Published in final edited form as:

Nat Rev Immunol. 2019 January ; 19(1): 55–62. doi:10.1038/s41577-018-0092-5.

Models of immune selection for multi-locus antigenic diversity of pathogens

Maria Georgieva^{1,*}, Caroline O. Buckee¹, and Marc Lipsitch^{1,2}

¹Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

²Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

Abstract

It is well accepted that pathogens can evade recognition and elimination by the host immune system by varying their antigenic targets. Thus, it has become a truism that host immunity is a major driver and determinant of the antigenic diversity of pathogens. However, it remains puzzling how host immunity selects for antigenic diversity at the level of the population, given that hosts have acquired immune responses to multiple antigens of most pathogens — sometimes through multiple effectors of both humoral and cellular immunity. In this Opinion article, we address this puzzle and the related question of why pathogens often have diversity at multiple antigenic loci. Here, we propose five hypotheses to explain the polymorphism of multiple antigens in a single species and highlight research relevant to our current models of thinking about multi-locus antigenic diversity.

Graphical Abstract

This Opinion article summarizes current models of thinking about multi-locus antigenic diversity. The authors propose five hypotheses to explain how host immunity selects for antigenic diversity at the level of the population.

Introduction

To counter innate and adaptive immune responses, microbial pathogens have developed multiple strategies to limit immune detection or to manipulate the host immune responses to their advantage. It is thought that pathogens that elicit an adaptive immune response evade recognition and elimination by the host immune system by varying their antigenic targets.

* maria.georgieva@unil.ch.

Current address: Department of Physiology, University of Lausanne, Lausanne, Switzerland.

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

M.L. has received consulting income/honoraria from Antigen Discovery, Pfizer, Merck and Affinivax, and has received institutional grant funding from Pfizer. The other authors declare no competing interests.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This creates selection for antigenic variation, because hosts who develop immunity to one variant of an antigen may still be susceptible to infection with pathogens that lack that antigen or express a different allelic variant of it^{1–14}. Consistent with this paradigm, antigenic determinants of viral, bacterial and protozoan pathogens are often highly variable^{15–17}. However, the population structure of this antigenic diversity exhibits different patterns in space and time¹⁸. For some chronic pathogens — including HIV, the human malaria parasite *Plasmodium falciparum* and the trypanosome parasites — diversity occurs within a host during the course of a single infection owing to active switching between antigenic variants or serial selection resulting from immune evasion. This within-host selection can result in extreme diversity in the circulating pathogen population (for a more in-depth discussion of this topic, see BOX 1). More commonly, antigenic diversity is observed at the level of the pathogen population; for these so-called multi-strain pathogen species, such as influenza virus and the bacterial pathogen *Streptococcus pneumoniae*, multiple clones or strains of pathogen co-circulate in the host population, each of which is defined by a particular suite of antigenic determinants.

It is widely accepted that multi-strain pathogens experience immune-mediated selection pressure to diversify. Despite there being other potential causes of antigenic diversity — such as variability stimulated by host shifts, which is particularly relevant for zoonotic viruses^{19,20} — it has become a truism that this host immune pressure selects for pathogen antigenic diversity. Nonetheless, some aspects of antigenic diversity are not fully explained. In particular, many models of multi-strain pathogen systems assume that the immune response is directed at a single antigenic target, such that diversification of that target offers a clear fitness advantage, which may require compensatory pathogen mutations to repair the fitness defects resulting from a loss of function of the pathogen target caused by the initial change^{21–25}. For most infections, however, host immunity involves multiple effectors targeting multiple antigens, and even multiple epitopes within a single antigen. Similarly, most pathogens have complex patterns of antigenic diversity at the population level, sharing some but not all antigenic targets between strains. These observations suggest two closely related puzzles, the first of which relates to mechanisms of immune escape by the pathogen and the second of which concerns the causes of the observed patterns of pathogen diversity at the population level. First, if specific immune responses directed against any one of multiple pathogen antigenic targets are sufficient to prevent infection, how does an individual pathogen gain an advantage by varying only one or a few of these targets, when this would leave it vulnerable to clearance by immune responses directed at the other targets? Second, why does a single pathogen species often have antigenic diversity at multiple loci?

These questions are striking by analogy to another selective pressure on pathogens, namely antimicrobial treatment. It is clear how pathogen mutants resistant to a single antimicrobial agent used for treatment can rise to high frequency in a treated host. However, multiple drugs are used to treat infections such as tuberculosis precisely because mutants resistant to any one of these drugs will still be killed by the other drugs^{26–30}. By analogy, one might expect that a multi-targeted host immune response would preclude any selective advantage for a pathogen variant that differs at a single locus, because responses directed at the other targets would be sufficient to eliminate the pathogen. In this scenario, host immunity

targeting multiple antigens and multiple epitopes on each antigen is analogous to combination therapy with a large number of drugs. Indeed, this comparison has been invoked to explain why resistance to vaccines is relatively uncommon, whereas resistance to drugs is common³¹. In the context of naturally acquired immune responses, however, the existence of high levels of sequence diversity in pathogens suggests that there is some benefit of immune escape at each of the individual loci to which hosts develop immune responses. What is the mechanism by which such diversity arises and is maintained in the context of multi-faceted immune responses?

We consider five hypotheses, each of which may partially help to resolve the question of how pathogens escape from multiple immune responses and maintain diversity at multiple antigenic loci. Here, we consider the case of acute infection, where reinfection depends on how distinct the new infecting strain is from the strain of previous exposure. Different models, not discussed here, may apply to chronic infections.

Hypothesis 1 posits that pathogens do not vary only one or a few important antigenic loci at a time, but rather that individual pathogen strains differ from one another at nearly all of the important antigenic loci. Hypothesis 2 questions the assumption that immune responses directed at any individual pathogen antigen are fully protective, proposing instead an additive model whereby each antigenic locus at which a pathogen can escape a preexisting immune response of the host will quantitatively improve the ability of the pathogen to infect that host. Hypothesis 3 suggests that varying one antigenic locus can render responses to other antigenic loci less effective, thereby creating escape from all immune responses with a single change to a pathogen antigen. Hypothesis 4 suggests that hosts are heterogeneous for various genetic and non-genetic reasons, so that an immune response targeting a particular antigen that is strongly protective in one host may be absent or less protective in another host. Hypothesis 5 emphasizes that antigens are not only targets of immunity but are also functional molecules that are often involved in host-pathogen interactions; it posits that the puzzling degree of sequence variation in some pathogen antigens may reflect selection for functional variability rather than antigenic escape.

These hypotheses are not mutually exclusive. For any particular pathogen species, several of these mechanisms may be operative. Although theoretical work has explored some of these ideas, most empirical studies of this issue have focused on a single hypothesis in a single pathogen. By bringing together the theoretical and empirical evidence for each hypothesis, we seek to stimulate further work to understand the phenomenon more deeply in individual pathogen species, as well as to inspire additional theoretical and comparative studies that may reveal the general mechanisms at work across host–pathogen systems.

Each hypothesis proposes that one of the common modelling assumptions about immune selection is incorrect. These hypotheses are not novel in themselves. Indeed, immunologists already know that many of these modelling assumptions do not fit some of the experimental data they observe. Rather, our goal in discussing these hypotheses as explanations for the polymorphism of multiple antigens in a single pathogen species is to suggest further directions to exploit the demonstrated synergy between pathogen population biology and immunology. In particular, immunological findings about the antigenicity of particular

pathogen moieties have stimulated advances in modelling the evolutionary epidemiology of those pathogens^{14, 32–36}, and population-level observations about patterns of variability in pathogen antigens have stimulated mechanistic studies to understand the causes and functional consequences of those variations^{37–39} and their implications for vaccination⁴⁰. With this Opinion article, we seek to highlight productive directions for future research along these lines but directed at understanding multiple antigens and their role in pathogen fitness, immunogenicity and immune clearance.

Hypotheses for multi-locus antigenic diversity

Hypothesis 1: non-overlapping antigenic repertoires.

We refer here to two studies that were among the first to highlight the puzzle of antigenic diversity at multiple pathogen loci and to propose a solution^{12, 41}. The authors noted the existence of antigenic diversity at multiple loci within two microbial populations — those of the malaria parasite *P. falciparum* and the bacterium *Neisseria meningitidis*. They proposed a conceptual model in which a host is completely immune to any pathogen strain that shares an antigenic variant with a strain that previously infected the same host. Under this model, if we consider three antigenic loci *A*, *B* and *C*, each having two alleles (Fig. 1a), a host who has previously been infected by a pathogen of genotype *A1B1C1* would be immune to infection with any pathogen carrying variant *A1* or *B1* or *C1* (Fig. 1b). In this setting, preexisting immunity at each antigenic locus would prevent a host from being infected by any pathogen genotype that shared any allele with a previous infection. Thus, if the pathogen cannot escape responses directed at all loci, what is the benefit of escaping at one of the antigenic loci? The resolution offered by the original authors^{12, 41} is that selective pressure from host immunity of this sort leads to pathogen populations within which several pathogen strains co-exist, each with completely non-overlapping antigenic repertoires. Thus, strain 1 may have variant 1 of each locus *A*, *B* and *C* (*A1B1C1*) and hosts immune to this strain can only be reinfected by strain *A2B2C2*, as they would be immune to strains with any shared antigenic variants, such as strain *A1B2C2* (Fig. 1b). At the population level, this leads to the emergence and coexistence of pathogen strains that are completely distinct in their antigenic epitopes and it selects against strains with partially overlapping antigens.

Some pathogens may indeed conform to this prediction. In particular, *N. meningitidis* has patterns of non-overlap between different antigenic loci that are consistent with the prediction^{3, 14}. However, it is challenging to evaluate the generality of this hypothesis because, for most pathogen species, the relative importance and identity of antigenic loci remain poorly defined. Also, the hypothesis predicts that recombination and mutation will constantly generate some pathogens that deviate from the non-overlapping pattern (although these would be eliminated by selection), which raises the practical question of what degree of deviation from purely non-overlapping antigenic repertoires would be required to falsify the hypothesis. For example, in *Streptococcus pneumoniae*, another bacterial pathogen with multiple antigens, there is evidence of association (linkage disequilibrium) between alleles at different antigenic loci, but this association is much weaker for some pairs of antigens than that observed in *N. meningitidis*⁴². A final challenge arises because in pathogens with low or moderate rates of recombination, some nonrandom associations between alleles at different

loci may exist because of clonal descent rather than because of selection acting to structure such variation⁴², although the non-overlapping pattern of antigens would be unlikely to be generated by chance among antigenic loci with many allelic variants. Stochastic mathematical models suggest that complex antigen patterns of overlapping structure are expected under moderate immune selection when there is competition for susceptible hosts or antigenic loci have variable levels of diversity^{35,43}, but more theoretical work is needed to quantify how much non-overlap of antigen repertoires is expected under different scenarios, and to definitively separate clonal descent from selection as an explanation.

Hypothesis 2: additive effects of host immunity.

The variation of multiple pathogen antigens to simultaneously escape from multiple immune responses is most puzzling if each immune response targeting a single antigen is individually lethal: that is, if a pathogen expressing one or more antigenic variants that have been seen before by a host is unable to infect that host. For several pathogen species, this seems not to be the case. Empirical evidence from pneumococcal colonization indicates that a single exposure to a particular strain does not provide sterilizing immunity to other strains that share one or more antigenic variants with the first strain, although it may reduce the rate of acquisition and/or increase the rate of clearance^{44–46}. Thus, part of the answer to the puzzle of multi-locus antigenic diversity may be that varying each antigen individually provides an incremental advantage to a pathogen strain encountering a host who has previously developed immunity to a strain that shares several alleles with the present strain. To illustrate, we consider three antigenic loci *A*, *B* and *C*, each of which has two variants (Fig. 2a). In this setting, diversification at each of these loci increases the antigenic distance between strains, such that *A1B1C1* and *A2B2C2* are the farthest apart (Fig. 2a). Under this hypothesis, increasing the antigenic distance incrementally increases the probability that a pathogen strain can infect a previously infected host (Fig. 2b). Intuitively, it seems that this could support antigenic variation at multiple sites, as variation at each site would contribute to pathogen fitness in a previously exposed host.

Hypothesis 3: variation at one locus affects responses to other loci.

As stated, the puzzling observations that motivate this article presume that variation at one antigenic locus decreases the effectiveness of immune responses directed at that locus, but not at other loci. However, if changes at one locus could reduce the effectiveness of immune responses directed at several loci, that single change would have a clear selective advantage even when responses are still directed at other loci. Under this model, we again consider three antigenic loci *A*, *B* and *C*. For simplicity, locus *A* has two variants (*A1* and *A2*), whereas loci *B* and *C* have a single variant (*B1* and *C1*, respectively) (Fig. 3a). A host who has previously experienced infection by a pathogen of genotype *A1B1C1* generates specific immunity against loci *A1*, *B1* and *C1*, thereby becoming immune to subsequent infection by a pathogen with the same genotype (Fig. 3b). Under hypothesis 3, variation of the antigenic locus *A* reduces the effectiveness of immune responses against additional loci by interfering with the access or effectiveness of antibodies against locus *B* and locus *C*. In these circumstances, a host who has been previously infected with a pathogen of genotype *A1B1C1* would be susceptible to a pathogen of genotype *A2B1C1*, despite preexisting immunity to loci *B1* and *C1* (Fig. 3b).

Studies of the haemagglutinin (HA) protein of influenza A virus provide an example of how this may occur⁴⁷. The HA protein, which is displayed on the surface of influenza A virus particles, binds sialic acid on the host cell surface and thereby promotes viral entry into the host cell⁴⁸. HA is the most important target of antibody-mediated immunity (both natural and vaccine-induced immune responses) and it is also the fastest-evolving antigen of influenza A virus. Mutational changes within the HA protein become fixed in the virus population between successive years of influenza A virus epidemics, which enables individuals to be repeatedly infected with the same virus subtype⁴⁹. It is typically posited that these amino acid changes in the HA protein enable repeated infection by allowing the virus to escape from neutralizing antibodies that can occlude the receptor-binding site and impede viral entry into the host cell⁵⁰. However, data suggest that virus escape from the polyclonal antibody response that is observed in human hosts would be unusual as the number of virus mutants with simultaneous point mutations at multiple epitopes of HA seems to be low⁵¹. By what mechanism, then, does sequence diversification in the HA protein of influenza A virus promote immune escape and enable repeated human infection?

A possible explanation is that antibody escape can occur without individually changing the antibody-binding affinity of each targeted pathogen epitope, by instead altering the kinetics of host receptor binding by the virus. The immunogenic regions of HA fall within the globular domain of the protein, a three-dimensional structure that is crucial for its biological function of binding to the sialic acid receptor on host cell surfaces⁵⁰. Mutations that increase the avidity of HA for the sialic acid receptor could facilitate rapid binding to host cells before antibodies can access the epitope regions of the molecule. Consequently, a single change to HA that promotes high receptor avidity would enable escape from the effects of polyclonal antibodies even in the absence of diversification in other epitopes⁴⁷. Indeed, the data are consistent with this idea as they show a positive correlation between receptor–HA binding avidity and escape from antibody responses⁴⁷. If this model is generally applicable, it could solve the paradox of escaping from multiple antibody responses because a single change to HA that increases receptor-binding avidity could lead to a simultaneous reduction in the effectiveness of antibodies directed at all epitopes of HA. Little is known about the evolutionary or physical constraints on the receptor avidity of HA, but these would undoubtedly have an important role in limiting the extent of immune escape via this mechanism⁵². One could imagine in other systems that a mutation that downregulates expression of a target could have a similar function of reducing the effectiveness of all immune responses directed at that target. We are unaware of examples or tests of this hypothesis related to natural immunity, although one report suggests that downregulated expression of meningococcal capsule may be a means of escape from vaccine-induced, polyclonal antibodies directed at the capsule⁵³.

Hypothesis 4: host heterogeneity.

Host heterogeneity in the pathogen antigens that are targeted by immune responses may provide another answer to the puzzle of multi-locus antigenic diversity. If individual hosts tend to respond primarily to one or a few pathogen antigens, but different hosts focus responses on different antigens, the variability of pathogens at multiple loci may reflect adaptation to their heterogeneous hosts, with variation at each locus being most important

for adaptation to the hosts who most strongly respond to the antigen encoded by that locus. An example of host heterogeneity that is relevant in this context could be diversity within the human leukocyte antigen (HLA) locus and differences in the host immune repertoire^{54–57}. This pattern of host heterogeneity in immune recognition has been found in human antibody responses during tuberculosis infection, where the identity of serologically reactive antigens differs greatly from individual to individual⁵⁸. It has been proposed that mechanisms regulating antigen availability and antigen presentation, as well as the process of immune exhaustion, underlie these between-host differences and shape the dynamics of individual host immune responses^{59–63} (for a more in-depth discussion of this topic, please see BOX 2). Another example of host heterogeneity is the existence of predominant germline mutations that are shared among some individuals and that specifically target functional antigens. For example, the IGHV1–69 germline mutation is associated with broadly neutralizing antibodies against the HA-stem region of the influenza A virus, and the IGHV1–2*02 germline mutation is associated with antibodies against the gp120 CD4-binding site of HIV-1^{64–66}.

Other, non-genetic factors, such as age, may add further diversity to the selective environments that are imposed by hosts on pathogens. For example, very young, old or immunocompromised hosts typically mount weak immune responses, whereas other age groups generate immunologically more robust and efficacious responses that clear the infection quickly^{67–69}. To the extent that age-related changes affect immune responses to different antigens in different hosts, this could be an additional source of host heterogeneity that provides pressure to vary different antigenic loci of the pathogen.

Hypothesis 5: selection for functional diversity.

A final candidate hypothesis to resolve these puzzles is that antigenic diversity of pathogens reflects selection for functional diversity rather than, or in addition to, immune escape. This alternative hypothesis addresses some examples of antigens that undergo sequence diversification without measurably escaping immunity⁷⁰. A good example of this is the pneumococcal surface protein C (PspC) of *S. pneumoniae* — a highly polymorphic protein that exists as multiple antigenic variants, each of which is marked by extensive sequence polymorphism and structural variation^{71,72}. Like other surface antigens, PspC is commonly assumed to be under diversifying selection to escape antibody-mediated immune responses. This conjecture seems to be supported by recent findings that individuals with strong antibody responses to a particular variant of PspC are more likely to carry a pneumococcal strain expressing a different PspC variant⁴⁴. Despite this, mechanistic studies do not always show measurable specificity of PspC-targeted antibody responses. For example, in a study of three recombinant PspC immunogens, PspC-binding antibodies were highly variant-specific in one case but substantially less so in others⁷⁰. Antibodies to PspC8 were shown to bind by Western blot and flow cytometry only to a bacterial strain containing that allele. By contrast, antibodies to PspC3 also bound to multiple other PspC alleles, and antibodies to PspC4 also showing binding to PspA, a different antigenic protein with strong homology to PspC⁷⁰. These findings of cross-reactive antibody responses against different PspC variants and the homologous PspA protein cast doubt on the classical immune selection hypothesis as being the sole explanation for antigenic diversity in *S. pneumoniae*.

If immune escape is not the only mechanism driving diversification of PspC, then an alternative mechanism is selection for functional diversity. This hypothesis is supported by data showing that variation across PspC alleles reflects between-variant differences in the ability to bind the host immunomodulatory factor H complement protein^{73,74,75}. Similarly, for influenza virus HA, selection for functional diversity in receptor binding activity may be involved because single mutations in HA have been shown to affect both antigenicity and receptor-binding avidity⁷⁶. If selection is maintaining functional diversity in pathogen molecules that are also antigenic, then what mechanisms underlie this selection? At present this is unknown, but some form of host heterogeneity — now referring to host heterogeneity in factors that would favor particular functional variants, rather than particular immune variants, of pathogen antigens — must underlie such selection.

Pathogens with low antigenic diversity

The hypotheses presented so far have all dealt with antigenically diverse pathogen species, and our questions have related to why such high levels of diversity should emerge and be maintained. However, given the many advantages of evading host immune responses that are discussed, an important part of the puzzle also relates to the other end of the spectrum in considering why some successful pathogens have almost no antigenic diversity. There are a few notable examples of antigenically stable pathogens, including viruses such as smallpox, mumps, rubella and measles, and several important bacterial pathogens. The basis for the absence of antigenic diversity in this group is incompletely understood. We discuss here two pathogens of public health importance in which antigenic conservation is common, namely *Mycobacterium tuberculosis* and the measles virus.

M. tuberculosis does not seem to use antigenic diversity as a mechanism of immune evasion. Two recent studies confirm that antigenic diversity is almost absent in *M. tuberculosis*, even among phylogenetically divergent strains^{77,78}. A possible explanation for this is that there is positive selection for epitope conservation, probably driven by a fitness benefit to epitope recognition by the host immune system in *M. tuberculosis*⁷⁹. One possibility is that this fitness benefit is mediated via increased transmission rates as an outcome of immune-driven lung tissue damage in the presence of activated T cells.

Similarly, the measles virus is antigenically stable and a single exposure to measles virus confers life-long immunity⁸⁰. However, *in vitro* experiments show that neutralizing monoclonal antibodies can select for escape variants in measles and suggest that, in laboratory settings at least, the virus has the capacity to diversify in the presence of immune pressure⁸¹. Why, then, does measles not exhibit antigenic diversity during natural infection? A possible explanation is that the polyclonal nature of host immunity, which simultaneously targets multiple antigens, diminishes the benefit of antigenic diversity in measles⁸⁰. Effectively, it seems that measles does indeed suffer from being targeted by multiple immune responses that are individually highly effective, which means that it cannot escape them all. It is perhaps the best example of why the highly variable nature of other pathogens is indeed a puzzle in terms of immune escape.

From an evolutionary perspective, an alternative explanation is that antigenic diversity may be less beneficial for highly transmissible viruses such as measles and, instead, a high reproductive number may be more important for these pathogen groups⁸². Therefore, it is plausible that in a tradeoff between antigenic diversity and high reproductive number, such viruses may favour specialization in terms of the latter. A theoretical study has suggested that certain features of the transmission network, as well as overall infectiousness, may promote such specialization in measles and other childhood infections⁹. An important limitation of that model, however, is that it does not explicitly consider reinfection of partially immune hosts, but focuses on the pattern of within-host diversification of the virus and its impact on transmission in a naive population as measured by the basic reproductive number.

Mechanisms of immune evasion

Pathogens use different strategies to evade or subvert the host immune response. Antigenic variation, latency, resistance to immune effector mechanisms and suppression of the immune response all contribute to immune evasion and, indeed, many pathogens use one or a combination of these strategies to subvert host immunity. From the point of view of the host, each of these mechanisms contributes to the same result — that the pathogen avoids normal host defence mechanisms and causes recurrent or persistent disease. For example, *S. pneumoniae* and *N. meningitidis* use strategies to circumvent complement-mediated immunity by limiting complement-mediated opsonization and subsequent phagocytosis by host immune cells^{83,84}, which probably allows these bacteria to persist during nasopharyngeal carriage and thus increases their fitness. Similarly, *P. falciparum* uses additional immune evasion strategies (other than antigenic variation) to subvert host immunity; these include dendritic cell modulation, apoptosis of parasite-specific T cells, and activation of T cell inhibitory pathways⁸⁵. It is worth pointing out that some of these other immune mechanisms may affect some of the observations made in relation to pathogen antigenic variation that we have discussed here. Thus, any of these processes may directly or indirectly promote genetic or phenotypic variability in pathogen antigens and favour different variants with improved functionality. In this context, antigenic diversity may be an ‘unintentional’ outcome of host-pathogen interaction rather than of immune pressure.

Concluding remarks

More work remains to be done on evolutionary models of antigenic variability to understand the set of biologically plausible assumptions under which multiple antigens can experience diversifying selection even in the presence of multiple immune responses. For example, many of the ‘classical’ models of immune selection make the assumption that a pathogen expressing one or more antigenic variants that have been seen before by a host is unable to infect that host, or has a much reduced probability of doing so. However, for many pathogens, as we discuss in relation to hypothesis 2, a host with immunity to several antigenic variants may still be readily infected or colonized by a strain expressing these variants because immunity is partial even to homologous variants. We have found, for example, that mice exposed to pneumococci can be recolonized by the identical strain despite having generated an immune response; the duration of that colonization episode

declines with repeated exposures but is not zero even after repeated colonization with the identical strain⁸⁶. In epidemiological data, immunity to repeated colonization with identical serotypes is only partial, perhaps reducing the probability of acquisition by ~30% for many serotypes^{34,45}. Similarly, weak epidemiological protection seems to occur for variant-specific responses to PspC⁴⁴. Future mathematical modelling should explore the evolutionary dynamics of multiple antigens under weak diversifying selection, whereby each variant that escapes existing immune responses modestly increases the fitness of the strain in an additive manner.

From a practical perspective, a better understanding of cross-reactivity in the context of variants of a single antigen can be valuable in advancing our efforts to find associations between the expression of specific antigenic variants and the generation of protective immune responses. In turn, this association could inform the future design and development of efficacious vaccines against highly variable pathogens. To uncover nuances of cross-reactivity, we argue for the use of truncated proteins as immunogens, to reflect the variable and unique antigenic portions of each antigen. This is a promising approach that has already been successfully applied in some contexts and permits the generation of variant-specific antibodies that can distinguish between antigenic variants in both binding and functional assays^{73,87}. Epitope mapping could be another valuable approach as highly parallel assays for serological reactivity to individual protein fragments and peptides become more accessible^{42,88}. In addition, more detailed mapping of antibody binding could help to disentangle epitope-specific immune responses from cross-reactivity against whole antigens. A better understanding of antibody function and the design of functional assays would further address these limitations.

Another challenge in understanding antibody cross-reactivity lies in finding the right tools to study this. Typically, *in vitro* techniques, such as Western blotting, are used to measure binding between antigenic variants and their antibodies to evaluate patterns of cross-reactivity. These assays are fairly crude, use denatured proteins and assess binding in isolation, away from the true physiological environment in which these interactions occur. Therefore, these *in vitro* techniques may be too blunt to detect biologically meaningful specificity in immune responses and results should be critically evaluated on a case-by-case basis.

A strong body of research suggests that host immunity has a role in the maintenance of antigenic diversity. However, some questions remain about the interplay between host immune pressure and sequence diversification in variable antigens. From the host side, it would be interesting to dissect the differences, if any, between the repertoires of antibody responses against different antigenic variants. From the pathogen side, it would be useful to know how much ‘distance’ in terms of sequence diversity between antigenic variants translates into immunologically distinct antigenic targets. Many questions remain, but going forward, both experimental work and theoretical models can be invaluable tools in deciphering these puzzles and advancing our knowledge of antigenic diversity and its biological principles.

Acknowledgements

The authors' work was supported by award R01AI048935 from the US National Institutes of Health.

References

1. Abu-Raddad LJ & Ferguson NM The impact of cross-immunity, mutation and stochastic extinction on pathogen diversity. *Proc. Biol. Sci* 271, 2431–2438 (2004). [PubMed: 15590592]
2. Abu-Raddad LJ & Ferguson NM Characterizing the symmetric equilibrium of multi strain host-pathogen systems in the presence of cross immunity. *J. Math. Biol* 50, 531–558 (2005). [PubMed: 15772826]
3. Callaghan MJ et al. The effect of immune selection on the structure of the meningococcal opa protein repertoire. *PLoS Pathog* 4, e1000020, doi:10.1371/journal.ppat.1000020 (2008). [PubMed: 18369470]
4. Ferguson N, Anderson R & Gupta S The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *PNAS* 96, 790–794 (1999). [PubMed: 9892712]
5. Gog JR & Swinton J A status-based approach to multiple strain dynamics. *Journal of Mathematical Biology* 44, 169–184 (2002). [PubMed: 11942531]
6. Gog JR & Grenfell BT Dynamics and selection of many-strain pathogens. *PNAS* 99, 17209–17214 (2002). [PubMed: 12481034]
7. Gomes MG, Medley GF & Nokes DJ On the determinants of population structure in antigenically diverse pathogens. *Proc. Biol. Sci* 269, 227–233 (2002).
8. Gupta S, Ferguson N & Anderson R Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280, 912–915 (1998). [PubMed: 9572737]
9. Lange A & Ferguson NM Antigenic diversity, transmission mechanisms, and the evolution of pathogens. *PLoS Comput. Biol* 5, e1000536, doi:10.1371/journal.pcbi.1000536 (2009). [PubMed: 19847288]
10. Luo S, Reed M, Mattingly JC & Koelle K The impact of host immune status on the within-host and population dynamics of antigenic immune escape. *J. R. Soc. Interface* 9, 2603–2613 (2012). [PubMed: 22572027]
11. Penman BS, Ashby B; Buckee CO & Gupta S Pathogen selection drives nonoverlapping associations between HLA loci. *PNAS* 110, 19645–19650 (2013). [PubMed: 24225852]
12. Recker M et al. Transient cross-reactive immune responses can orchestrate antigenic variation in malaria. *Nature* 429, 555–558 (2004). [PubMed: 15175751]
13. Recker M, Pybus OG, Nee S & Gupta S The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *PNAS* 104, 7711–7716 (2007). [PubMed: 17460037]
14. Watkins ER, Grad YH, Gupta S & Buckee CO Contrasting within- and between- host immune selection shapes *Neisseria* Opa repertoires. *Sci. Rep* 4, 6554, doi:10.1038/srep06554 (2014). [PubMed: 25296566]
15. Janeway CA, Travers P, Walport M & Shlomchik M in *Immunobiology* (ed. Gibbs S) (Garland Science, New York, 2001).
16. DeFranco AL, Locksley RM & Robertson M in *Immunity: The immune response in infectious and inflammatory disease* (Oxford University Press, 2007).
17. Gupta S & Maiden MCJ Exploring the evolution of diversity in pathogen populations. *Trends in Microbiology* 9, 181–185 (2001). [PubMed: 11286883]
18. Lipsitch M & O'Hagan JJ Patterns of antigenic diversity and the mechanisms that maintain them. *J. R. Soc. Interface* 4, 787–802 (2007). [PubMed: 17426010]
19. Mandl JN et al. Reservoir host immune responses to emerging zoonotic viruses. *Cell* 160, 20–35 (2015). [PubMed: 25533784]
20. Lipsitch M et al. Viral factors in influenza pandemic risk assessment. *Elife* 5, doi:10.7554/eLife.18491 (2016).

21. Das SR et al. Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *PNAS* 108, E1417–E1422 (2011). [PubMed: 22106257]
22. Friedrich TC et al. Reversion of CTL escape-variant immunodeficiency viruses in vivo. *Nat. Med* 10, 275–281 (2004). [PubMed: 14966520]
23. Troyer RM et al. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS Pathog* 5, e1000365, doi:10.1371/journal.ppat.1000365 (2009). [PubMed: 19343217]
24. Kosik I et al. Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs. *PLoS Pathog* 14, e1006796, doi:10.1371/journal.ppat.1006796 (2018). [PubMed: 29346435]
25. Lyons DM & Lauring AS Mutation and Epistasis in Influenza Virus Evolution. *Viruses* 10, doi: 10.3390/v10080407 (2018).
26. zur Wiesch PA, Kouyos R, Engelstädter J, Regoes RR & Bonhoeffer S Population biological principles of drug-resistance evolution in infectious diseases. *The Lancet Infectious Diseases* 11,236–247 (2011). [PubMed: 21371657]
27. Consortium REX Heterogeneity of selection and the evolution of resistance. *Trends Ecol. Evol* 28, 110–118 (2013). [PubMed: 23040463]
28. Gillespie SH Evolution of drug resistance in *Mycobacterium tuberculosis*: clinical and molecular perspective. *Antimicrobial Agents and Chemotherapy* 46, 267–274 (2002). [PubMed: 11796329]
29. Dye C Doomsday postponed? Preventing and reversing epidemics of drug-resistant tuberculosis. *Nature Reviews Microbiology* 7, 81–87 (2009). [PubMed: 19079354]
30. Ehrlich P Address in pathology on chemotherapeutics: scientific principles, methods, and results. *The Lancet* 182, 445–451 (1913).
31. Kennedy DA & Read AF Why does drug resistance readily evolve but vaccine resistance does not? *Proc. Biol. Sci* 284, doi:10.1098/rspb.2016.2562 (2017).
32. Koelle K, Cobey S, Grenfell B & Pascual M Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science* 314, 1898–1903 (2006). [PubMed: 17185596]
33. Zinder D, Bedford T, Gupta S & Pascual M The roles of competition and mutation in shaping antigenic and genetic diversity in influenza. *PLoS Pathog* 9, e1003104, doi:10.1371/journal.ppat.1003104 (2013). [PubMed: 23300455]
34. Cobey S & Lipsitch M Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science* 335, 1376–1380 (2012). [PubMed: 22383809]
35. Buckee CO et al. Role of selection in the emergence of lineages and the evolution of virulence in *Neisseria meningitidis*. *PNAS* 105, 15082–15087 (2008). [PubMed: 18815379]
36. Buckee CO, Gupta S, Kriz P, Maiden MC & Jolley KA Long-term evolution of antigen repertoires among carried meningococci. *Proc. Biol. Sci* 277, 1635–1641 (2010). [PubMed: 20129981]
37. Hausdorff WP, Feikin DR & Klugman KP Epidemiological differences among pneumococcal serotypes. *The Lancet Infectious Diseases* 5, 83–93 (2005). [PubMed: 15680778]
38. Li Y et al. Distinct effects on diversifying selection by two mechanisms of immunity against *Streptococcus pneumoniae*. *PLoS Pathog* 8, e1002989, doi:10.1371/journal.ppat.1002989 (2012). [PubMed: 23144610]
39. Ndung'u T & Weiss RA On HIV diversity. *AIDS* 26,1255–1260 (2012). [PubMed: 22706010]
40. Korber B, Hraber P, Wagh K & Hahn BH Polyvalent vaccine approaches to combat HIV-1 diversity. *Immunol. Rev* 275, 230–244 (2017). [PubMed: 28133800]
41. Gupta S et al. The maintenance of strain structure in populations of recombining infectious agents. *Nature Medicine* 2, 437–442 (1996).
42. Croucher NJ et al. Diverse evolutionary patterns of pneumococcal antigens identified by pangenome-wide immunological screening. *PNAS* 114, E357–E366 (2017). [PubMed: 28053228]
43. Buckee C, Recker M, Watkins ER & Gupta S Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations. *PNAS* 108, 15504–15509 (2011). [PubMed: 21876129]

44. Azarian T et al. Association of pneumococcal protein antigen serology with age and antigenic profile of colonizing Isolates. *J. Infect. Dis* 215, 713–722 (2017). [PubMed: 28035010]
45. Weinberger DM et al. Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. *J. Infect. Dis* 197, 1511–1518 (2008). [PubMed: 18471062]
46. Turner P et al. A longitudinal study of *Streptococcus pneumoniae* carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. *PLoS One* 7, e38271, doi:10.1371/journal.pone.0038271 (2012). [PubMed: 22693610]
47. Hensley SE et al. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science* 326, 734–736 (2009). [PubMed: 19900932]
48. Gamblin SJ & Skehel JJ Influenza hemagglutinin and neuraminidase membrane glycoproteins. *J. Biol. Chem* 285, 28403–28409 (2010). [PubMed: 20538598]
49. Plotkin JB, Dushoff J & Levin SA Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *PNAS* 99, 6263–6268 (2002). [PubMed: 11972025]
50. Wilson IA & Cox NJ. Structural basis of immune recognition of influenza virus hemagglutinin. *Annu. Rev. Immunol* 8, 737–771 (1990). [PubMed: 2188678]
51. Yewdell JW, Webster RG & Gerhard WU Antigenic variation in three distinct determinants of an influenza type A haemagglutinin molecule. *Nature* 279, 246–248 (1979). [PubMed: 86955]
52. Benton DJ, Martin SR, Wharton SA & McCauley JW Biophysical measurement of the balance of influenza a hemagglutinin and neuraminidase activities. *J. Biol. Chem* 290, 6516–6521 (2015). [PubMed: 25586179]
53. Maiden MC et al. Impact of meningococcal serogroup C conjugate vaccines on carriage and herd immunity. *J. Infect. Dis* 197, 737–743 (2008). [PubMed: 18271745]
54. Su LF, Kidd BA, Han A, Kotzin JJ & Davis MM Virus-specific CD4(+) memory- phenotype T cells are abundant in unexposed adults. *Immunity* 38, 373–383 (2013). [PubMed: 23395677]
55. Correia-Neves M, Waltzinger C, Mathis D & Benoist C The shaping of the T cell repertoire. *Immunity* 14, 21–32 (2001). [PubMed: 11163227]
56. von Boehmer H Shaping the T cell repertoire. *The Journal of Immunology* 175, 7067–7068 (2005). [PubMed: 16301607]
57. Bevan MJ The major histocompatibility complex determines susceptibility to cytotoxic T cells directed against minor histocompatibility antigens. *Journal of Experimental Medicine* 142, 1349–1364 (1975). [PubMed: 53263]
58. Lyashchenko K et al. Heterogeneous antibody responses in tuberculosis. *Infect. Immun* 66, 3936–3940 (1998). [PubMed: 9673283]
59. Weidt G, Utermohlen O, Heukeshoven J, Lehmann-Grube F & Deppert W Relationship among immunodominance of single CD8 T cell epitopes, virus load, and kinetics of primary antiviral CTL response. *The Journal of Immunology* 160, 2923–2931 (1998). [PubMed: 9510196]
60. Yewdell JW & Bennink JR Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu. Rev. Immunol* 17, 51–88 (1999). [PubMed: 10358753]
61. Burrows SR, Khanna R, Burrows JM & Moss DJ An alloresponse in humans is dominated by cytotoxic T lymphocytes (CTL) cross-reactive with a single Epstein-Barr virus CTL epitope: implications for graft-versus-host disease. *J. Exp. Med* 179, 1155–1161 (1994). [PubMed: 7511682]
62. Burrows SR et al. T cell receptor repertoire for a viral epitope in humans is diversified by tolerance to a background major histocompatibility complex antigen. *J. Exp. Med* 182, 1703–1715 (1995). [PubMed: 7500015]
63. Moir S et al. Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. *J. Exp. Med* 205, 1797–1805 (2008). [PubMed: 18625747]
64. Corti D et al. Heterosubtypic neutralizing antibodies are produced by individuals immunized with a seasonal influenza vaccine. *J. Clin. Invest* 120, 1663–1673 (2010). [PubMed: 20389023]
65. Avnir Y et al. Molecular signatures of hemagglutinin stem-directed heterosubtypic human neutralizing antibodies against influenza A viruses. *PLoS Pathog* 10, e1004103, doi:10.1371/journal.ppat.1004103 (2014). [PubMed: 24788925]

66. Wu X et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* 333, 1593–1602 (2011). [PubMed: 21835983]
67. Weng NP Aging of the immune system: how much can the adaptive immune system adapt? *Immunity* 24, 495–499 (2006). [PubMed: 16713964]
68. Goronzy JJ & Weyand CM Immune aging and autoimmunity. *Cell Mol. Life Sci* 69, 1615–1623 (2012). [PubMed: 22466672]
69. Shaw AC, Joshi S, Greenwood H, Panda A & Lord JM Aging of the innate immune system. *Curr. Opin. Immunol* 22, 507–513 (2010). [PubMed: 20667703]
70. Moreno AT et al. Cross-reactivity of antipneumococcal surface protein C (PspC) antibodies with different strains and evaluation of inhibition of human complement factor H and secretory IgA binding via PspC. *Clin. Vaccine Immunol* 19, 499–507 (2012). [PubMed: 22336290]
71. Iannelli F, Oggioni MR & Pozzi G Allelic variation in the highly polymorphic locus *pspC* of *Streptococcus pneumoniae*. *Gene* 284, 63–71 (2002). [PubMed: 11891047]
72. Brooks-Walter A, Briles DE & Hollingstead SK The *pspC* gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia. *Infect. Immun* 67, 6533–6542 (1999). [PubMed: 10569772]
73. Georgieva M, Kagedan L, Lu YJ, Thompson CM & Lipsitch M Antigenic variation in *Streptococcus pneumoniae* PspC promotes immune escape in the presence of variant-specific immunity. *MBio* 9, doi:10.1128/mBio.00264-18 (2018).
74. Hyams C et al. *Streptococcus pneumoniae* capsular serotype invasiveness correlates with the degree of factor H binding and opsonization with C3b/iC3b. *Infect. Immun* 81, 354–363 (2013). [PubMed: 23147038]
75. Yuste J et al. The effects of PspC on complement-mediated immunity to *Streptococcus pneumoniae* vary with strain background and capsular serotype. *Infect. Immun* 78, 283–292 (2010). [PubMed: 19884335]
76. Li Y et al. Single hemagglutinin mutations that alter both antigenicity and receptor binding avidity influence influenza virus antigenic clustering. *J. Virol* 87, 9904–9910 (2013). [PubMed: 23824816]
77. Comas I et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat. Genet* 42, 498–503 (2010). [PubMed: 20495566]
78. Coscolla M et al. *M. tuberculosis* T Cell Epitope Analysis Reveals Paucity of Antigenic Variation and Identifies Rare Variable TB Antigens. *Cell Host Microbe* 18, 538–548 (2015). [PubMed: 26607161]
79. Reichler MR et al. Evaluation of investigations conducted to detect and prevent transmission of tuberculosis. *JAMA* 287, 991–995 (2002). [PubMed: 11866646]
80. Griffin DE The immune response in measles: virus control, clearance and protective immunity. *Viruses* 8, doi:10.3390/v8100282 (2016).
81. Birrer MJ, Udem S, Nathenson S & Bloom BR Antigenic variants of measles virus. *Nature* 293, 67–69 (1981). [PubMed: 6167866]
82. Frank SA & Bush RM Barriers to antigenic escape by pathogens: trade-off between reproductive rate and antigenic mutability. *BMC Evol. Biol* 7, 229, doi:10.1186/1471-2148-7-229 (2007). [PubMed: 18005440]
83. Kadioglu A, Weiser JN, Paton JC & Andrew PW The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat. Rev. Microbiol* 6, 288–301 (2008). [PubMed: 18340341]
84. Serruto D, Rappuoli R, Scarselli M, Gros P & van Strijp JA Molecular mechanisms of complement evasion: learning from staphylococci and meningococci. *Nat. Rev. Microbiol* 8, 393–399 (2010). [PubMed: 20467445]
85. Stanisic DI, Barry AE & Good MF Escaping the immune system: how the malaria parasite makes vaccine development a challenge. *Trends Parasitol* 29, 612–622 (2013). [PubMed: 24176554]
86. Malley R et al. CD4 T cells mediate antibody-independent acquired immunity to pneumococcal colonization. *PNAS* 102, 4848–4853 (2005). [PubMed: 15781870]

87. Yassine HM et al. Use of hemagglutinin stem probes demonstrate prevalence of broadly reactive group 1 influenza antibodies in human sera. *Sci. Rep* 8, 8628, doi:10.1038/s41598-018-26538-7 (2018). [PubMed: 29872070]
88. Xu GJ et al. Viral immunology. Comprehensive serological profiling of human populations using a synthetic human virome. *Science* 348, aaa0698, doi:10.1126/science.aaa0698 (2015). [PubMed: 26045439]
89. Wei X et al. Antibody neutralization and escape by HIV-1. *Nature* 422, 307–312 (2003). [PubMed: 12646921]
90. Deng K et al. Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature* 517, 381–385 (2015). [PubMed: 25561180]
91. Asquith B, Edwards CT, Lipsitch M & McLean AR Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells in vivo. *PLoS Biol* 4, e90, doi:10.1371/journal.pbio.0040090 (2006). [PubMed: 16515366]
92. Cnops J, Magez S & De Trez C Escape mechanisms of African trypanosomes: why trypanosomiasis is keeping us awake. *Parasitology* 142, 417–427 (2015). [PubMed: 25479093]
93. Peck A & Mellins ED Precarious balance: Th17 cells in host defense. *Infect. Immun* 78, 32–38 (2010). [PubMed: 19901061]
94. Trzcinski K et al. Protection against nasopharyngeal colonization by *Streptococcus pneumoniae* is mediated by antigen-specific CD4+ T cells. *Infect. Immun* 76, 2678–2684 (2008). [PubMed: 18391006]
95. Li Y, Thompson CM, Trzcinski K & Lipsitch M. Within-host selection is limited by an effective population of *Streptococcus pneumoniae* during nasopharyngeal colonization. *Infect. Immun* 81, 4534–4543 (2013). [PubMed: 24082074]
96. Bonhoeffer S, Nowak MA Intra-host versus inter-host selection: Viral strategies of immune function impairment. *PNAS* 91, 8062–8066 (1994). [PubMed: 8058757]
97. Frank SA in *Immunology and evolution of infectious disease* (Princeton University Press, 2002)
98. Sette A & Sundaran R in *Immunodominance - the choice of the immune system* (ed. Frelinger JA) 55–71 (Weinheim: Wiley-VCH Verlag GmbH & Co, 2006).
99. Sadegh-Nasseri S & Kim A Selection of immunodominant epitopes during antigen processing is hierarchical. *Mol. Immunol* doi:10.1016/j.molimm.2018.08.011 (2018).
100. Angeletti D & Yewdell JW Understanding and manipulating viral immunity: antibody immunodominance enters center stage. *Trends Immunol* 39, 549–561 (2018). [PubMed: 29789196]
101. Giuliani M et al. Human protective response induced by meningococcus B vaccine is mediated by the synergy of multiple bactericidal epitopes. *Sci. Rep* 8, 3700, doi:10.1038/s41598-018-22057-7 (2018). [PubMed: 29487324]

BOX 1 |**Determinants of within-host selection for immune escape**

The main text discusses selection for variability of pathogen antigens at the level of between-host transmission. That is, selection exerted by a host with immunity to pathogen strains they have been exposed to subsequently determines which strains are still able to infect that host. Interesting observations also arise in considering within-host selection for antigenic variation: that is, the fate of a pathogen mutant arising during infection of a single host (or a small pathogen subpopulation that may initiate a secondary infection of an already infected host). The advantage of a pathogen mutant that lacks or varies a particular antigenic epitope in the context of an infection that has generated an immune response to that epitope depends on the details of the immune mechanism stimulated against that epitope. A pathogen mutant that escapes recognition by a specific cytotoxic T lymphocyte (CTL) lineage or a specific antibody will typically have a selective advantage and thus increase in frequency within the host. For example, this has been shown to occur during the course of infection with HIV-1^{89–91} and trypanosomes⁹². Selection efficiently results in an increase in the frequency of the escape mutants because the target of the antibody (a free parasite or virus) must display the antigen for the antibody to bind; similarly, the target of the CTL (a virus-infected host cell) must display the antigen on its surface (as an MHC class I-restricted peptide) for the CTL to kill the cell. By contrast, other immune effectors, notably T helper (T_H) cells, function mainly by secreting cytokines, which may have effects beyond the immediate area of the antigen-presenting cell that stimulates them. For example, when stimulated by cognate antigen, T_H17 cells release cytokines that recruit non-antigen-specific innate effector cells such as monocytes, macrophages and neutrophils⁹³. If a subpopulation of mutant pathogens lacking an important epitope were to arise during an infection caused by a pathogen population of which a majority of pathogens still has that epitope, the antigen-specific stimulation of T_H17 cells by the majority population would persist and the innate effector cells would continue to be recruited. As the innate effectors are not antigen-specific, they would not discriminate between pathogens with or without the epitope. For example, colonization of mice with a mixture of *Streptococcus pneumoniae* that do or do not contain a specific antigen to which the mice have been previously exposed leads to clearance of both subpopulations⁹⁴ at nearly equal rates³⁸. This should lead to less evidence of diversifying selection in the sequences of antigenic epitopes targeted by T_H17 cells than in the sequences of antigenic epitopes targeted by antibodies and B cells, a prediction that has been confirmed in *S. pneumoniae*³⁸. Evolutionary responses to selection within a single infection may also be limited by the small effective population size in some infections, as small selective advantages may not lead to escape if, for example, stochastic factors overwhelm the selective advantage of immune escape⁹⁵. The differences between selection at the within-host and between-host levels have been characterized more generally for several mechanisms of immune escape⁹⁶.

BOX 2 |**Immunodominance**

Pathogens present many antigens and epitopes to the host immune system but in the case of common acute viral and bacterial infections, a strong immune response develops against only a few of these targets^{97,98}. This phenomenon is termed immunodominance. It is best understood for T cells; only a few peptides that can bind the peptide-binding groove of MHC class II molecules are selected and presented to T cells, and this defines which pathogen epitopes stimulate an immune response. The mechanisms that determine which epitopes form the peptide-MHC class II complex, and the factors that lead to the availability of these epitopes and their recognition by T cells, are complex and involve multiple levels of selection and processing of particular pathogen epitopes⁹⁹. Epitope accessibility, the selective processing of certain peptides by proteases and the stability of peptide-MHC class II binding are all likely to be important factors in determining which T cell epitopes are responded to by the host. Antibody immunodominance is less well characterized. It remains unclear why particular epitopes of an antigen are immunodominant for B cells, but the distribution of avidities and frequency of naive precursor B cells are likely to have a role in determining which B cell lineages are most competitive within the germinal centres during affinity maturation¹⁰⁰. Immunodominance is thought to be adaptive for the host because it maximizes the power of functional effector responses while retaining a manageable number of memory cell specificities, and minimizing the risk of autoimmunity and excessive immune activation that can result in immune-mediated pathology. However, misdirected immunodominance may lead to an insufficiently broad immune response, which hinders clearance of the infection, or it can skew the immune response and subsequent immune memory towards particular epitopes, such that the host is primed to make suboptimal responses to pathogens expressing similar, but not identical, antigens. In broad terms it explains why, for many pathogens, only a subset of immunogenic epitopes are under strong diversifying selection; in this Opinion article, we primarily describe the population-level consequences of immunity with respect to immunodominant antigenic epitopes.

Subject categories

Biological sciences / Immunology / Immunogenetics

[URI/631/250/248]

Biological sciences / Evolution

[URI/631/181]

Biological sciences / Microbiology / Pathogens

[URI/631/326/421]

Biological sciences / Immunology / Antimicrobial responses

[URI/631/250/2499]

Biological sciences / Immunology / Adaptive immunity

[URI/631/250/2152]

Figure 1 | Non-overlapping antigenic repertoires.

a | We consider three antigenic loci *A*, *B* and *C*, each of which has two variants, such that *A1*, *A2*, *B1*, *C1* and *C2* together comprise the antigenic space for the example pathogen. **b |** A host who has previously been infected by a pathogen of genotype *A1B1C1* generates immune responses against *A1*, *B1* and *C1* antigens and is therefore immune to infection with any pathogen with a genotype carrying *A1* or *B1* or *C1*. Therefore, a host who has previously been infected by a pathogen of genotype *A1B1C1* would be subsequently susceptible to only a pathogen expressing a different variant at each antigenic locus — namely, *A2B2C2*.

Figure 2 | Additive effects of host immunity.

a | We consider three antigenic loci *A*, *B* and *C*, each of which has two variants, such that *A1*, *A2*, *B1*, *B2*, *C1* and *C2* together comprise the antigenic space for the example pathogen. Sequence diversification of these loci increases the antigenic distance between strains, such that a pathogen of genotype *A2B2C2*, which is marked by antigenic diversity in all three loci, has the greatest antigenic distance from a pathogen of genotype *A1B1C1*. **b** | Increases in antigenic distance between strains incrementally increase the probability that a pathogen strain can infect a previously infected host. Although the relationship shown here between antigenic distance and probability of infection is linear, it could also be nonlinear as a result of a synergistic relationship between immune effectors such as antibodies, for example¹⁰¹.

Figure 3 |. Variation at one locus affects responses at other loci.

a | We consider three antigenic loci *A*, *B* and *C*. Locus *A* has two variants (*A1* and *A2*), whereas loci *B* and *C* have a single variant each (*B1* and *C1*, respectively). Therefore, *A1*, *A2*, *B1* and *C1* together comprise the antigenic space for the example pathogen. **b** | A host who has previously been infected by a pathogen of genotype *A1B1C1* generates specific immunity against loci *A1*, *B1* and *C1*, and is therefore immune to subsequent infection by a pathogen with the same genotype. Under the model of hypothesis 2, the single variable antigenic locus *A* with two variants (*A1* and *A2*) could reduce the effectiveness of immunity against additional loci (locus *B* and locus *C*), thereby rendering the *A1B1C1*-immune host susceptible to any pathogen with a genotype containing *A2*, despite pre-existing immunity to loci *B* and *C*.