explain variation in susceptibility between people to the same pathogen. But how does it arise?

Repertoire diversity in the first exposure to antigen is regulated by opposing forces of poorly known strength. Affinity maturation acts on only a subset of the naive repertoire, but from this initial pool, somatic hypermutation introduces mutations that increase genetic and phenotypic variation (phenotype here refers to the affinity and specificity of B cells to various epitopes). Strong competition between B cells for both antigen and T cell help inside germinal centres discards many mutations and promotes the growth of high-affinity cells. In this issue, Childs *et al.* [119] show that parallel evolution in multiple germinal centres promotes adaptation to multiple epitopes. More complex antigens can compromise this response. In addition to providing a null model for immunodominance, Childs *et al.* [119] show that competition between antibodies and B cells inside germinal centres attenuates the bias in epitope targeting and increases the breadth of the repertoire. This model thus suggests that the immune system may have evolved to optimize a trade-off between mounting an effective response to any one epitope versus many, a trade-off that may be exploitable by pathogens. In related work, Mayer *et al.* [53] propose that diverse repertoires are especially advantageous when pathogens have high growth rates, and early recognition of any epitope can attenuate the severity of an infection.

Several examples suggest that intrinsic differences in epitopes' attractiveness can influence immunodominance. The antibody response to *Staphylococcus aureus* is highly constricted: the vast majority of plasmablasts target SpA, a bacterial virulence factor [120]. Because antibodies to SpA are non-neutralizing, this immunogenic and immunodominant epitope, or 'superantigen,' may be an effective means for *S. aureus* to escape immunity. Immunodominant responses to influenza may arises similarly. Epitopes on the globular head of the influenza haemagglutinin are more accessible and thus more immunogenic than epitopes on the recessed stalk, potentially explaining the preponderance of antibodies to the head despite broader protection associated with anti-stalk antibodies [121]. It has been speculated that influenza may evolve immunogenic decoy epitopes, far from neutralizing epitopes, that may sterically interfere with neutralizing antibodies [122]. Pauli *et al.* [120] observed that most anti-SpA antibodies in their experiments derived from the IGVH3 idiotype, which suggests that epitopes' immunogenicity might account for not only the phenotypic but also the genotypic diversity of the repertoire.

Studies of repeated exposures to similar antigens reveal contrasting observations: repertoires show signatures of both contingent and convergent evolution. In contingent evolution, chance events, like mutations or infections with other pathogens, affect the induced repertoire. In 1960, Thomas Francis Jr., proposed that contingent evolution is at work in adaptive immunity: that a first infection with influenza might indelibly shape responses to later strains, a phenomenon he dubbed 'Original Antigenic Sin' [123]. This hypothesis is consistent with recent research: people indeed have especially high antisera to influenza strains circulating in early childhood [36], and researchers have investigated potential mechanisms [124]. These early exposures may entrain responses in ways that are clinically and epidemiologically important. For example, two recent studies have shown that the specific epitopes targeted on H1N1 strains depend on hosts' year of birth [26,35]. People born between 1965 and 1979, for instance, frequently have antibodies targeting a site that mutated in 2013. Linderman *et al.* [35] showed that this mutation leads to loss of antibody binding, potentially explaining the increased incidence of influenza-like illness in this cohort in the 2013–2014 influenza season. These age-specific patterns of epitope binding were recapitulated by sequentially immunizing ferrets with different H1N1 strains, demonstrating that the order in which strains are encountered can explain some of the differences between cohorts [35]. Interestingly, ferrets with the same sequence of immunizations did not always develop immunodominant responses to the same sites: a minority of ferrets (three of eight) had the particular antibody under investigation. Similarly, only 40% of humans born between 1965 and 1979 had this antibody. Thus, even though similar infection histories may lead to similar patterns of immunodominance between individuals, differences remain. By contrast, seemingly diverse responses may eventually converge via evolution, as seen in dengue [42], HIV [43] and influenza [41,45,96,125]. Zarnitsyna *et al.* [126] in this issue investigate competition between B cells in a setting that models antigenic drift and shift in influenza, including multiple antigenic sites. The free variables in this model are population sizes of free and bound antigen on one hand and B cells with various affinities on the other. This set-up allows them to quantify the effect of epitope masking and derive a ratio quantifying original antigenic sin.

It is tempting to assume that experienced repertoires to fast-mutating pathogens reflect intimate coevolution, with adaptation by B cells followed by pathogen escape. Although individual B-cell clones show traces of coevolution with HIV [95], Luo & Perelson [127] argue in this issue that there is little evidence linking repertoires to specific selective pressures or to viral adaptation. Also in this issue, Hoehn *et al.* [128] find no observable impact of antiretroviral therapy on the genetic structure of B-cell populations in HIV-positive patients over time. This is surprising, as one might expect shifts in clonal size distributions to accompany therapy-induced changes in the viral population. In the same vein, Luo & Perelson [127] note that the emergence of detectable broadly neutralizing antibodies to HIV has not been correlated with bottlenecks in the viral population.

Predicting immunodominance, or the outcome of competition between evolving B-cell populations, appears limited by the data available to fit models. Although changes in population size and selective pressures can to some extent be inferred from sequences, and sequences can be related to antigen-specific serum responses [57], it is currently impossible to obtain high-throughput information about receptors' precise affinities to different epitopes [129]. Highly related sequences may have similar affinities [130], but it is difficult to imagine how a thorough understanding of clonal competition and evolution could be acquired without more precise information on B cells' phenotypes. Rates of antibody secretion and clearance, which vary between tissues, also

Box 1. Future directions/major questions about repertoire dynamics.

Future directions

— Measurement of genetic variation in people and model organisms at B-cell receptor loci.
— Models of germinal centre dynamics that incorporate more types of data, such as B-cell receptor sequences, expression information [138], antigen availability and B-cell position.
— Phylodynamics models to evaluate spatial dynamics in germinal centres and statistical models of evolutionary descent.
— Improved models of B-cell memory formation and recall, especially those that infer the amount of competition between memory and naive responses for entry into germinal centres and between secreted antibodies and affinity-maturing B cells.
— Development of phylogenetic methodology specialized to the intricacies of B-cell receptor sequence evolution.
— Measurements of epitopes' relative immunogenicities across individuals.
— Between-species comparative analysis, especially for vaccine model organisms such as ferrets and macaques.
— Variation of B-cell response across human subpopulations, especially in response to shared exposures such as vaccines.
— Specific impacts of autoimmune checkpoints on the evolution of naive and experienced repertoires.
— Diversity and evolution of germline genes among vertebrates (i.e. evolution of presence-absence).
— Better understanding of the effects of age and co-infection, in particular, for autoimmunity and allergies.

Questions

— How can we approximate the genotype to phenotype map of B-cell receptors [139]?
— What are good models of sequence-based fitness landscapes for B-cell receptors? Are pairwise interactions between sites enough, as found by the Ising versus Potts analysis in Mann *et al*. [140]?
— How does T cell help impact the general dynamics of affinity maturation and the selective pressures on specific clones?
— How do the general dynamics of affinity maturation differ between individuals and change with age?
— When two genetically identical and naive hosts are immunized to the same antigen, how do their repertoires differ genetically and phenotypically? How would differences in their naive repertoires, chance recruitment of naive B cells to the response, stochastic dynamics of affinity maturation and other factors contribute?
— Can we use immune information to infer asymptomatic infections?
— Can we relate sequences from sampled repertoires to protection?
— Can we use germline gene loci or a sample of the naive repertoire to predict an individual's responsiveness to a vaccine [141]?
— How is vaccine responsiveness affected by immune memory to other antigens?
— Can immune systems across individuals be classified into meaningful types, and can we use immune 'type' information for stratified sampling in clinical trials?
— Holding infection history constant, are differences in B-cell repertoires important for pathogen evolution [142]?

affect the measured concentration of serum antibodies. Fitting models to these kinds of data could shed light on dynamics that are difficult to observe directly, such as the rules for memory B-cell and plasmablast generation, the rate at which memory B cells pre-empt naive B cells during repeated exposures, and the kinds of competition between secreted antibodies and evolving B-cell clones.

6. Vaccines to direct evolution

It is notoriously difficult to effectively vaccinate against antigenically variable pathogens. Such pathogens may evolve escape mutations, or vaccination may select for escape mutants already present at low frequencies. We have also seen in the previous sections that vaccine design is not a simple process of injecting an antigen and getting out a high-affinity antibody. Challenges in vaccine design are fundamental questions about the ecological and evolutionary dynamics of B cells and their coevolving pathogens.

One such challenge is to evolve high-affinity broadly neutralizing antibodies to an antigen that is poorly bound by naive B cells. For example, broadly neutralizing antibodies in the VRC01 group bind to the gp120 protein of HIV and descend from the IGHV1–69 family; although VRC01 antibodies are diverse on the level of amino acids, they form similar binding sites [43]. These broadly neutralizing antibodies tend to be highly diverged from their unmutated ancestors, requiring months to years to appear [95]. Gao *et al.* [131] showed that mutations in HIV induced by one B-cell lineage enabled binding by another lineage, which ultimately developed broadly neutralizing properties. Given the vast space of possible mutations, it is perhaps unsurprising that diverged antibodies with specific properties appear to arise rarely, and under seemingly contrived circumstances. Identifying the antigens that bind precursors of broadly neutralizing antibodies, and incrementally evolving lineages through sequential immunizations with these antigens, is an active area of research [132,133]. A more tactical evolutionary approach might be to infer the mutations that limit adaptation, such as those with high fitness cost, and to focus immunogen design on selecting viable solutions. If crucial mutations are rare, it may also be beneficial to increase the size of the relevant B-cell population. Finally, the probability of evolving an antibody with particular properties might depend on host genotype: would people without particular IGVH1–69 alleles not benefit from vaccines

to induce VRC01, or can some B-cell phenotypes be evolved from many different gene families?

Another challenge is to ensure that protective B-cell populations induced by vaccination remain at protective levels. McGuire *et al*. [134] investigated how narrowly neutralizing antibodies outcompete broadly neutralizing antibodies after immunization with a specially engineered HIV antigen. Paradoxically, they found that the narrowly neutralizing antibodies had germline ancestors that bound a wider array of antigens and were thus more likely to be stimulated. Does their result imply that once induced, broadly neutralizing antibodies—which by definition bind diverse antigens—should dominate? Broadly neutralizing antibodies can be thought of as having, on average, more available antigen than narrowly neutralizing antibodies. The idea that antibodies to abundant antigens, such as conserved epitopes, should proliferate more than antibodies to specific epitopes has been used to explain patterns of immunodominance in response to malaria [135] and is consistent with basic principles of population growth. However, the subdominance of anti-stalk antibodies in influenza suggests that structure complicates simple assessments of antigen availability. In addition, the factors determining antigen presentation of viral proteins to B cells [136], and how this presentation shapes immunogenicity [137], are not well understood. Considering likely differences in epitope immunogenicity, a larger concern is thus whether the immediate response to the vaccine might eventually be eroded by competition with pre-existing or other induced responses. It is unclear what strategies may help avoid this, or even anticipate its occurrence.

B-cell repertoires lie at the intersection of immunology and evolution, supplying intriguing questions for researchers in both, and other, fields (box 1). Addressing these questions would go a long way to measuring protection and predicting response. Understanding the dynamics of antibody repertoires is a first step.

Authors' contributions. All the authors wrote the paper and approved the final version.

Competing interests. We declare we have no competing interests.

Funding. S.C. is supported by NIH award DP2AI117921 and the James S. McDonnell Foundation. P.W. acknowledges funding from NIH/NIAID grants no. 2U19AI082724, 1P01AI097092, 2U19AI057266, 1U19AI109946, 5U19AI090023, HHSN272201400005C, and the James S. McDonnell Foundation. F.M. is supported by NIH R01 GM113246-01, by the University of Washington eScience Institute through its Seed Grants programme in Translational Health Sciences, and by the University of Washington Center for AIDS Research (CFAR), an NIH funded programme under award number P30AI027757 which is supported by the following NIH Institutes and Centers (NIAID, NCI, NIMH, NIDA, NICHD, NHLBI, NIA, NIGMS, NIDDK).

Acknowledgements. We thank Greg Dwyer and Duncan Ralph for comments.

References

1. Jenner E. 1801 *An inquiry into the causes and effects of the variolae vaccinae, a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox*. London, UK: DN Shury.
2. Fenner F, Henderson DA, Arita I, Jezek Z, Ladnyi ID, World Health Organization. 1988 *Smallpox and its eradication*. Geneva, Switzerland: World Health Organization.
3. Von Behring E, Kitasato S. 1890 Über das zustandekommen der diphtherie-immunität und der tetanus-immunität bei thieren. Dtsch Med. Wochenschr. [Internet]. Philipps-Universität Marburg. See http://archiv.ub.uni-marburg.de/eb/2013/0164/.
4. Ehrlich P. 1897 *Die Wertbemessung des Diphtherieheilserums und deren theoretische Grundlagen*. Jena, Germany: G. Fischer.
5. Winau F, Winau R. 2002 Emil von Behring and serum therapy. *Microbes Infect.* **4**, 185–188. (doi:10.1016/S1286-4579(01)01526-X)
6. Tempest PR, Bremner P, Lambert M, Taylor G, Furze JM, Carr FJ, Harris WJ. 1991 Reshaping a human monoclonal antibody to inhibit human respiratory syncytial virus infection *in vivo*. *Biotechnology* **9**, 266–271. (doi:10.1038/nbt0391-266)
7. The IMpact-RSV Study Group. 1998 Palivizumab, a humanized respiratory syncytial virus monoclonal antibody, reduces hospitalization from respiratory syncytial virus infection in high-risk infants. *Pediatrics* **102**, 531–537. (doi:10.1542/peds.102.3.531)
8. McLaughlin P *et al*. 1998 Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program. *J. Clin. Oncol.* **16**, 2825–2833.
9. Slamon DJ *et al*. 2001 Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med.* **344**, 783–792. (doi:10.1056/NEJM200103153441101)
10. Keating MJ *et al*. 2002 Therapeutic role of alemtuzumab (Campath-1H) in patients who have failed fludarabine: results of a large international study. *Blood* **99**, 3554–3561. (doi:10.1182/blood.V99.10.3554)
11. Karapetis CS *et al*. 2008 K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N. Engl. J. Med.* **359**, 1757–1765. (doi:10.1056/NEJMoa0804385)
12. Yang K, Wang Y-J, Chen X-R, Chen H-N. 2010 Effectiveness and safety of bevacizumab for unresectable non-small-cell lung cancer: a meta-analysis. *Clin. Drug Invest.* **30**, 229–241. (doi:10.2165/11532260-000000000-00000)
13. Silverstein AM. 2001 Autoimmunity versus horror autotoxicus: the struggle for recognition. *Nat. Immunol.* **2**, 279–281. (doi:10.1038/86280)
14. Marrack P, Kappler J, Kotzin BL. 2001 Autoimmune disease: why and where it occurs. *Nat. Med.* **7**, 899–905. (doi:10.1038/90935)
15. Endo K, Kasagi K, Konishi J, Ikekubo K, Okuno T, Takeda Y, Mori T, Torizuka K. 1978 Detection and properties of TSH-binding inhibitor immunoglobulins in patients with Graves' disease and Hashimoto's thyroiditis. *J. Clin. Endocrinol. Metab.* **46**, 734–739. (doi:10.1210/jcem-46-5-734)
16. Tan EM, Schur PH, Carr RI, Kunkel HG. 1966 Deoxybonucleic acid (DNA) and antibodies to DNA in the serum of patients with systemic lupus erythematosus. *J. Clin. Invest.* **45**, 1732–1740. (doi:10.1172/JCI105479)
17. Alexander EL, Provost TT. 1981 Ro (SSA) and La (SSB) antibodies. *Semin. Immunopathol.* **4**, 253–273. (doi:10.1007/BF01892181)
18. Salminen K *et al*. 2003 Enterovirus infections are associated with the induction of β-cell autoimmunity in a prospective birth cohort study. *J. Med. Virol.* **69**, 91–98. (doi:10.1002/jmv.10260)
19. Barzilai O, Ram M, Shoenfeld Y. 2007 Viral infection can induce the production of autoantibodies. *Curr. Opin. Rheumatol.* **19**, 636–643. (doi:10.1097/BOR.0b013e3282f0ad25)
20. Fernando MMA, Stevens CR, Walsh EC, De Jager PL, Goyette P, Plenge RM, Vyse TJ, Rioux JD. 2008 Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet. Public Libr. Sci.* **4**, e1000024. (doi:10.1371/journal.pgen.1000024)
21. Dhillon AK *et al*. 2007 Dissecting the neutralizing antibody specificities of broadly neutralizing sera from human immunodeficiency virus type 1-infected donors. *J. Virol.* **81**, 6548–6562. (doi:10.1128/JVI.02749-06)
22. Li Y *et al*. 2007 Broad HIV-1 neutralization mediated by CD4-binding site antibodies. *Nat. Med.* **13**, 1032–1034. (doi:10.1038/nm1624)

23. Binley JM *et al*. 2008 Profiling the specificity of neutralizing antibodies in a large panel of plasmas from patients chronically infected with human immunodeficiency virus type 1 subtypes B and C. *J. Virol.* **82**, 11 651–11 668. (doi:10.1128/JVI.01762-08)
24. Wrammert J *et al*. 2011 Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. *J. Exp. Med.* **208**, 181–193. (doi:10.1084/jem.20101352)
25. Corti D *et al*. 2010 Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine. *J. Clin. Invest.* **120**, 1663–1673. (doi:10.1172/JCI41902)
26. Li G-M *et al*. 2012 Pandemic H1N1 influenza vaccine induces a recall response in humans that favors broadly cross-reactive memory B cells. *Proc. Natl Acad. Sci. USA* **109**, 9047–9052. (doi:10.1073/pnas.1118979109)
27. Balazs AB, Chen J, Hong CM, Rao DS, Yang L, Baltimore D. 2012 Antibody-based protection against HIV infection by vectored immunoprophylaxis. *Nature* **481**, 81–84. (doi:10.1038/nature10660)
28. Shingai M *et al*. 2014 Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques. *J. Exp. Med.* **211**, 2061–2074. (doi:10.1084/jem.20132494)
29. Lambotte O *et al*. 2009 Heterogeneous neutralizing antibody and antibody-dependent cell cytotoxicity responses in HIV-1 elite controllers. *AIDS* **23**, 897–906. (doi:10.1097/QAD.0b013e328329f97d)
30. Hancock K *et al*. 2009 Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. *N. Engl. J. Med.* **361**, 1945–1952. (doi:10.1056/NEJMoa0906453)
31. Sui J, Sheehan J, Hwang WC, Bankston LA, Burchett SK, Huang C-Y, Liddington RC, Beigel JH, Marasco WA. 2011 Wide prevalence of heterosubtypic broadly neutralizing human anti-influenza A antibodies. *Clin. Infect. Dis.* **52**, 1003–1009. (doi:10.1093/cid/cir121)
32. Henry Dunand CJ *et al*. 2015 Preexisting human antibodies neutralize recently emerged H7N9 influenza strains. *J. Clin. Invest.* **125**, 1255–1268. (doi:10.1172/JCI74374)
33. Skountzou I *et al*. 2010 Immunity to pre-1950 H1N1 influenza viruses confers cross-protection against the pandemic swine-origin 2009 A (H1N1) influenza virus. *J. Immunol.* **185**, 1642–1649. (doi:10.4049/jimmunol.1000091)
34. Li Y *et al*. 2013 Immune history shapes specificity of pandemic H1N1 influenza antibody responses. *J. Exp. Med.* **210**, 1493–1500. (doi:10.1084/jem.20130212)
35. Linderman SL *et al*. 2014 Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013–2014 influenza season. *Proc. Natl Acad. Sci. USA* **111**, 15 798–15 803. (doi:10.1073/pnas.1409171111)
36. Lessler J *et al*. 2012 Evidence for antigenic seniority in influenza A (H3N2) antibody responses in southern China. *PLoS Pathog.* **8**, e1002802. (doi:10.1371/journal.ppat.1002802)
37. Jiang N *et al*. 2013 Lineage structure of the human antibody repertoire in response to influenza vaccination. *Sci. Transl. Med.* **5**, 171ra19. (doi:10.1126/scitranslmed.3004794)
38. Wang C *et al*. 2013 Effects of aging, cytomegalovirus infection, and EBV infection on human B cell repertoires. *J. Immunol.* **192**, 603–611. (doi:10.4049/jimmunol.1301384)
39. Glanville J *et al*. 2011 Naive antibody gene-segment frequencies are heritable and unaltered by chronic lymphocyte ablation. *Proc. Natl Acad. Sci. USA* **108**, 20 066–20 071. (doi:10.1073/pnas.1107498108)
40. Wang C *et al*. 2015 B-cell repertoire responses to varicella-zoster vaccination in human identical twins. *Proc. Natl Acad. Sci. USA* **112**, 500–505. (doi:10.1073/pnas.1415875112)
41. Krause JC, Tsibane T, Tumpey TM, Huffman CJ, Briney BS, Smith SA, Basler CF, Crowe JE. 2011 Epitope-specific human influenza antibody repertoires diversify by B cell intraclonal sequence divergence and interclonal convergence. *J. Immunol.* **187**, 3704–3711. (doi:10.4049/jimmunol.1101823)
42. Parameswaran P *et al*. 2013 Convergent antibody signatures in human dengue. *Cell Host Microbe* **13**, 691–700. (doi:10.1016/j.chom.2013.05.008)
43. Zhou T *et al*. 2013 Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* **39**, 245–258. (doi:10.1016/j.immuni.2013.04.012)
44. Jackson KJL, Kidd MJ, Wang Y, Collins AM. 2013 The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. *Front. Immunol.* **4**, 263. (doi:10.3389/fimmu.2013.00263)
45. Trück J, Ramasamy MN, Galson JD, Rance R, Parkhill J, Lunter G, Pollard AJ, Kelly DF. 2015 Identification of antigen-specific B cell receptor sequences using public repertoire analysis. *J. Immunol.* **194**, 252–261. (doi:10.4049/jimmunol.1401405)
46. Talmage DW. 1959 Immunological specificity, unique combinations of selected natural globulins provide an alternative to the classical concept. *Science* **129**, 1643–1648. (doi:10.1126/science.129.3364.1643)
47. Burnet M. 1959 *The clonal selection theory of acquired immunity*. Cambridge, UK: Cambridge University Press.
48. Eisen HN, Siskind GW. 1964 Variations in affinities of antibodies during the immune response. *Biochemistry* **3**, 996–1008. (doi:10.1021/bi00895a027)
49. Cooper MD, Peterson RD, Good RA. 1965 Delineation of the thymic and bursal lymphoid systems in the chicken. *Nature* **205**, 143–146. (doi:10.1038/205143a0)
50. Zhang J, Shakhnovich EI. 2010 Optimality of mutation and selection in germinal centers. *PLoS Comput. Biol.* **6**, e1000800. (doi:10.1371/journal.pcbi.1000800)
51. Purtha WE, Tedder TF, Johnson S, Bhattacharya D, Diamond MS. 2011 Memory B cells, but not long-lived plasma cells, possess antigen specificities for viral escape mutants. *J. Exp. Med.* **208**, 2599–2606. (doi:10.1084/jem.20110740)
52. Glanville J *et al*. 2009 Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc. Natl Acad. Sci. USA* **106**, 20 216–20 221. (doi:10.1073/pnas.0909775106)
53. Mayer A, Balasubramanian V, Mora T, Walczak AM. 2015 How a well-adapted immune system is organized. *Proc. Natl Acad. Sci. USA* **112**, 5950–5955. (doi:10.1073/pnas.1421827112)
54. Saini J, Hershberg U. 2015 B cell Variable genes have evolved their codon usage to focus the targeted patterns of somatic mutation on the complementarity determining regions. *Mol. Immunol.* **65**, 157–167. (doi:10.1016/j.molimm.2015.01.001)
55. Wei L, Chahwan R, Wang S, Wang X, Pham PT, Goodman MF, Bergman A, Scharff MD, MacCarthy T. 2015 Overlapping hotspots in CDRs are critical sites for V region diversification. *Proc. Natl Acad. Sci. USA* **112**, E728–E737. (doi:10.1073/pnas.1500788112)
56. Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. 2003 Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374–1377. (doi:10.1126/science.1086907)
57. Wine Y *et al*. 2013 Molecular deconvolution of the monoclonal antibodies that comprise the polyclonal serum response. *Proc. Natl Acad. Sci. USA* **110**, 2993–2998. (doi:10.1073/pnas.1213737110)
58. Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. 2014 The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat. Biotechnol.* **32**, 158–168. (doi:10.1038/nbt.2782)
59. DeKosky BJ, Kojima T, Rodin A, Charab W, Ippolito GC, Ellington AD, Georgiou G. 2015 In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat. Med.* **21**, 86–91. (doi:10.1038/nm.3743)
60. Wrammert J *et al*. 2008 Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* **453**, 667–671. (doi:10.1038/nature06890)
61. Victora GD, Schwickert TA, Fooksman DR, Kamphorst AO, Meyer-Hermann M, Dustin ML, Nussenzweig MC. 2010 Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* **143**, 592–605. (doi:10.1016/j.cell.2010.10.032)
62. Hessell AJ *et al*. 2007 Fc receptor but not complement binding is important in antibody protection against HIV. *Nature* **449**, 101–104. (doi:10.1038/nature06106)
63. Bournazos S, Klein F, Pietzsch J, Seaman MS, Nussenzweig MC, Ravetch JV. 2014 Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for *in vivo* activity. *Cell* **158**, 1243–1253. (doi:10.1016/j.cell.2014.08.023)

64. Tonegawa S. 1983 Somatic generation of antibody diversity. *Nature* **302**, 575–581. (doi:10.1038/302575a0)
65. Arnaout R, Lee W, Cahill P, Honan T, Sparrow T, Weiand M, Nusbaum C, Rajewsky K, Koralov SB. 2011 High-resolution description of antibody heavy-chain repertoires in humans. *PLoS ONE* **6**, e22365. (doi:10.1371/journal.pone.0022365)
66. Thompson CB, Neiman PE. 1987 Somatic diversification of the chicken immunoglobulin light chain gene is limited to the rearranged variable gene segment. *Cell* **48**, 369–378. (doi:10.1016/0092-8674(87)90188-7)
67. Becker RS, Knight KL. 1990 Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits. *Cell* **63**, 987–997. (doi:10.1016/0092-8674(90)90502-6)
68. Nossal G, Austin CM, Ada GL. 1965 Antigens in immunity: VII. Analysis of immunological memory. *Immunology* **9**, 333.
69. Weigert MG, Cesari IM, Yonkovich SJ, Cohn M. 1970 Variability in the lambda light chain sequences of mouse antibody. *Nature* **228**, 1045–1047. (doi:10.1038/2281045a0)
70. McKean D, Huppi K, Bell M, Staudt L, Gerhard W, Weigert M. 1984 Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc. Natl Acad. Sci. USA* **81**, 3180–3184. (doi:10.1073/pnas.81.10.3180)
71. Kepler TB *et al*. 2014 Immunoglobulin gene insertions and deletions in the affinity maturation of HIV-1 broadly reactive neutralizing antibodies. *Cell Host Microbe* **16**, 304–313. (doi:10.1016/j.chom.2014.08.006)
72. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. 2000 Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563. (doi:10.1016/S0092-8674(00)00078-7)
73. Revy P *et al*. 2000 Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell* **102**, 565–575. (doi:10.1016/S0092-8674(00)00079-9)
74. Harris RS, Petersen-Mahrt SK, Neuberger MS. 2002 RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* **10**, 1247–1253. (doi:10.1016/S1097-2765(02)00742-6)
75. Rogozin IB, Pavlov YI, Bebenek K, Matsuda T, Kunkel TA. 2001 Somatic mutation hotspots correlate with DNA polymerase η error spectrum. *Nat. Immunol.* **2**, 530–536. (doi:10.1038/88732)
76. Zheng N-Y, Wilson K, Jared M, Wilson PC. 2005 Intricate targeting of immunoglobulin somatic hypermutation maximizes the efficiency of affinity maturation. *J. Exp. Med.* **201**, 1467–1478. (doi:10.1084/jem.20042483)
77. Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS. 1993 Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. *Proc. Natl Acad. Sci. USA* **90**, 2385–2388. (doi:10.1073/pnas.90.6.2385)
78. Hershberg U, Luning Prak ET. 2015 The analysis of clonal expansions in normal and autoimmune B cell repertoires. *Phil. Trans. R. Soc. B* **370**, 20140239. (doi:10.1098/rstb.2014.0239)
79. Wittenbrink N, Klein A, Weiser AA, Schuchhardt J, Or-Guil M. 2011 Is there a typical germinal center? A large-scale immunohistological study on the cellular composition of germinal centers during the hapten-carrier–driven primary immune response in mice. *J. Immunol.* **187**, 6185–6196. (doi:10.4049/jimmunol.1101440)
80. Clarke SH, Huppi K, Ruezinsky D, Staudt L, Gerhard W, Weigert M. 1985 Inter- and intraclonal diversity in the antibody response to influenza hemagglutinin. *J. Exp. Med.* **161**, 687–704. (doi:10.1084/jem.161.4.687)
81. Victora GD, Nussenzweig MC. 2012 Germinal centers. *Annu. Rev. Immunol.* **30**, 429–457. (doi:10.1146/annurev-immunol-020711-075032)
82. Zhang Y *et al*. 2013 Germinal center B cells govern their own fate via antibody feedback. *J. Exp. Med.* **210**, 457–464. (doi:10.1084/jem.20120150)
83. Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR. 2013 Genetic measurement of memory B-cell recall using antibody repertoire sequencing. *Proc. Natl Acad. Sci. USA* **110**, 13 463–13 468. (doi:10.1073/pnas.1312146110)
84. Luther SA, Gulbranson-Judge A, Acha-Orbea H, MacLennan IC. 1997 Viral superantigen drives extrafollicular and follicular B cell differentiation leading to virus-specific antibody production. *J. Exp. Med.* **185**, 551–562. (doi:10.1084/jem.185.3.551)
85. Boyd SD *et al*. 2010 Individual variation in the germline Ig gene repertoire inferred from variable region gene rearrangements. *J. Immunol.* **184**, 6986–6992. (doi:10.4049/jimmunol.1000445)
86. Collins AM, Wang Y, Roskin KM, Marquis CP, Jackson KJL. 2015 The mouse antibody heavy chain repertoire is germline-focused and highly variable between inbred strains. *Phil. Trans. R. Soc. B* **370**, 20140236. (doi:10.1098/rstb.2014.0236)
87. Nadel B, Tang A, Escuro G, Lugo G, Feeney AJ. 1998 Sequence of the spacer in the recombination signal sequence affects V(D)J rearrangement frequency and correlates with nonrandom Vκ usage *in vivo*. *J. Exp. Med.* **187**, 1495–1503. (doi:10.1084/jem.187.9.1495)
88. Feeney AJ, Tang A, Ogwaro KM. 2000 B-cell repertoire formation: role of the recombination signal sequence in non-random V segment utilization. *Immunol. Rev.* **175**, 59–69. (doi:10.1111/j.1600-065X.2000.imr017508.x)
89. Watson CT *et al*. 2013 Complete haplotype sequence of the human immunoglobulin heavy-chain variable, diversity, and joining genes and characterization of allelic and copy-number variation. *Am. J. Hum. Genet.* **92**, 530–546. (doi:10.1016/j.ajhg.2013.03.004)
90. Watson CT *et al*. 2014 Sequencing of the human IG light chain loci from a hydatidiform mole BAC library reveals locus-specific signatures of genetic diversity. *Genes Immun.* **16**, 24–34. (doi:10.1038/gene.2014.56)
91. Sasso EH, Johnson T, Kipps TJ. 1996 Expression of the immunoglobulin VH gene 51p1 is proportional to its germline gene copy number. *J. Clin. Invest.* **97**, 2074–2080. (doi:10.1172/JCI118644)
92. Volpe JM, Kepler TB. 2008 Large-scale analysis of human heavy chain V(D)J recombination patterns. *Immunome Res.* **4**, 3. (doi:10.1186/1745-7580-4-3)
93. Boyd SD *et al*. 2009 Measurement and clinical monitoring of human lymphocyte clonality by massively parallel V-D-J pyrosequencing. *Sci. Transl. Med.* **1**, 12ra23. (doi:10.1126/scitranslmed.3000540)
94. Kwong PD, Mascola JR. 2012 Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies. *Immunity* **37**, 412–425. (doi:10.1016/j.immuni.2012.08.012)
95. Liao H-X *et al*. 2013 Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469–476. (doi:10.1038/nature12053)
96. Jackson KJL *et al*. 2014 Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. *Cell Host Microbe* **16**, 105–114 (doi:10.1016/j.chom.2014.05.013)
97. Dunand CJH, Wilson PC. 2015 Restricted, canonical, stereotyped and convergent immunoglobulin responses. *Phil. Trans. R. Soc. B* **370**, 20140238. (doi:10.1098/rstb.2014.0238)
98. Martin V, Wu Y-C, Kipling D, Dunn-Walters D. 2015 Ageing of the B-cell repertoire. *Phil. Trans. R. Soc. B* **370**, 20140237. (doi:10.1098/rstb.2014.0237)
99. Alamyar E, Giudicelli V, Li S, Duroux P, Lefranc M-P. 2012 IMGT/HighV-QUEST: the IMGT® web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. *Immunome Res.* **8**, 26.
100. Ye J, Ma N, Madden TL, Ostell JM. 2013 IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res.* **41**, W34–W40. (doi:10.1093/nar/gkt382)
101. Gaëta BA, Malming HR, Jackson KJL, Bain ME, Wilson P, Collins AM. 2007 iHMMune-align: hidden Markov model-based alignment and identification of germline genes in rearranged immunoglobulin gene sequences. *Bioinformatics* **23**, 1580–1587. (doi:10.1093/bioinformatics/btm147)
102. Munshaw S, Kepler TB. 2010 SoDA2: a hidden Markov model approach for identification of immunoglobulin rearrangements. *Bioinformatics* **26**, 867–872. (doi:10.1093/bioinformatics/btq056)
103. Frost SDW, Murrell B, Hossain ASMM, Silverman GJ, Pond SLK. 2015 Assigning and visualizing germline genes in antibody repertoires. *Phil. Trans. R. Soc. B* **370**, 20140240. (doi:10.1098/rstb.2014.0240)
104. Elhanati Y, Sethna Z, Marcou Q, Callan Jr CG, Mora T, Walczak AM. 2015 Inferring processes underlying B-cell repertoire diversity. *Phil. Trans. R. Soc. B* **370**, 20140243. (doi:10.1098/rstb.2014.0243)
105. Kepler TB. 2013 Reconstructing a B-cell clonal lineage. I. Statistical inference of unobserved

ancestors. *F1000Res* **2**, 103. (doi:10.12688/f1000research.2-103.v1)

106. Gadala-Maria D, Yaari G, Uduman M, Kleinstein SH. 2015 Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. *Proc. Natl Acad. Sci. USA* **112**, E862–E870. (doi:10.1073/pnas.1417683112)
107. Jiang N, Weinstein JA, Penland L, White 3rd RA, Fisher DS, Quake SR. 2011 Determinism and stochasticity during maturation of the zebrafish antibody repertoire. *Proc. Natl Acad. Sci. USA* **108**, 5348–5353. (doi:10.1073/pnas.1014277108)
108. Bashford-Rogers RJM, Palser AL, Huntly BJ, Rance R, Vassiliou GS, Follows GA, Kellam P. 2013 Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations. *Genome Res.* **23**, 1874–1884. (doi:10.1101/gr.154815.113)
109. Doria-Rose NA *et al*. 2014 Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* **509**, 55–62. (doi:10.1038/nature13036)
110. McCoy CO, Bedford T, Minin VN, Bradley P, Robins H, Matsen IV FA. 2015 Quantifying evolutionary constraints on B-cell affinity maturation. *Phil. Trans. R. Soc. B* **370**, 20140244. (doi:10.1098/rstb.2014.0244)
111. Yaari G, Benichou JIC, Vander Heiden JA, Kleinstein SH, Louzoun Y. 2015 The mutation patterns in B-cell immunoglobulin receptors reflect the influence of selection acting at multiple time-scales. *Phil. Trans. R. Soc. B* **370**, 20140242. (doi:10.1098/rstb.2014.0242)
112. Łuksza M, Lässig M. 2014 A predictive fitness model for influenza. *Nature* **507**, 57–61. (doi:10.1038/nature13087)
113. Bedford T *et al*. 2014 Integrating influenza antigenic dynamics with molecular evolution. *Elife* **3**, e01914. (doi:10.7554/eLife.01914)
114. Vaughan TG, Kühnert D, Popinga A, Welch D, Drummond AJ. 2014 Efficient Bayesian inference under the structured coalescent. *Bioinformatics* **30**, 2272–2279. (doi:10.1093/bioinformatics/btu201)
115. Frank SA. 2002 *Immunology and evolution of infectious disease*. Princeton, NJ: Princeton University Press.
116. Wicker LS, Katz M, Sercarz EE, Miller A. 1984 Immunodominant protein epitopes I. Induction of suppression to hen egg white lysozyme is obliterated by removal of the first three N-terminal amino acids. *Eur. J. Immunol.* **14**, 442–447. (doi:10.1002/eji.1830140511)
117. Scheerlinck JP, DeLeys R, Saman E, Brys L, Geldhof A, De Baetselier P. 1993 Redistribution of a murine humoral immune response following removal of an immunodominant B cell epitope from a recombinant fusion protein. *Mol. Immunol.* **30**, 733–739. (doi:10.1016/0161-5890(93)90144-Z)
118. Vijayakrishnan L, Kumar V, Agrewala JN, Mishra GC, Rao KV. 1994 Antigen-specific early primary humoral responses modulate immunodominance of B cell epitopes. *J. Immunol.* **153**, 1613–1625.
119. Childs LM, Baskerville EB, Cobey S. 2015 Trade-offs in antibody repertoires to complex antigens. *Phil. Trans. R. Soc. B* **370**, 20140245. (doi:10.1098/rstb.2014.0245)
120. Pauli NT *et al*. 2014 *Staphylococcus aureus* infection induces protein A-mediated immune evasion in humans. *J. Exp. Med.* **211**, 2331–2339. (doi:10.1084/jem.20141404)
121. Krammer F, Palese P. 2015 Advances in the development of influenza virus vaccines. *Nat. Rev. Drug Discov.* **14**, 167–182. (doi:10.1038/nrd4529)
122. Ndifon W, Wingreen NS, Levin SA. 2009 Differential neutralization efficiency of hemagglutinin epitopes, antibody interference, and the design of influenza vaccines. *Proc. Natl Acad. Sci. USA* **106**, 8701–8706. (doi:10.1073/pnas.0903427106)
123. Francis Jr T. 1960 On the doctrine of original antigenic sin. *Proc. Am. Philos. Soc.* **104**, 572–578.
124. Kim JH, Skountzou I, Compans R, Jacob J. 2009 Original antigenic sin responses to influenza viruses. *J. Immunol.* **183**, 3294–3301. (doi:10.4049/jimmunol.0900398)
125. Schmidt AG *et al*. 2015 Viral receptor-binding site antibodies with diverse germline origins. *Cell* **161**, 1026–1034. (doi:10.1016/j.cell.2015.04.028)
126. Zarnitsyna VI, Ellebedy AH, Davis C, Jacob J, Ahmed R, Antia R. 2015 Masking of antigenic epitopes by antibodies shapes the humoral immune response to influenza. *Phil. Trans. R. Soc. B* **370**, 20140248. (doi:10.1098/rstb.2014.0248)
127. Luo S, Perelson AS. 2015 The challenges of modelling antibody repertoire dynamics in HIV infection. *Phil. Trans. R. Soc. B* **370**, 20140247. (doi:10.1098/rstb.2014.0247)
128. Hoehn KB *et al*. 2015 Dynamics of immunoglobulin sequence diversity in HIV-1 infected individuals. *Phil. Trans. R. Soc. B* **370**, 20140241. (doi:10.1098/rstb.2014.0241)
129. Lavinder JJ, Horton AP, Georgiou G, Ippolito GC. 2015 Next-generation sequencing and protein mass spectrometry for the comprehensive analysis of human cellular and serum antibody repertoires. *Curr. Opin. Chem. Biol.* **24**, 112–120. (doi:10.1016/j.cbpa.2014.11.007)
130. Zhu J *et al*. 2013 De novo identification of VRC01 class HIV-1–neutralizing antibodies by next-generation sequencing of B-cell transcripts. *Proc. Natl Acad. Sci. USA* **110**, E4088–E4097. (doi:10.1073/pnas.1306262110)
131. Gao F *et al*. 2014 Cooperation of B cell lineages in induction of HIV-1-broadly neutralizing antibodies. *Cell* **158**, 481–491. (doi:10.1016/j.cell.2014.06.022)
132. Bonsignori M *et al*. 2011 Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors. *J. Virol.* **85**, 9998–10 009. (doi:10.1128/JVI.05045-11)
133. Wang S *et al*. 2015 Manipulating the selection forces during affinity maturation to generate cross-reactive HIV antibodies. *Cell* **160**, 785–797. (doi:10.1016/j.cell.2015.01.027)
134. McGuire AT, Dreyer AM, Carbonetti S, Lippy A, Glenn J, Scheid JF, Mouquet H, Stamatatos L. 2014 Antigen modification regulates competition of broad and narrow neutralizing HIV antibodies. *Science* **346**, 1380–1383. (doi:10.1126/science.1259206)
135. Chaudhury S, Reifman J, Wallqvist A. 2014 Simulation of B cell affinity maturation explains enhanced antibody cross-reactivity induced by the polyvalent malaria vaccine AMA1. *J. Immunol.* **193**, 2073–2086. (doi:10.4049/jimmunol.1401054)
136. Yewdell JW, Frank E, Gerhard W. 1981 Expression of influenza A virus internal antigens on the surface of infected P815 cells. *J. Immunol.* **126**, 1814–1819.
137. Bodewes R, Geelhoed-Mieras MM, Wrammert J, Ahmed R, Wilson PC, Fouchier RAM, Osterhaus ADME, Rimmelzwaan GF. 2013 *In vitro* assessment of the immunological significance of a human monoclonal antibody directed to the influenza A virus nucleoprotein. *Clin. Vaccine Immunol.* **20**, 1333–1337. (doi:10.1128/CVI.00339-13)
138. McHeyzer-Williams LJ, Milpied PJ, Okitsu SL, McHeyzer-Williams MG. 2015 Class-switched memory B cells remodel BCRs within secondary germinal centers. *Nat. Immunol.* **16**, 296–305. (doi:10.1038/ni.3095)
139. Sela-Culang I *et al*. 2014 Using a combined computational-experimental approach to predict antibody-specific B cell epitopes. *Structure* **22**, 646–657. (doi:10.1016/j.str.2014.02.003)
140. Mann JK, Barton JP, Ferguson AL, Omarjee S, Walker BD, Chakraborty A, Ndung T. 2014 The fitness landscape of HIV-1 gag: advanced modeling approaches and validation of model predictions by *in vitro* testing. *PLoS Comput. Biol.* **10**, e1003776. (doi:10.1371/journal.pcbi.1003776)
141. Boyd SD, Jackson KJL. 2015 Predicting vaccine responsiveness. *Cell Host Microbe* **17**, 301–307. (doi:10.1016/j.chom.2015.02.015)
142. Cobey S. 2014 Pathogen evolution and the immunological niche. *Ann. NY Acad. Sci.* **1320**, 1–15. (doi:10.1111/nyas.12493)