

# Polyvalent vaccine approaches to combat HIV-1 diversity

Bette Korber<sup>1,2</sup> | Peter Hrabec<sup>1</sup> | Kshitij Wagh<sup>1</sup> | Beatrice H. Hahn<sup>3,4</sup>

<sup>1</sup>Theoretical Biology and Biophysics, T6, Los Alamos National Laboratory, Los Alamos, NM, USA

<sup>2</sup>New Mexico Consortium, Los Alamos, NM, USA

<sup>3</sup>Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA

<sup>4</sup>Department of Microbiology, University of Pennsylvania, Philadelphia, PA, USA

## Correspondence

Bette Korber, T6, Theoretical Biology and Biophysics, MS-K710, Los Alamos National Laboratory, Los Alamos, NM, USA.  
Email: btk@lanl.gov

## Funding information

NIH, NIAID, Division of AIDS, Duke Center for HIV/AIDS Vaccine Immunology-Immunogen Discovery (CHAVI-ID), Grant/Award Number: UM-1-AI100645.

## Summary

A key unresolved challenge for developing an effective HIV-1 vaccine is the discovery of strategies to elicit immune responses that are able to cross-protect against a significant fraction of the diverse viruses that are circulating worldwide. Here, we summarize some of the immunological implications of HIV-1 diversity, and outline the rationale behind several polyvalent vaccine design strategies that are currently under evaluation. Vaccine-elicited T-cell responses, which contribute to the control of HIV-1 in natural infections, are currently being considered in both prevention and treatment settings. Approaches now in preclinical and human trials include full proteins in novel vectors, concatenated conserved protein regions, and polyvalent strategies that improve coverage of epitope diversity and enhance the cross-reactivity of responses. While many barriers to vaccine induction of broadly neutralizing antibody (bNAb) responses remain, epitope diversification has emerged as both a challenge and an opportunity. Recent longitudinal studies have traced the emergence of bNAbs in HIV-1 infection, inspiring novel approaches to recapitulate and accelerate the events that give rise to potent bNAb in vivo. In this review, we have selected two such lineage-based design strategies to illustrate how such in-depth analysis can offer conceptual improvements that may bring us closer to an effective vaccine.

## KEYWORDS

AIDS, antibodies, antigens/peptides/epitopes, B cells, vaccination, viral

## 1 | INTRODUCTION

Human Immunodeficiency Virus Type 1 (HIV-1) establishes chronic infections that are never cleared, but persist for life. Combination antiretroviral therapy (ART) suppresses viremia and halts disease progression,<sup>1</sup> and also reduces virus transmission.<sup>2</sup> But without continuous, lifelong treatment, HIV-1 eventually leads to the development of immunodeficiency (AIDS) and death in most infected individuals. HIV/AIDS has remained a public health threat for many decades now, and concerted global efforts have resulted in successful prevention

and treatment strategies that are beginning to show success in curtailing the spread and pathogenicity of this disease. Importantly, death rates and new HIV-1 infections have declined in many regions of the world.<sup>3</sup> Despite these advances, HIV-1 remains a devastating global health problem. Over half of the people living with HIV-1 are not receiving treatment, and many infected people are not aware of their infection. In 2015, there were an estimated 36.7 million people living with HIV-1, and approximately 1 million of these died AIDS-related deaths.<sup>3</sup> Moreover, by compromising the immune system, HIV-1 increases the morbidity and mortality of other devastating diseases, including tuberculosis<sup>4</sup> and malaria,<sup>5</sup> which are endemic to regions with a high HIV-1 prevalence. Thus, an effective vaccine that protects against HIV-1 infection is urgently needed.

This article is part of a series of reviews covering B cells and Immunity to HIV appearing in Volume 275 of *Immunological Reviews*.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Published 2017. This article is a U.S. Government work and is in the public domain in the USA. *Immunological Reviews* published by John Wiley & Sons Ltd

Within each of the more than 70 million people who have been infected with HIV-1 since the beginning of the pandemic,<sup>3</sup> an extraordinary evolutionary process has played out. High baseline rates of HIV-1 mutation<sup>6</sup> and replication<sup>7</sup> lay a foundation of viral diversity, and this combined with continuous immune-driven selection generates a complex and unique viral quasispecies in each infected person.<sup>8–12</sup> An extensive range of mutations and combinations of mutations, even within single epitopes, can be explored by the virus in response to immune pressures,<sup>8,9,13</sup> and the delicate balance of immune escape and replicative fitness shifts over time as other mutations, sometimes compensatory, come into play.<sup>14</sup> Successful ART constrains this dynamic, although the extent to which the virus continues to replicate during ART is still not fully resolved.<sup>15</sup> Unless ART is given immediately after infection, HIV-1 will have established a highly complex population of variants by the time therapy is initiated, including a latent reservoir that persists despite therapy and reactivates when ART is discontinued or when drug-resistance develops.<sup>16</sup>

Thus, immunologically relevant mutations accrue in every infected individual, creating a unique and highly complex mixture of variants within each host. When an infected person passes the virus to another individual, one or more of these variants will establish a new infection in the next individual in the chain, ultimately giving rise to a highly diverse global virus population. A successful vaccine must counter this diversity.<sup>17,18</sup> The term “wildtype” virus, which traditionally refers to the dominant form in a viral population, makes little sense in the context of HIV-1, yet the term is often used. Population-based viral consensus sequences are sometimes referred to as wildtype, but even within a genetic clade, they are quite distant from any natural isolate.<sup>17</sup> Each of the 35 million HIV-1 infections in the world today is comprised of numerous coexisting variants. Thus, we are confronted with 35 million related, but distinct “wildtypes”, each of which embodies a complex viral quasispecies. Given the very high cost of vaccine development, it is generally considered prohibitive to include more than two or possibly three, antigens in a polyvalent vaccine. Two or three against 35 million is a daunting ratio, and it is no wonder that an effective vaccine still eludes us! As a result, much current effort is focused on the design and selection of sets of vaccine antigens that may ultimately be able to effectively counter the entirety of HIV-1 diversity.

This review will discuss vaccine strategies that utilize sets of proteins for use as polyvalent immunogens. These strategies include bioinformatic designs of artificial proteins that can improve epitope coverage,<sup>19,20</sup> as well as lineage-based designs intended to recreate the immunological events that resulted in broadly cross-reactive neutralizing antibodies during natural infection.<sup>21</sup> These strategies begin to address the challenge of devising vaccines that can elicit responses capable of protecting against a significant fraction of the spectrum of highly diverse circulating HIV-1 strains.

## 2 | VACCINE APPROACHES TO CONTEND WITH HIV-1 DIVERSITY

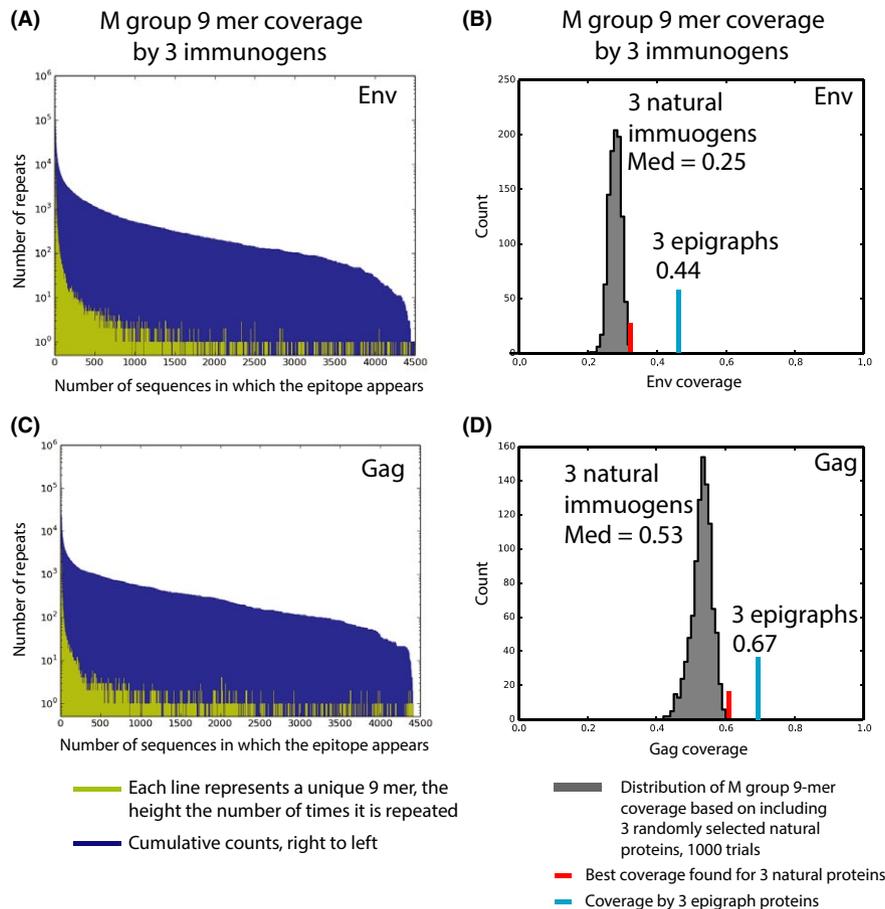
The highly diverse set of viruses that make up most of the global HIV-1 epidemic are classified into four groups (M, N, O, and P) based

on their phylogenetic relationships and zoonotic origins. Group M, the “Main” group of HIV-1, is further subdivided into several distinct clades (or subtypes), which have been labeled A–K.<sup>22</sup> In addition many inter-subtype recombinants are circulating, some of which establish important epidemic lineages.<sup>23</sup> A number of regional epidemics are dominated by a single clade—the B clade in North America, and the C clade in Southern Africa and India—but in most places in the world different clades and recombinant forms co-circulate.<sup>22</sup> The Los Alamos HIV database provides an interactive map showing the global distribution of different clades and circulating recombinant forms, as they have been sampled and sequenced (<http://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp>).

An early vaccine approach to cover HIV-1 diversity was to include one virus sampled from each of several major clades.<sup>24</sup> Unfortunately, the results of the human efficacy phase 2b trial HVTN 505, which tested a combination of three envelope (Env) immunogens from different clades as a vaccine, were disappointing.<sup>25</sup> The HVTN 505 vaccine included a multivalent (six plasmid) DNA prime, each plasmid expressing either a clade B *gag*, *pol* or *nef* or a clade A, B or C *env* gene, followed by a boost with four recombinant adenovirus (rAd5) vectors expressing either a clade B Gag/Pol fusion protein, or clade A, B or C Envs. The vaccine was tested in a population at increased risk of infection in the United States, a predominantly B clade HIV-1 epidemic. Hence, the vaccine-induced Gag- and Pol-specific T-cell responses had to act only in a within-clade context, while the polyvalent Env combination was designed to counter a more diverse set of viruses. This approach did not reduce either the rate of acquisition or set point viral load of new HIV-1 infections.<sup>25</sup>

One potential way to improve the breadth of vaccine responses relative to simply using a natural variant is to minimize rare amino acids at positions that might focus the immune system on epitopes that yield type-specific responses. A recently discovered class of rare mutations that can result in potent, but type-specific, neutralizing antibody responses are unusual gaps in the Env glycan shield.<sup>26</sup> For example, the loss of a highly conserved glycosylation site at position 241 (based on HXB2 numbering) in the BG505 Env resulted in such a “glycan hole”.<sup>27</sup> When the BG505 glycoprotein was delivered in a vaccine as a soluble, near-native SOSIP trimer, the resulting autologous antibody responses specifically targeted this rare glycan hole.<sup>27</sup>

There are several approaches that can be used to avoid targeting rare epitope variants. One approach involves Mosaic or Epigraph vaccines, which are artificial proteins that resemble natural proteins, but are designed in silico to maximize inclusion of the most common forms of linear epitopes.<sup>19,20</sup> By design, Mosaic vaccines disfavor inclusion of very rare amino acids at any given position. In addition, they disfavor unique local combinations of amino acids, including the loss or gain of rare potential N-linked glycosylation sites. Such rare amino acids and pairings of amino acid combinations in local regions are common, and present in virtually all naturally occurring HIV Env proteins<sup>20</sup> (Figure 1). The intentional minimization of such rare amino acids in Mosaic vaccines may help elicit greater breadth of not only T-cell but also B-cell responses. Indeed, a recent study showed that Env mosaic vaccines elicited both cellular and humoral responses, and that vaccine-induced



**FIGURE 1** The diversity of HIV-1 Env and Gag considered in terms of epitope-length fragments (9-mers). The left-hand panels (A and C) summarize the frequency of each unique 9-mer in the global 2015 HIV M group alignments curated at the Los Alamos database (<http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html>, 08/31/2015). Note the log scale on the y-axis. While the overall distribution is captured here, the very high numbers of completely unique epitopes that was brought out in the main text are difficult to discern, as the green line corresponding to number of distinct 9-mers that are only repeated once is very thin and on the extreme left. There are, precisely, 89 957 9-mers only found once of 136 828 unique 9-mers found in 4500 Gag sequences, and 556 401 9-mers found only once out of 719 578 unique 9-mers among 4600 Env sequences. The right-hand panels (B and D) show the coverage, as number of perfect 9-mer matches, that can be obtained with a 3-valent antigen vaccine. The coverage distribution shown in grey is based on 1000 trials, each a random selection of three natural strains from the data set, as a putative polyvalent immunogen. The best 9-mer coverage possible is indicated by a red line, and the median value for the set is noted. The coverage achieved by a 3-valent Epigraph vaccine is indicated by the blue line, and is expected to be comparable to a Mosaic design. The new epigraph tool suite was used to create this figure<sup>20</sup>

antibodies correlated with protection from acquisition in a SHIV challenge model.<sup>28</sup> A second approach to improve induction of cross-reactive antibodies is based on the idea of tracking B-cell lineage development in chronically infected subjects who have generated potent and broad neutralizing antibody (bNAb) responses. Envs that preferentially stimulate B-cell lineages along pathways known to have the potential to produce bNAbs are empirically identified and then utilized as immunogens, a strategy referred to as B-cell lineage-based design.<sup>21</sup>

### 3 | T-CELL VACCINE DESIGN STRATEGIES

The HIV-1 T-cell vaccine field has made significant advances since the first T-cell based vaccine failed to prevent infection or improve viral control in vaccinated individuals who became infected.<sup>29,30</sup> The

Ad5 vectored vaccines used in these trials may even have enhanced HIV-1 infection rates.<sup>29,30</sup> Despite these early setbacks, there are several reasons for renewed optimism regarding the value of T-cell approaches, in both preventative and therapeutic settings. These include the use of cytomegalovirus (CMV) vectors, the use of polyvalent Mosaic antigens, and the design of immunogens specifically based on highly conserved regions. Each of these approaches may be successful when used alone, but may have even greater potential when used in combination.

#### 3.1 | CMV vectors

A novel CMV vector developed by Louis Picker et al.<sup>31-33</sup> has elicited unusual and protective T-cell responses in rhesus macaques. Over 50% of SIVmac infections were effectively cleared in animals that

were vaccinated with SIV antigens delivered by this vector. To date, no other HIV or SIV vaccine has elicited immune responses that could suppress established infections to prolonged undetectable levels. The unique protection that CMV-vectored SIV vaccines have generated in this non-human primate (NHP) model has resulted in a seismic shift in our understanding of CTL-mediated immunity. By using rhesus CMV vectors that lacked particular genes (Rh157.5, Rh157.4, and Rh157.6), a prolific T-cell response was generated, and a very high density of epitopes was targeted in the SIVmac Gag protein that was included in the vaccine. The responses were restricted either by class II major histocompatibility complex (MHC) molecules,<sup>32</sup> or by the conserved, non-classical, ubiquitously expressed MHC-E protein.<sup>33</sup> The MHC-E presented epitopes were non-canonical, meaning that they generally did not contain previously defined MHC-E anchor motifs.<sup>33</sup> The biology underlying this highly unusual protection is not yet understood, but if the uncharacteristically large number of CTL responses induced by the rhCMV vectors contribute to the protection, designing HIV inserts that are likely to produce a large number of cross-reactive responses may be important for vaccine success. Hence, polyvalent HIV Epigraph vaccines<sup>20</sup> delivered in rhCMV vectors are currently being evaluated in preclinical NHP studies. Epigraphs are a second-generation Mosaic-like vaccine antigen design; like Mosaics, Epigraphs optimize T-cell epitope coverage but with additional features (see below).

### 3.2 | Computational design of polyvalent T-cell vaccines

A second important advance for promising T-cell vaccine strategies is the finding that polyvalent HIV-1 Mosaic antigens result in significantly greater breadth and potency of vaccine-elicited T-cell responses than do natural proteins in NHP studies.<sup>34–39</sup> Natural proteins used in the unsuccessful STEP/Phambili human trials elicited T-cells that targeted only a small number of epitopes in each vaccinated individual.<sup>35</sup> If the enhanced responses to Mosaic vaccines observed in NHP translate to humans, it is possible that Mosaics may improve outcomes in future vaccine studies.

The Mosaic code utilizes a genetic algorithm to design small sets of artificial proteins, typically two to four, comprised of complementary vaccine antigens for use in a polyvalent vaccine. Mosaics are intact proteins, so can be delivered in the same manner as natural proteins, or alternatively can be divided into conserved sub-regions of proteins. Mosaics sets collectively optimize coverage of diverse potential epitopes in a targeted population<sup>19</sup> for a given set size, or valency. Theoretically, such Mosaics offer much improved epitope coverage relative to combinations of natural protein sequences of the same valency.<sup>19</sup> In practice, when tested in NHP models, HIV-1 mosaic vaccines significantly improved both the number and potency of vaccine-elicited T-cell responses that cross-reacted with natural variants.<sup>36–39</sup> While the level of vaccine responses to HIV-1 proteins can be evaluated in NHPs, their protective capacity cannot be assessed, except for responses to the HIV-1 Env glycoprotein, which can be tested using simian-human immunodeficiency virus (SHIV) challenge models.<sup>40</sup> Using SIVmac proteins instead, the number of

vaccine-elicited SIVmac Gag CTL responses has been correlated with improved viral control in NHPs of SHIVs and SIVmac.<sup>41,42</sup> These observations suggest that the improved T-cell responses to HIV mosaics in NHPs may translate into better outcomes in human subjects.

A bivalent mosaic vaccine including pairs of HIV-1 Gag, Pol, and Env mosaics delivered in adenovirus and poxvirus vectors conferred significant protection against serial, low-dose rectal challenges with a difficult-to-neutralize simian-human immunodeficiency virus SHIV-SF162P3.<sup>28</sup> As noted above, the only mosaic proteins that would have been able to elicit cross-reactive protective responses in this vaccine were the two HIV-1 Env glycoproteins, since the HIV-1 Gag and Pol directed responses are unlikely to cross-protect against the SIVmac Gag and Pol proteins of the SHIV challenge virus. The Env responses were sufficient, however, to reduce viral acquisition as well as set point viral loads, improving animal survival. Antibodies to the HIV-1 Env Mosaic protein were elicited, and were correlated with protection against acquisition.<sup>28</sup>

That mosaic proteins elicited beneficial anti-HIV-1 Env antibodies<sup>28</sup> is not unexpected, even though they were optimized for linear epitope coverage. Mosaic Env glycoproteins fold well and retain the ability to interact with monoclonal antibodies that are known to target discontinuous Env epitopes.<sup>28,36,39</sup> The mosaic design also optimizes for coverage of local covariation patterns (within 9 amino acids), and even though antibodies generally bind discontinuous epitopes, such epitopes have linear components. Mosaics also tend to exclude rare amino acids at given positions, and rare local combinations of amino acids, while they favor inclusion of common variants. Thus, Mosaics tend to be immunologically more “central”. They also provide complementary coverage of the most common amino acid variants in a polyvalent formulation. Taken together, Mosaics should improve coverage of both B-cell and T-cell epitopes relative to natural strains.

It should be noted that in an SIVmac challenge study, where SIV (rather than HIV-1) mosaic Envs were evaluated, no vaccine-mediated protection by antibodies was observed in the mosaic Env vaccine group, while animals vaccinated with a natural protein showed reduced rates of acquisition.<sup>42</sup> The mosaic-Env-vaccinated macaques in this study did, however, have lower peak viremia, and reduced viral loads in early infection. There are a number of reasons why the SIV mosaics did not work as well as their HIV-1 mosaic counterparts in terms of preventing infection.<sup>43</sup> First, the SIV Env mosaic was designed based on a very small number (approximately 20) of SIVmac and SIVsmm Env proteins, compared to the hundreds or thousands of Env proteins used for HIV-1 designs. Second, the extensive diversity sampled in the SIV virus input set had evolved over hundreds of thousands of years in sooty mangabeys, and weaving together proteins separated by such ancient divergence times into Mosaic antigens may have been problematic.<sup>43</sup> Finally, the mosaic SIV Env glycoproteins might not have folded as well as their Mosaic HIV-1 Env counterparts. HIV-1 mosaic Envs have been extensively tested and shown to exhibit native antigenic profiles that are recognized by many neutralizing antibodies that bind discontinuous epitopes. Moreover, HIV-1 mosaic Envs undergo the appropriate conformational shift when bound to CD4,<sup>36</sup> and so are

assumed to form native-like proteins. However, the SIV mosaics were not similarly evaluated.<sup>42</sup> Also of note, SIV Gag mosaic vaccinated animals in this same study had reduced viral load and improved survival upon SIVmac infection.<sup>42</sup>

Given the encouraging NHP results overall, there are currently several mosaic antigen HIV-1 human vaccine trials underway, including HVTN 106 and HIV-V-A004/IPCAVD 009. A listing of all human HIV vaccine trials is maintained by the AIDS Vaccine Advocacy Coalition, AVAC, at <http://www.avac.org>. Mosaic vaccines have also shown promise against other highly variable pathogens, including Hepatitis C,<sup>44</sup> Ebola,<sup>45</sup> Influenza,<sup>46</sup> Chlamydia,<sup>47</sup> and foot-and-mouth disease (W. Fischer and E. Rieder, personal communication).

To further improve mosaic vaccines, a second-generation design strategy was recently developed to create mosaic-like vaccines, called Epigraphs.<sup>20</sup> Epigraphs use a computationally highly efficient graph theory approach to solve essentially the same design problem as Mosaics, thereby reducing computer run times from hours to seconds. Thus, the increased computational efficiency allows the exploration of subtle vaccine-design problems that were previously intractable using the Mosaic code. These include exploring the coverage impact of excluding rare epitopes from the vaccine antigens, and exploring coverage optimization including imperfect matches between potential epitopes as well as perfect matches. In addition, Epigraphs enable a strategy to develop a set of antigens for a tailored approach in the context of therapeutic vaccines.<sup>20</sup> In this scenario, an infected subject's viruses could be sequenced, and those sequences used to inform the vaccine selection. As a new vaccine could not feasibly be produced for every subject, a reasonable number of immunogens could be produced (eg, six), and the two best matches with the subject's HIV-1 sequences from that manufactured set of six could be delivered. These could be selected to maximize epitope matches and minimize mismatches between the vaccine and the within-host viral quasispecies.

One of the outputs of the Epigraph tool suite (<http://www.hiv.lanl.gov/content/sequence/EPIGRAPH/epigraph.html>) is a count of each unique potential epitope (9-mer) in a dataset<sup>20</sup>; this is the fundamental information that forms the basis of the nodes in the graph, and it is useful for identifying conserved regions from the perspective of epitope length fragments. This count is also informative in that it provides a more immunologically relevant, albeit more sobering, lens with which to view HIV-1 diversity compared to a standard phylogeny (Figure 1). Env is the most variable protein in HIV-1, and the target of choice for vaccine-induced antibodies. The 2015 Los Alamos HIV database "filtered" Env alignment (filtered means it excludes sequences with multiple errors or ambiguities, and the alignment includes only one sequence per infected person) has 4600 sequences, each between 850-890 amino acids in length. Found within those sequences are 720 000 distinct 9-mers. An astonishing 560 000 of those 9-mers occur only once in the whole set. This means that every HIV-1 Env protein contains, on average, approximately 120 completely unique 9-mers not found in any other sequence in the database. Only 53 9-mers in Env occur in over 80% of the sequences. Here, we also consider, Gag, one of the more conserved proteins in HIV-1. Gag is about 500 amino acids long, and is an important target of CTL responses.<sup>48</sup>

Among 4500 Gag sequences in the 2015 database alignment, there are 137 000 distinct 9-mers, 90 000 of which are unique. Thus, there are approximately 20 completely unique 9-mers per Gag sequence, and only 71 9-mers occur in over 80% of the sequences. Given this, it is clear that Env proteins will have ample opportunity to generate autologous type-specific responses, and that even the more conserved Gag has very serious issues related to epitope diversity. Since immune responses can tolerate some variation in their target, these estimates, which are based on the perfect identity of potential linear epitopes, somewhat exaggerate the challenge. Nonetheless, it is prudent to take them into consideration when designing polyvalent vaccines.

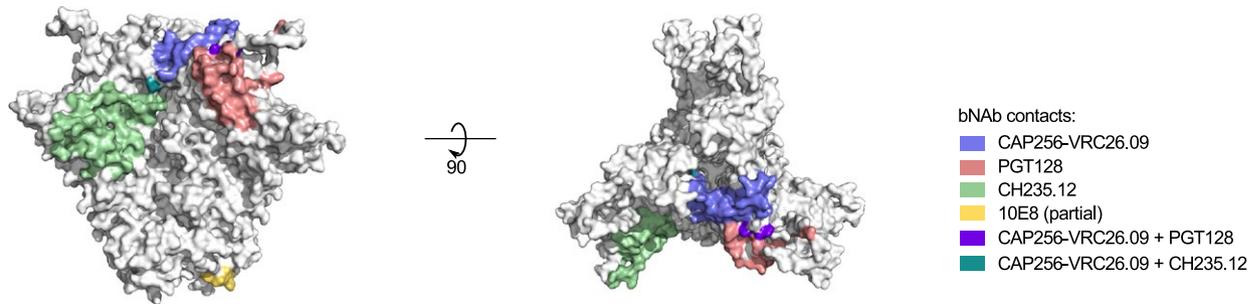
### 3.3 | Combining conserved region and polyvalent approaches

A third area of interest for HIV-1 T-cell vaccine design is the idea of focusing responses on epitopes that are presumed to be beneficial because they are either highly conserved,<sup>49-52</sup> require compensatory mutations to escape,<sup>53</sup> or are associated with better viral control in natural infections.<sup>54</sup> While particular epitopes can be identified that are good candidates for such design strategies, vaccines based on the delivery of concatenated epitopes have not been immunogenic in humans.<sup>55,56</sup> In contrast, immunogens that link larger conserved protein regions, which provide at least some natural context for processing within the larger peptide, have elicited strong CTL responses in preclinical studies<sup>50</sup> and in a human Phase I trial.<sup>57</sup> This inspired a second-generation conserved-region approach using the Mosaic tool suite,<sup>58</sup> first to define the most conserved regions in the HIV-1 proteome, and then, to design complementary paired Mosaic vaccines that span the conserved regions with vaccine antigens that provide optimal epitope coverage.<sup>51,52</sup> Because even the most conserved regions of HIV-1 vary at the epitope level, polyvalent Mosaic approaches are useful to optimize variant coverage.<sup>34</sup> In a recent mosaic vaccine described by Ondondo et al.<sup>51</sup>, included regions were restricted not only by being the most conserved regions, but were further limited to include only regions that spanned epitopes that were deemed protective, in that responses to these epitopes were associated with low viral loads in natural infections.<sup>54</sup> Indeed, it was found that infected people with enriched responses to epitopes embedded in the second-generation conserved region vaccine had lower viral loads, supporting the utility of the strategy of immune response focusing.<sup>51</sup>

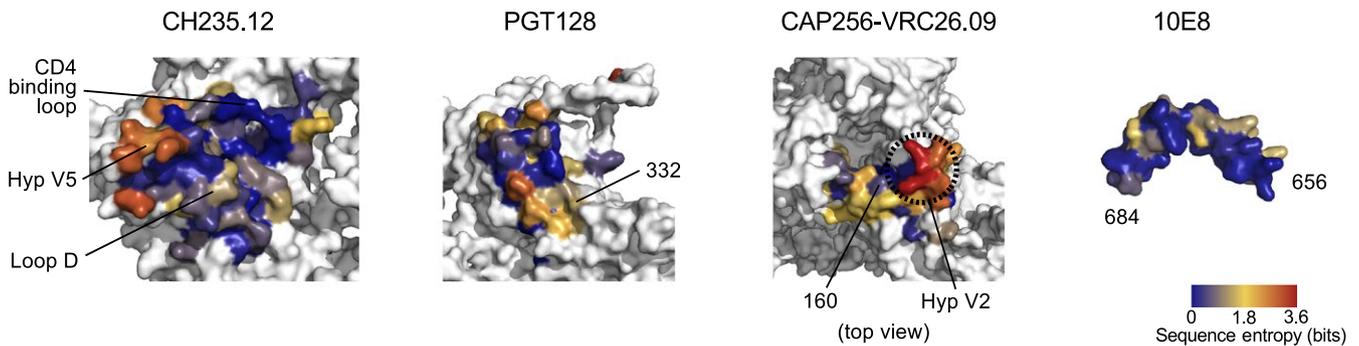
## 4 | ANTIBODY EPITOPE DIVERSITY

B-cell clonal lineages have an additional capability relative to T cells in that they can adapt to sustained generation of viral diversity through somatic hypermutation. Thus, at any given moment in an infected person, a swarm of viral variants is interacting with a swarm of antibody variants. Selection in this scenario goes both ways. The outcome for both antibody and virus is driven by the frequencies of their variant forms in the population, their relative levels of affinity, and the antibody's functional capacity to counteract and select against

## (A) bNAb contacts



## (B) Sequence entropy for contact sites



**FIGURE 2** (A) Contact regions of representative bNAb that bind four regions commonly targeted by bNAb isolated from different individuals. The contact regions are shown on the crystal structure of a subtype G Env SOSIP trimer<sup>60</sup> (PDB: 5FYJ), colored according to the bNAb epitope, with different colors for overlapping contact sites of multiple bNAb. For calculations of bNAb contact regions, crystal structures for antibodies CH235.12<sup>68</sup> (PDB:5F96), PGT128<sup>89</sup> (PDB: 5C7K) and 10E8<sup>90</sup> (PDB:4G6F), and a structural model for CAP256-VRC26.09<sup>62</sup> were used. A generous cutoff of 8.5 Å between heavy atoms of antibody and Env was used to define contacts, so as to capture the full region where Env amino acid substitutions might directly impact binding. For 10E8, the core epitope is an alpha helix between positions 671–683 in HIV-1; this region is not part of the trimer structure used here. HXB2 positions 656–665 are highlighted here, however, as 10E8 contacts sites, since they are in the structure and within 8.5 Å of the bound 10E8 surface. (B) Sequence entropy of bNAb contact sites. Env sequences from a often-used 207 M-group virus neutralization panel were used to calculate sequence entropy, and deletions were included in the entropy calculation. Entropy scores are mapped on the crystal structure for a subtype G Env SOSIP trimer<sup>60</sup> (PDB:5FYJ) and the MPER peptide<sup>90</sup> (PDB:4G6F) using the color scheme indicated, which uses blue for low entropy (high sequence conservation), yellow for intermediate entropy and red for high entropy (high sequence variation). The view of CAP256-VRC26.09 contacts is from the top as in right panel in (A). Hyp = hypervariable

viral variants. When intensive longitudinal sampling was undertaken, “cooperative” autologous neutralizing antibody lineages were observed<sup>13</sup> (M. Bonsignori and B.F. Haynes, unpublished data), adding a further dimension of complexity to this system. In these studies, it was shown that escape from one B-cell lineage can lead to developing sensitivity to another, and the development of neutralizing antibody breadth was facilitated by these interactions.

bNAb are defined as those antibodies that potently neutralize a broad spectrum of heterologous viruses. The most potent bNAb, all of which have been isolated from infected individuals, have been found to frequently target a relatively small set of well-defined regions on the surface of the Env. Here, we discuss representative bNAb that bind to one of four commonly targeted regions<sup>59</sup>: the CD4 binding site (CD4bs), the membrane proximal external region (MPER), the V2 glycan region (V2g), and the V3 glycan region (V3g). Figure 2A shows the Env contact regions for representative bNAb from each class on the crystal structure of a subtype G SOSIP trimer.<sup>60</sup> bNAb with shared

epitope specificity tend to share reactivity profiles across large Env panels.<sup>61</sup> This results at least in part from bNAb critically interacting with the same structural motif in the targeted region. The two clearest examples of this phenomenon (Figure 2B) are the V2g bNAb (eg, CAP256), which generally target the N-linked glycan at Env position 160, eg, Gorman et al.<sup>62</sup>, and the V3g bNAb (eg, PGT128) which generally target an N-linked glycan at position 332, eg, Kong et al.<sup>63</sup>

Because of the cross-reactive potential of bNAb, their epitopes are often described as conserved.<sup>64</sup> In terms of structural interactions between a bNAb and the spectrum of sensitive Env variants, this is a valid point. However, while certain Env positions in some parts of these epitope regions are highly conserved, other positions that are also in close proximity or embedded directly in the antibody-binding site can be highly variable. CD4bs bNAb such as VRC01 and CH235.12 (Figure 2B), clearly illustrate this. VRC01 binds to highly conserved amino acid residues which are almost invariant and maintained across viruses, critical for CD4 binding and viral entry (eg, the

conserved CD4 binding loop contacts shown in Figure 2, that are important for all VRC01-like antibodies).<sup>65,66</sup> However, VRC01 also directly contacts amino acids in the hypervariable part of V5 (positions 460–465).<sup>65,66</sup> This short region of Env is extremely variable, and laden with insertions and deletions to the extent that it is extremely difficult to align in a meaningful way. Thus, VRC01 comes into close contact with both conserved and extremely variable regions of Env, and yet VRC01 maintains great breadth and potency. We argue that mutations in the high diversity regions in Env drive autologous immune escape *in vivo*, and that this is an essential component of the factors that collectively induce neutralization breadth.<sup>13,67</sup> The same holds for other CD4bs bNAbs, as well as other bNAb classes. Figure 2B offers a high-level view of this, but more a comprehensive and detailed accounting of bNAb contacts juxtaposed with viral diversity measures can be readily explored using the HIV database genome browser ([http://www.hiv.lanl.gov/content/sequence/genome\\_browser/browser.html](http://www.hiv.lanl.gov/content/sequence/genome_browser/browser.html)).

It should be emphasized that the concept of sequence conservation interleaved with extreme variability in bNAb contact residues offers an entirely different perspective to the commonly held view that bNAbs bind to conserved epitopes. In this alternative view, bNAbs are not broad because the epitopes are highly conserved. Rather, selective pressures from effective autologous neutralizing antibodies that arise during natural infection drive mutations that confer some level of resistance *in vivo*. B cells are then selected that can still bind the resistant viruses, and these interactions lead to enhanced heterologous neutralization breadth, as many of these same antibody-resistance mutations are found in the heterologous virus population. In this view, bNAb lineages start with limited neutralization breadth, which increases as they are selected for their ability to tolerate epitope variation, by exposure to the diversity that arises as a consequence of continuous cycles of autologous immune escape and selection for recognition.<sup>13</sup> If these mutations do not reduce viral fitness, and if the respective epitopes are commonly targeted, then the sampled escape mutations will be reflected in the diversity found at the population level.

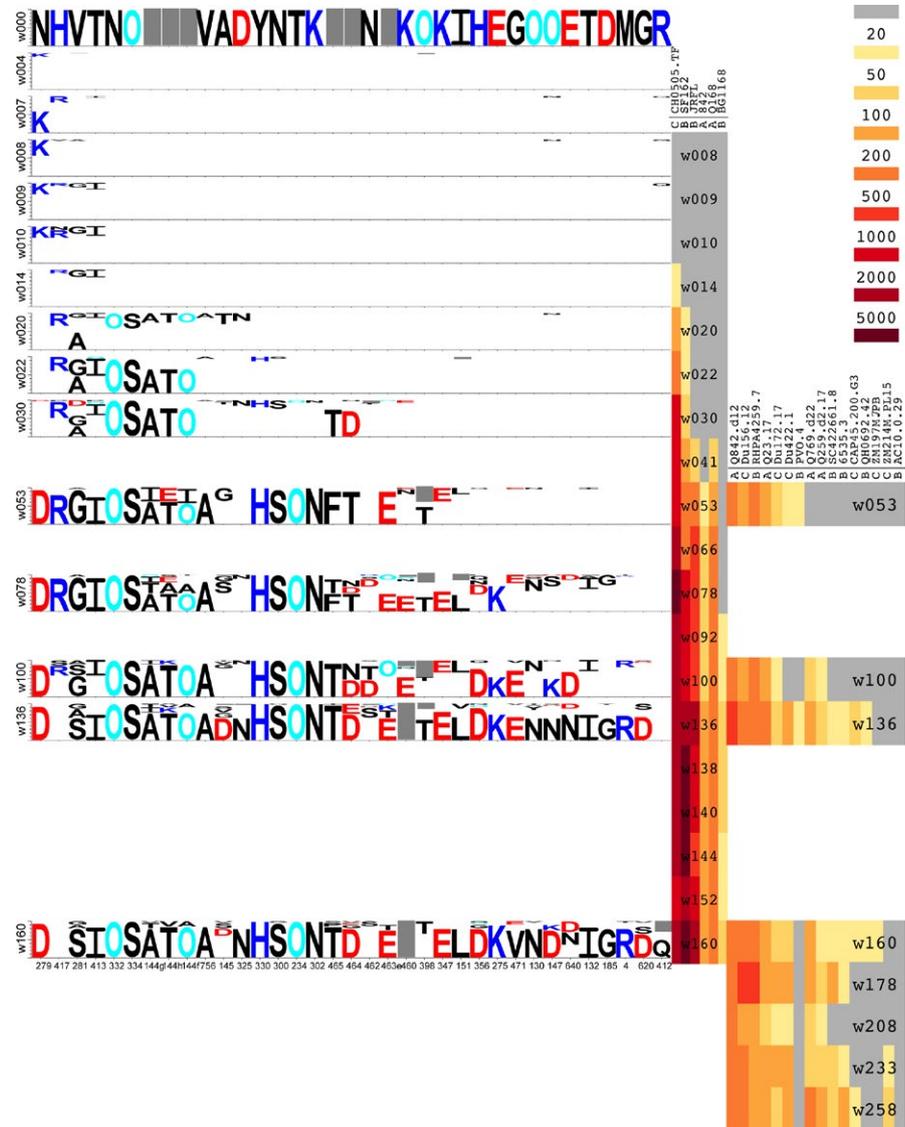
This is an important distinction in outlook, because it directly impacts the priorities for vaccine design. If, as we propose, the exposure of B-cell lineages to viral mutations that confer relative resistance is necessary to select for antibodies able to overcome that resistance, then it follows that polyvalent vaccines representing immunologically relevant epitope diversity may be essential to achieve breadth and potency. This hypothesis is supported by the *in vivo* development of antibody breadth, as it has been traced over years in infected subjects who are followed over time.<sup>11,13,68,69</sup> The B-cell lineages that eventually produce bNAbs indeed start out with limited or no heterologous neutralization breadth, but can bind the autologous virus that stimulated them. These interactions eventually drive selection of epitope variants. Importantly, epitope diversification arises *in vivo* prior to the development of heterologous breadth and concurrent with expansion of autologous breadth<sup>11,13,68–70</sup> (also M. Bonsignori and B.F. Haynes, unpublished data). These observations, plus the general observation that serum neutralization breadth is overall slow to develop during human infection<sup>71</sup> (see Figure 3 for an example), supports the premise

that exposure to epitope diversity may be an essential aspect of the evolution of antibody breadth *in vivo*.

One polyvalent vaccine design strategy that derives from these observations is a lineage-based design,<sup>21,68,72</sup> where the Envs selected for use as immunogens are intended to recapitulate the events that induced bNAb development in the infected individual. This strategy involves tracking the co-evolution of Env and antibodies in HIV-infected individuals who eventually develop potent bNAbs over years of infection. In one take on this strategy, an Env prime is identified with the ability to bind to and stimulate the B-cell germline, then other Envs are included for serial boosting, which are chosen based on their ability to bind with high affinity to key antibody intermediates along a particular B-cell evolutionary lineage.<sup>21</sup> Supporting this idea is the observation that bNAbs can share  $V_H$ -gene-restriction. For example CD4bs bNAbs that are CD4 mimics typically arise from one of two highly similar  $V_H1-2$  and  $V_H1-46$  genes.<sup>73</sup> In such a case, an initial triggering of a germ-line B-cell carrying an appropriate  $V_H$  heavy gene is likely to be an important first step for vaccine-induced CD4bs bNAb development,<sup>68</sup> and it is possible that particular Envs have a propensity to bind to these germ-line B-cells. The choices of prime and successive boost Env immunogens using the lineage-based design approach are illustrated below using the data from the subject CH505, who developed CD4bs bNAbs.

In contrast to an affinity-based B-cell lineage-design strategy, it will also be important to test the hypothesis that exposure to common Env variants, even if they do not represent the highest affinity binders of B-cell lineage members, may be critical for inducing neutralization breadth. One such strategy is a DNA-based swarm vaccine (designed according to the CD4bs bNAb ontogeny in subject CH505), which includes many natural variants.<sup>74</sup> Still another strategy is to trigger appropriate B-cell lineages with the potential to develop breadth, but then boost with specific Env variants chosen to represent more neutralization-resistant Env forms, which are modestly but not completely resistant to antibodies in the developing lineage. These variations of the lineage-design strategy are intended to break down specific barriers to expanded heterologous breadth and potency in a polyvalent vaccine. The inclusion of variants with some level of resistance is based on the premise that exposing the evolving B-cell lineages to these variants will select for antibodies better able to recognize them (see the V3 glycan bNAb vaccine example below, modeled on bNAb development and expansion of breadth in subject CH848). This approach is distinct from a pure B-cell lineage-based vaccine design, which focuses only on the Envs with the highest affinity, but both could start with the same Env, selected to trigger an appropriate germline B cell.

Vaccine strategies that induce antibodies with other anti-viral functionalities, apart from neutralization, are also important to consider. While bNAb induction remains a key focus in the HIV-1 vaccine field, an immune correlates analysis of the RV144 trial in Thailand,<sup>75</sup> the only successful HIV-1 vaccine trial to date, with a modest efficacy of 31.2% (95% CI, 1.1–52.1), raised the possibility that V2-directed antibody dependent cellular cytotoxicity (ADCC) mediated a protective effect.<sup>76</sup> Non-neutralizing antibodies can limit the number of viruses



**FIGURE 3** Env diversity in subject CH505 accompanies development and expansion of heterologous neutralization breadth. (Left) Frequency of mutations among sites in CH505 sequences with at least 80% TF loss in any time point sampled through week 160; these are the sites we consider candidates for being under the greatest selective pressure from the immune response. (Right) Breadth develops over longitudinal plasma neutralization ID<sub>50</sub> against Tier 1 (autologous CH0505.TF, then B|SF162 through B|BG1168) and Tier 2 (A|Q842.d12 through B|AC10.0.29) Env-pseudotyped viruses. Heatmaps summarize neutralization ID<sub>50</sub> values

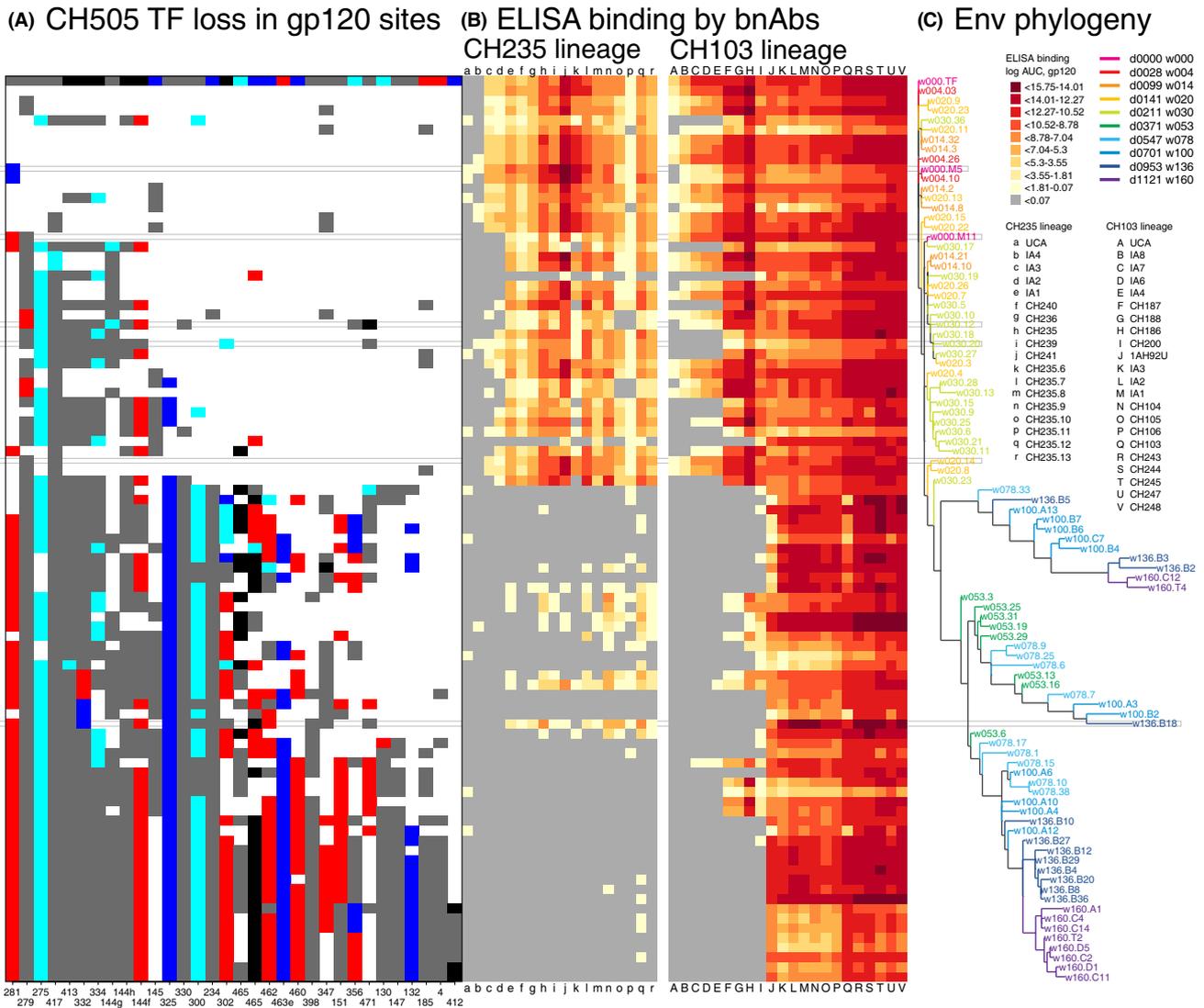
in passive protection studies.<sup>77</sup> The challenge of contending with HIV-1 diversity remains, regardless of the functionalities that protective antibodies employ. Some of the approaches outlined here could also be applied to developing vaccines that target other antibody functionalities.

## 5 | IMMUNE SELECTION IN LONGITUDINAL SAMPLES: ONTOGENY OF CD4BS BNABS IN SUBJECT CH505

As summarized above, one approach for vaccine development is to recapitulate key events in bNAb induction that have actually occurred in a specific HIV-1 infected individual who made particularly good bNAb response. A case in point is the study of subject CH505, for whom Env-antibody coevolution was chronicled starting at the time of acute infection.<sup>11</sup> Extensive antibody and viral sequence data, as well as immunological data were collected, and this yielded a rare view of the complexities involved in the development of CD4bs BNABs.

One lineage that developed in CH505 culminated in a bNAb designated CH103, which neutralized about half of a panel of 196 viruses chosen to represent genetic and geographical diversity.<sup>11</sup> Another lineage, designated CH235, eventually developed greater breadth and potency. The unmutated germline ancestor and ancestral intermediates of these B-cell lineages were inferred, synthesized, and their neutralization and binding sensitivity were tested against panels of autologous and heterologous Envs.<sup>11</sup> The gains and losses in binding affinity and neutralization during bNAb ontogeny were evaluated, point mutations and insertions in Env were identified that were associated with neutralization escape, and the viral diversification was characterized as the viral quasispecies evolved in response to immune selection.<sup>11,67</sup>

CH103 maturation was initially facilitated by escape mutants, which were selected by early antibodies in the cooperating CH235 lineage<sup>13,68</sup> (Figure 4). Unlike CH103, CH235 is a CD4 mimic, similar to the VRC01-class of CD4bs BNABs.<sup>73</sup> Development of the CH235 lineage breadth progressed over 6 years of longitudinal sampling. CH235 antibodies eventually acquired over 20% somatic hypermutations, and could



**FIGURE 4** Env mutations, antibody binding phenotypes, and the phylogeny representing CH505 sequences selected for immunological testing. (A) Accruing mutations among a subset of 32 LASSIE selected gp120 sites are indicated as colored boxes. Blanks match the TF, red and blue depict negative and positive charge change, black shows insertions/deletions, cyan a gain of an asparagine in a potential N-linked glycosylation motif, and gray shows other mutations. (B) Log AUCs of ELISA binding are displayed as heatmaps to show the relative affinities for the CH235 and CH103 bnAb lineage members, one column per antibody and one row per Env. Deeper reds indicate higher affinity, gray undetected. Each row in this figure follows the ordering of leaves in the tree on the right (C), which was made by maximum likelihood (in phyML, with HIVw+ $\gamma$ 4+I) from gp120 sequences. The CH505 TF sequence was used to root the tree. Note that these sequences were selected for experimental evaluation prior to the availability of LASSIE. The six lineage-based Envs [ref. 68] are outlined

neutralize 90% of diverse HIV-1 isolates.<sup>13,68</sup> Of note, high levels of somatic hypermutation were required for breadth, despite an early, nearly optimal binding orientation of the CH235 lineage. Other CD4bs-directed antibodies face similar developmental impediments, whether  $V_H$ -gene restricted or CDR H3-dominated.<sup>73</sup> Studying coevolution of HIV-1 with two bnAb lineages and their cooperative interactions<sup>13,68</sup> has led to unique insights into the complexities of bnAb development and suggested vaccination strategies for simultaneous induction of both lineages.

## 5.1 | From quasispecies to swarms

Longitudinal studies of Env-antibody coevolution such as those undertaken for CH505 involve sequencing hundreds to thousands of

Env variants from one subject. It is, however, feasible to test binding and neutralization for only a fraction of these. We therefore developed a method to select a set of variants that represents the gradual acquisition of selected mutations among longitudinally sampled sequences (Longitudinal Antigenic Sequences and Sites from Intra-Host Evolution, or simply LASSIE).<sup>74</sup> LASSIE uses loss of the transmitted founder (TF) viral sequence to identify specific sites under positive immune selection, and then selects Envs that represent the gradual acquisition of all recurrent mutations at the selected sites for developing reagent panels.

We applied LASSIE retrospectively to CH505 data for a performance assessment.<sup>74</sup> TF loss above 80% at any given time point identified 35 sites under immune selection, from among 398

single-genome-amplification (SGA) derived Envs that were sampled over a period spanning the first 3 years of infection. These sites corresponded to verified immune targets in this donor. The algorithm next identified 54 Envs that represent all recurrent mutations in the 35 selected sites in CH505. Importantly, an Env was selected to represent each of these mutations as they initially occurred, so that each mutation of interest was represented in the 54 Env set in the context of the natural Env where it was first observed. These variants are more likely to display an immunologically natural phenotype, compared to, for example, an out-of-context mutation introduced into the TF virus by site-directed mutagenesis. The 54 Envs were well-dispersed throughout the CH505 phylogeny, and represented the development of antibody binding and neutralization observed in a larger set of previously studied handpicked Envs<sup>11,13,74</sup> (Figure 4). The LASSIE algorithm chooses sequence sets with more recurrent mutations represented, and less redundancy, than random or “by hand” selection.<sup>74</sup> Thus, the algorithm objectively provides a minimal, manageable number of Envs to represent diversity in natural infection, to aid in the study virus-antibody coevolution. We call this minimal representation of antigenic diversity an “antigen swarm.” We initially developed LASSIE to explore mutational patterns *in vivo* and for reagent design in longitudinal studies. However, given the emergence of new vaccine technologies that may enable the use of high valency antigen cocktails, this approach could also be used to design a vaccine that mimics viral evolution in an individual who made potent bNAbs responses.<sup>74</sup>

## 5.2 | Lineage-based vaccine design and swarm immunogens

In patient CH505, plasma began to show increasing neutralization of Tier 1 (easily neutralized) viruses through the first year of infection, followed by a gradual development of breadth against Tier 2 viruses that peaked by week 136, within the first 3 years of infection (Figure 3). A small set of candidate Envs to use for lineage-based immunogens were selected based on this data, including four natural variants and two mutants with high affinity for the UCAs of the CD4bs lineages (Figure 4).<sup>68</sup> The selection of this set was informed by the coevolutionary dynamic of escape from immune surveillance. The Env variants chosen for the vaccine represented a gradual progression of binding affinity and neutralization sensitivities, starting with the germline B-cell precursor and then moving on to intermediate antibodies. Representation of mutations found in CH505 from among sites that were established to be involved with immune escape was also considered.

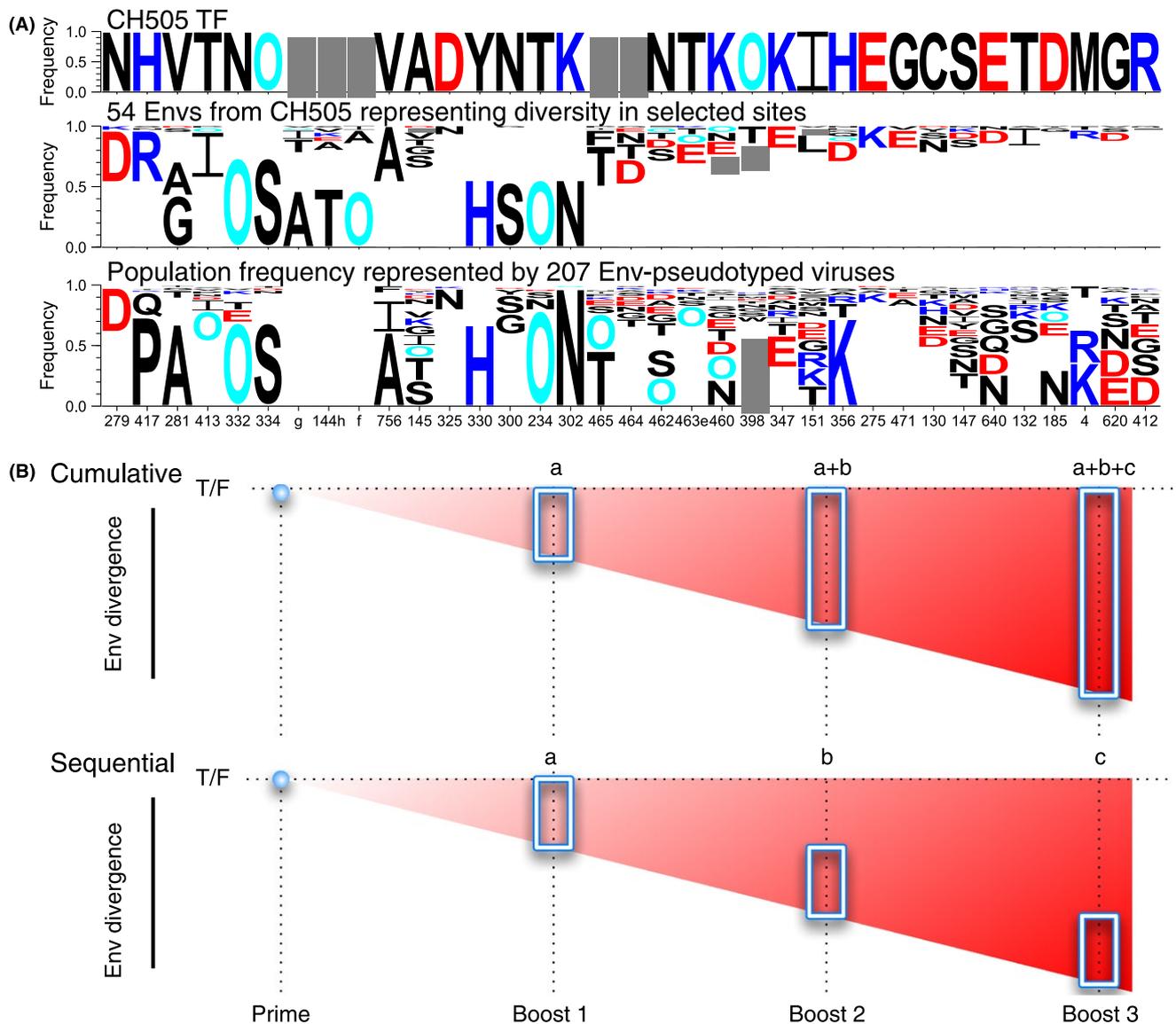
The lineage-based vaccine antigen series was thus carefully selected with the goal of enabling lineage-based selection of bNAbs, to attempt to recreate the known evolutionary trajectories of the CD4bs bNAbs in CH505. Studies to test this vaccine are underway in animal models. Still, the diversity that arose *in vivo* in CH505 is greatly underrepresented by these four Envs. The global diversity of heterologous viruses that vaccine-elicited antibodies must ultimately interact with to be effective is far greater. It is not yet known if a representation of that greater diversity in a vaccine will be important, or necessary, for the development of heterologous neutralization breadth. Therefore,

our group is beginning to explore ways to deliver larger sets of Envs as an alternative vaccine strategy.<sup>74</sup>

To visualize the impact of epitope diversity in the CH505 example under discussion, we considered the diversity coverage of heterologous diversity at sites under selective pressure in CH505 by a hypothetical vaccine comprising the full antigenic swarm of the 54 LASSIE-identified Envs that represents the accumulation of mutations in immunologically relevant sites (Figure 5A) to comprise a hypothetical vaccine. Mutations in the sites away from the TF (by definition) accumulate over time (Figure 4A), in conjunction with changing immune susceptibility to antibodies in the two CD4bs lineages (Figure 4B). We compared the diversity coverage in CH505 LASSIE sites in 54 Env CH505 swarm with the variation found among the 207 M group pseudoviruses used to assess antibody breadth and potency of subject CH505 antibodies.<sup>68</sup> This set of 207 M group viruses is a diverse global sample that represents all major HIV-1 clades. The LASSIE swarm, based solely on the evolution within the single patient CH505, makes significant inroads into sampling the diversity seen in relevant sites among the 207 diverse heterologous viruses (Figure 5A). The CH505 quasispecies does, however, carry some rare amino acids that occur very infrequently in the general population, and the full 54 Env CH505 swarm Envs does not cover all of the common M group variants. For example, position 417 in Figure 5A (second from the left), is mostly Arg (R) or His (H) in subject CH505, usually Pro (P) or Gln (Q) in the pseudotyped virus panel. The CH505 Envs were deliberately chosen from a within-host infection that yielded heterologous breadth, with the working premise that CH505 lineage-based or swarm-based vaccine will also elicit broad responses. On the other hand, a deliberate exposure to more common amino acid variants in the circulating population may also improve both the breadth and potency of the responses relative to designs that are based exclusively on CH505. One way to achieve such exposure through vaccination might be through a lineage-based design that is followed by a mosaic boost. This possibility is addressed in an example below, from subject CH848.

## 5.3 | Swarm immunogen delivery

There are serious technical and regulatory challenges associated with delivery of a highly multivalent vaccine; still the concepts may merit exploration in animal models. DNA administration, which exploits host cellular processes for antigen expression,<sup>78–80</sup> is the simplest approach that may be feasible. In addition to administration of naked DNA with electroporation, several biotechnology alternatives exist as DNA vaccine delivery methods, and new technologies continue to offer promise.<sup>81–84</sup> Another unresolved question is how best to deliver such a complex set of antigens. Figure 5B illustrates two alternatives for increasing immunogen diversity using serial boosts, again using CH505 as an example. We would start with a prime intended to trigger appropriate B-cell lineages. In the case of CH505, this prime would include three Envs, starting with the TF and two mutant Envs that bound with high affinity to the unmutated ancestors of the CH103 and CH235 lineages.<sup>68</sup> We would divide the swarm of 54 Envs

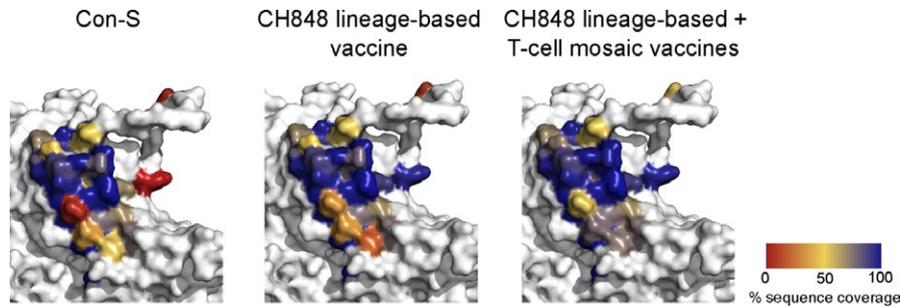


**FIGURE 5** (A) Sequence logos summarize variant amino acid frequency in 35 sites selected for having high TF loss in CH505. The CH505 TF is shown at the top. The letter O represents an asparagine in a potential N-linked glycosylation motif. A grey box is a gap. Red indicates a positively charged amino acid; blue, negative; and black, all other amino acids. The sites shown are the 35 sites selected by LASSIE for evidence of positive selection over time in CH505. In the sequence logo plots, white space stands in for the TF amino acids at the top; these are left out of the logo to emphasize differences. The height of the letter in the sequence logo indicates the frequency in the population. (B) Sequential and cumulative diversity increase for swarm immunization illustrated using a CH505 example. A priming immunization for a lineage-based vaccine modeled on CH505 includes 3 Envs, the TF Env, plus two addition Envs with single mutations introduced that increase binding susceptibility in the Loop D of the CD4 binding site.<sup>68</sup> In subsequent boosts, the 54 swarm immunogens selected using LASSIE would be divided into three sets of 18 each, the first set (*a*) the least diverse and favoring earlier time, with the subsequent two sets (*b* and *c*) to represent increasing divergence sampled over time in the subject

into three different groups of boost immunogens, including 18 Envs each (*a*, *b*, and *c*). These groups gradually increase sequence diversity among sites known to be associated with immune escape in CH505. Envs in group *a* are the set sampled earliest in infection, and have the fewest mutations relative to the CH505 TF. Groups *b* and *c* have progressively greater levels of diversity and later sampling times. These sets could either be administered either sequentially (*a*, followed by *b* alone, and finally *c* alone) or cumulatively (*a*, followed by *a* and *b*, then finally *a* and *b* together with *c*) (Figure 5B).

## 6 | V3 GLYCAN BNAB ONTOGENY IN SUBJECT CH848: LINEAGE PLUS DIVERSITY DESIGN

Another roadmap for lineage-based vaccine development was the ontogeny of bNAbs in subject CH848, who was infected by a single clade C TF virus. CH848, like CH505, was followed from early infection for many years, spanning the period when broadly neutralizing antibodies against heterologous strains began to develop roughly 3.5 years into



**FIGURE 6** Structural mapping of global sequence coverage at V3 glycan bNAb contact sites by candidate vaccine immunogens. Position-wise sequence coverage is calculated as the percentage of sequences in a commonly used 207 M-group virus neutralization panel that have the same amino acid as at least one of the sequences for a given vaccine strategy. Vaccine strategies considered are: Con-S, an M group consensus; the CH848 is a 5-valent CH848/DH270 lineage-based design, and a 3 valent Env mosaic.<sup>36</sup> Con S and the mosaic Envs are currently being evaluated in the HVTN106 Phase I human trial. Color-coding for sequence coverage is indicated by the color bar, with red shades indicating low, yellow indicating intermediate and blue indicating high sequence coverage. Crystal structure for a subtype G Env SOSIP trimer<sup>60</sup> (PDB:5FYJ) was used, and PGT128 contacts from Figure 2 are used. This figure shows a close-up view of the PGT128 contacts as in Figure 1B

the infection (M. Bonsignori and B.F. Haynes, personal communication, and M. Bonsignori and B.F. Haynes, unpublished data). Over 1200 Env sequences were generated from CH848 plasma viruses, collected at 26 time points over 5 years. Three antibody lineages were isolated that each showed dependence on the N-linked glycosylation site N332, the hallmark of V3 glycan antibodies. Two of these lineages, DH272 and DH475, did not develop heterologous neutralization breadth. The third, DH270, eventually developed the capacity to neutralize 55% of a diverse panel of 207 M group Envs in a TZM-bl assay. The evolutionary route to achieve this breadth was complex and again involved cooperative interactions. The DH272 and DH475 B-cell lineages expanded, and the virus in subject CH848 was selected for resistance to these antibodies. The escaped viruses eventually provided the sensitive Env forms that stimulated the DH270 antibody lineage that ultimately developed breadth. A lineage-based vaccine strategy based on the development of DH270 bNAbs in CH848 is currently being evaluated (M. Bonsignori and B. Haynes, personal communication, and M. Bonsignori and B.F. Haynes, unpublished data). As in the CH505 example, this design concept features particular Envs that were selected with the intent of recapitulating key events in DH270 lineage development, triggering an appropriate B-cell germline with the potential to develop neutralization, and then including a polyvalent set of antigens with the goal of selecting for antibodies with increasing breadth.

The CH848/DH270 lineage-based design includes a new feature relative to CH505. Changes in the hypervariable V1 region of Env appeared to be critical in the ontogeny the DH270 in subject CH848. The TF virus in CH848 had an atypically long V1 loop; the two cooperating lineages, DH272 and DH475, bound and neutralized the TF and other early CH848 viruses that carried this long V1. Shorter V1s were selected, as they conferred resistance to one of these early cooperating antibodies. In contrast, the earliest antibodies of the DH270 lineage, which we hope to emulate with a vaccine, targeted only Envs with short V1 loops, and the DH270 B-cell lineage was not triggered until viruses with short V1 loops dominated the CH848 quasispecies. After the DH270 lineage was initiated, longer V1 loops began to be replenished in the viral quasispecies; these were more resistant to early

DH270 antibodies generally, so were likely to have been selected by DH270 antibodies. Viruses with longer V1s were increasingly sensitive to later members of the lineage, which were in turn likely to have been selected to tolerate them. The patterns of V1 loop length restrictions on neutralization sensitivity were perfectly mirrored between autologous and heterologous viruses, and only the late DH270 antibodies could neutralize heterologous viruses with moderate or long V1 loops, and this impacted their breadth. Thus, for a DH270 polyvalent lineage-based vaccine, we selected a CH848 Env with short V1 length Env and high affinity for very early DH270 lineage-members as a prime, and sequentially add boosts with increasingly longer V1 loop lengths (B. Korber, P. Hraber, and B. Haynes, unpublished).

## 6.1 | Augmented vaccine compositions

If lineage-based design strategies as exemplified by those described for CH848 (M. Bonsignori and B. Haynes, personal communication) and CH505<sup>68,74</sup> above can indeed trigger B-cell lineages that initiate heterologous breadth at the population level, they may also benefit from a boost with a mosaic vaccine, comprised of three antigens designed to best capture variation in linear epitopes at the population level.<sup>36</sup> A case for this is made in Figure 6. Some residues in the PGT128 contact surface (thought to be similar to DH270.6 contacts) that are shown in Figure 2 are quite variable. DH270.6 was the best antibody within the DH270 lineage in terms of breadth and potency, but still it only neutralized 55% of M group viruses tested. Moreover, DH270 bNAbs never acquired the very high potency that is observed for some of the other V3 glycan bNAbs.<sup>61</sup> We hypothesize that this may be in part because the CH848 viruses carry some non-consensus rare amino acids near the contact surface of DH270 antibodies. It is possible that, in spite of the CH848 viral quasispecies selecting for antibodies that are able to interact with many heterologous Envs, the rare CH848-specific Env substitutions may have a negative impact, limiting affinity and breadth of DH270 antibodies at the population level. The enhanced coverage of heterologous viral variation afforded by augmenting lineage-based design with a mosaic boost is

summarized in Figure 6. While a pentavalent CH848/DH270 lineage-based design does a better job of covering V3 glycan epitope amino acid diversity than a single M group consensus, the lineage-based designed complemented by a mosaic boost would do better still.

A similar story is found for the CH505 Env quasiespecies which has its own set of quirks, ie, unusual amino acids in the CD4 binding region (Figure 5). Again a mosaic boost might shift vaccine-elicited antibody recognition toward improved heterologous breadth and potency. Because mosaics by definition carry common and complementary variants, we propose that they could be used generally to augment any lineage-based design, by including the more common forms of an epitope that are missing from the viral quasiespecies isolated from a single individual. This kind of mosaic coverage of the most common variants would be throughout Env, and analogous to the V3 glycan example shown here, could also potentially support breadth in any antibody target region.

## 7 | CONCLUSIONS AND FUTURE PERSPECTIVES

HIV/AIDS has become a manageable disease with respect to therapy, but the need to curb viral transmission through a vaccine remains a high and urgent priority. Fortunately, significant advances have been made in recent years, giving ample reason for hope that the challenges are not insurmountable. Overcoming viral diversity, the primary focus of this review, is a major issue, but many additional challenges will need to be addressed before a vaccine can be successful. For example, the bNAbs that arise during natural HIV-1 infections tend to have unusual features likely to pose challenges for vaccination, such as high levels of somatic mutation, autoreactivity with self proteins, and long complementarity determining regions.<sup>21</sup> Progress has been made on many fronts that may help us deal with these and other unresolved issues. For antibody induction, advances have been made in terms of better understanding of the structure of the native Env trimer and the role of the glycan shield in antibody interactions.<sup>60</sup> More native-like forms of the Env trimer for vaccine delivery are being explored.<sup>85,86</sup> Novel safe ways to modulate immune tolerance and enable vaccine induction of polyreactive and autoreactive antibodies are also being explored.<sup>87</sup> For cytotoxic T lymphocyte (CTL) focused vaccines, new delivery approaches are being used,<sup>33,88</sup> and preclinical testing of conserved region polyvalent mosaics vaccines is underway.<sup>51</sup> A desirable long-term goal may be to merge parallel B-cell and T-cell focused vaccine strategies into an immunological "one-two punch". This combined approach would incorporate vaccine elements that enable elicitation of antibodies that effectively block infection, coupled with elements that elicit favorable T-cell responses to provide immune-mediated control of breakthrough infections. Inclusion of diversity considerations, as outlined here, may contribute in important ways to this success.

### ACKNOWLEDGEMENTS

This work was funded by UM-1-AI100645 from NIH, NIAID, Division of AIDS, Duke Center for HIV/AIDS Vaccine Immunology-Immunogen

Discovery (CHAVI-ID). We thank Will Fischer for helpful suggestions on clarity and wording.

### CONFLICT OF INTEREST

The authors have no conflicts of interest.

### REFERENCES

1. Gunthard HF, Aberg JA, Eron JJ, et al. Antiretroviral treatment of adult HIV infection: 2014 recommendations of the International Antiviral Society-USA Panel. *JAMA*. 2014;312:410-425.
2. Abdul-Jawad S, Ondondo B, van Hateren A, et al. Increased valency of conserved-mosaic vaccines enhances the breadth and depth of epitope recognition. *Mol Ther*. 2016;24:375-384.
3. Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral therapy for the prevention of HIV-1 transmission. *N Engl J Med*. 2016;375:830-839.
4. World Health Organization. *Global AIDS Update*. Geneva: World Health Organization; 2016. Available at: [www.who.int/hiv/pub/arv/global-aids-update-2016-pub/en/](http://www.who.int/hiv/pub/arv/global-aids-update-2016-pub/en/). Accessed June, 2016.
5. Bates M, Mudenda V, Shibemba A, et al. Burden of tuberculosis at post mortem in inpatients at a tertiary referral centre in sub-Saharan Africa: a prospective descriptive autopsy study. *Lancet Infect Dis*. 2015;15:544-551.
6. Huson MA, Kalkman R, Hoogendijk AJ, et al. Impact of HIV infection on the haemostatic response during sepsis and malaria. *Br J Haematol*. 2016;173:918-926.
7. Mansky LM, Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol*. 1995;69:5087-5094.
8. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*. 1996;271:1582-1586.
9. Bar KJ, Tsao CY, Iyer SS, et al. Early low-titer neutralizing antibodies impede HIV-1 replication and select for virus escape. *PLoS Pathog*. 2012;8:e1002721.
10. Fischer W, Ganusov VV, Giorgi EE, et al. Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS ONE*. 2010;5:e12303.
11. Goonetilleke N, Liu MK, Salazar-Gonzalez JF, et al. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *J Exp Med*. 2009;206:1253-1272.
12. Liao HX, Lynch R, Zhou T, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature*. 2013;496:469-476.
13. Liu MK, Hawkins N, Ritchie AJ, et al. Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. *J Clin Invest*. 2013;123:380-393.
14. Gao F, Bonsignori M, Liao HX, et al. Cooperation of B cell lineages in induction of HIV-1-broadly neutralizing antibodies. *Cell*. 2014;158:481-491.
15. Brockman MA, Brumme ZL, Brumme CJ, et al. Early selection in Gag by protective HLA alleles contributes to reduced HIV-1 replication capacity that may be largely compensated for in chronic infection. *J Virol*. 2010;84:11937-11949.
16. Martinez-Picado J, Deeks SG. Persistent HIV-1 replication during antiretroviral therapy. *Curr Opin HIV AIDS*. 2016;11:417-423.
17. Murray AJ, Kwon KJ, Farber DL, Siliciano RF. The latent reservoir for HIV-1: how immunologic memory and clonal expansion contribute to HIV-1 persistence. *J Immunol*. 2016;197:407-417.
18. Gaschen B, Taylor J, Yusim K, et al. Diversity considerations in HIV-1 vaccine selection. *Science*. 2002;296:2354-2360.

19. Korber BT, Letvin NL, Haynes BF. T-cell vaccine strategies for human immunodeficiency virus, the virus with a thousand faces. *J Virol*. 2009;83:8300–8314.
20. Fischer W, Perkins S, Theiler J, et al. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nat Med*. 2007;13:100–106.
21. Theiler J, Yoon H, Yusim K, Picker LJ, Frueh K, Korber B. Epigraph: a vaccine design tool applied to an HIV therapeutic vaccine and a pan-filovirus vaccine. *Sci Rep*. 2016;6:33987.
22. Haynes BF, Moody MA, Alam M, et al. Progress in HIV-1 vaccine development. *J Allergy Clin Immunol*. 2014;134:3–10; quiz 11.
23. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science*. 2000;288:55–56.
24. Zhang M, Foley B, Schultz AK, et al. The role of recombination in the emergence of a complex and dynamic HIV epidemic. *Retrovirology*. 2010;7:25.
25. Wang S, Pal R, Mascola JR, et al. Polyvalent HIV-1 Env vaccine formulations delivered by the DNA priming plus protein boosting approach are effective in generating neutralizing antibodies against primary human immunodeficiency virus type 1 isolates from subtypes A, B, C, D and E. *Virology*. 2006;350:34–47.
26. Hammer SM, Sobieszczyk ME, Janes H, et al. Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. *N Engl J Med*. 2013;369:2083–2092.
27. Crooks ET, Tong T, Chakrabarti B, et al. Vaccine-elicited tier 2 HIV-1 neutralizing antibodies bind to quaternary epitopes involving glycan-deficient patches proximal to the CD4 binding site. *PLoS Pathog*. 2015;11:e1004932.
28. McCoy LE, van Gils MJ, Ozorowski G, et al. Holes in the glycan shield of the native HIV envelope are a target of trimer-elicited neutralizing antibodies. *Cell Rep*. 2016;16:2327–2338.
29. Barouch DH, Stephenson KE, Borducchi EN, et al. Protective efficacy of a global HIV-1 mosaic vaccine against heterologous SHIV challenges in rhesus monkeys. *Cell*. 2013;155:531–539.
30. Buchbinder SP, Mehrotra DV, Duerr A, et al. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet*. 2008;372:1881–1893.
31. Gray GE, Moodie Z, Metch B, et al. Recombinant adenovirus type 5 HIV gag/pol/nef vaccine in South Africa: unblinded, long-term follow-up of the phase 2b HVTN 503/Phambili study. *Lancet Infect Dis*. 2014;14:388–396.
32. Hansen SG, Ford JC, Lewis MS, et al. Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature*. 2011;473:523–527.
33. Hansen SG, Sacha JB, Hughes CM, et al. Cytomegalovirus vectors violate CD8<sup>+</sup> T cell epitope recognition paradigms. *Science*. 2013;340:1237874.
34. Hansen SG, Wu HL, Burwitz BJ, et al. Broadly targeted CD8(+) T cell responses restricted by major histocompatibility complex E. *Science*. 2016;351:714–720.
35. Barouch DH, Korber B. HIV-1 vaccine development after STEP. *Annu Rev Med*. 2010;61:153–167.
36. Hulot SL, Korber B, Giorgi EE, et al. Comparison of immunogenicity in rhesus macaques of transmitted-founder, HIV-1 group M consensus, and trivalent mosaic envelope vaccines formulated as a DNA prime, NYVAC, and envelope protein boost. *J Virol*. 2015;89:6462–6480.
37. Santra S, Liao HX, Zhang R, et al. Mosaic vaccines elicit CD8<sup>+</sup> T lymphocyte responses that confer enhanced immune coverage of diverse HIV strains in monkeys. *Nat Med*. 2010;16:324–328.
38. Santra S, Muldoon M, Watson S, et al. Breadth of cellular and humoral immune responses elicited in rhesus monkeys by multi-valent mosaic and consensus immunogens. *Virology*. 2012;428:121–127.
39. Barouch DH, O'Brien KL, Simmons NL, et al. Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. *Nat Med*. 2010;16:319–323.
40. Li H, Wang S, Kong R, et al. Envelope residue 375 substitutions in simian-human immunodeficiency viruses enhance CD4 binding and replication in rhesus macaques. *Proc Natl Acad Sci USA*. 2016;113:E3413–3422.
41. Liu J, O'Brien KL, Lynch DM, et al. Immune control of an SIV challenge by a T-cell-based vaccine in rhesus monkeys. *Nature*. 2009;457:87–91.
42. Roederer M, Keele BF, Schmidt SD, et al. Immunological and virological mechanisms of vaccine-mediated protection against SIV and HIV. *Nature*. 2014;505:502–508.
43. Fischer W, Apetrei C, Santiago ML, et al. Distinct evolutionary pressures underlie diversity in simian immunodeficiency virus and human immunodeficiency virus lineages. *J Virol*. 2012;86:13217–13231.
44. Yusim K, Dilan R, Borducchi E, et al. Hepatitis C genotype 1 mosaic vaccines are immunogenic in mice and induce stronger T-cell responses than natural strains. *Clin Vaccine Immunol*. 2013;20:302–305.
45. Fenimore PW, Muhammad MA, Fischer WM, et al. Designing and testing broadly-protective filoviral vaccines optimized for cytotoxic T-lymphocyte epitope coverage. *PLoS ONE*. 2012;7:e44769.
46. Kamlangdee A, Kingstad-Bakke B, Osorio JE. Mosaic H5 hemagglutinin provides broad humoral and cellular immune responses against influenza viruses. *J Virol*. 2016;90:6771–6783.
47. Badamchi-Zadeh A, McKay PF, Korber BT, et al. A multi-component prime-boost vaccination regimen with a consensus MOMP antigen enhances chlamydia trachomatis clearance. *Front Immunol*. 2016;7:162.
48. Williamson AL, Rybicki EP. Justification for the inclusion of Gag in HIV vaccine candidates. *Expert Rev Vaccines*. 2016;15:585–598.
49. Kulkarni V, Valentin A, Rosati M, et al. HIV-1 conserved elements p24CE DNA vaccine induces humoral immune responses with broad epitope recognition in macaques. *PLoS ONE*. 2014;9:e111085.
50. Letourneau S, Im EJ, Mashishi T, et al. Design and pre-clinical evaluation of a universal HIV-1 vaccine. *PLoS ONE*. 2007;2:e984.
51. Ondondo B, Murakoshi H, Clutton G, et al. Novel conserved-region T-cell mosaic vaccine with high global HIV-1 coverage is recognized by protective responses in untreated infection. *Mol Ther*. 2016;24:832–842.
52. Stephenson KE, SanMiguel A, Simmons NL, et al. Full-length HIV-1 immunogens induce greater magnitude and comparable breadth of T lymphocyte responses to conserved HIV-1 regions compared with conserved-region-only HIV-1 immunogens in rhesus monkeys. *J Virol*. 2012;86:11434–11440.
53. Ferguson AL, Mann JK, Omarjee S, Ndung'u T, Walker BD, Chakraborty AK. Translating HIV sequences into quantitative fitness landscapes predicts viral vulnerabilities for rational immunogen design. *Immunity*. 2013;38:606–617.
54. Mothe B, Llano A, Ibarrondo J, et al. Definition of the viral targets of protective HIV-1-specific T cell responses. *J Transl Med*. 2011;9:208.
55. Gorse GJ, Baden LR, Wecker M, et al. Safety and immunogenicity of cytotoxic T-lymphocyte poly-epitope, DNA plasmid (EP HIV-1090) vaccine in healthy, human immunodeficiency virus type 1 (HIV-1)-uninfected adults. *Vaccine*. 2008;26:215–223.
56. Jaoko W, Nakwagala FN, Anzala O, et al. Safety and immunogenicity of recombinant low-dosage HIV-1 A vaccine candidates vectored by plasmid pTHr DNA or modified vaccinia virus Ankara (MVA) in humans in East Africa. *Vaccine*. 2008;26:2788–2795.
57. Hayton EJ, Rose A, Ibrahimsa U, et al. Safety and tolerability of conserved region vaccines vectored by plasmid DNA, simian adenovirus and modified vaccinia virus ankara administered to human immunodeficiency virus type 1-uninfected adults in a randomized, single-blind phase I trial. *PLoS ONE*. 2014;9:e101591.
58. Thurmond J, Yoon H, Kuiken C, et al. Web-based design and evaluation of T-cell vaccine candidates. *Bioinformatics*. 2008;24:1639–1640.

59. Kelsoe G, Verkoczy L, Haynes BF. Immune system regulation in the induction of broadly neutralizing HIV-1 antibodies. *Vaccines*. 2014;2:1–14.
60. Stewart-Jones GB, Soto C, Lemmin T, et al. Trimeric HIV-1-Env structures define glycan shields from clades A, B, and G. *Cell*. 2016;165:813–826.
61. Wagh K, Bhattacharya T, Williamson C, et al. Optimal combinations of broadly neutralizing antibodies for prevention and treatment of HIV-1 clade C infection. *PLoS Pathog*. 2016;12:e1005520.
62. Gorman J, Soto C, Yang MM, et al. Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design. *Nat Struct Mol Biol*. 2016;23:81–90.
63. Kong L, Lee JH, Doores KJ, et al. Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat Struct Mol Biol*. 2013;20:796–803.
64. Burton DR, Mascola JR. Antibody responses to envelope glycoproteins in HIV-1 infection. *Nat Immunol*. 2015;16:571–576.
65. Wu X, Zhou T, Zhu J, et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science*. 2011;333:1593–1602.
66. Zhou T, Georgiev I, Wu X, et al. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science*. 2010;329:811–817.
67. Fera D, Schmidt AG, Haynes BF, et al. Affinity maturation in an HIV broadly neutralizing B-cell lineage through reorientation of variable domains. *Proc Natl Acad Sci USA*. 2014;111:10275–10280.
68. Bonsignori M, Zhou T, Sheng Z, et al. Maturation pathway from germline to broad HIV-1 neutralizer of a CD4-mimic antibody. *Cell*. 2016;165:449–463.
69. Doria-Rose NA, Bhiman JN, Roark RS, et al. New member of the V1V2-directed CAP256-VRC26 lineage that shows increased breadth and exceptional potency. *J Virol*. 2016;90:76–91.
70. Bhiman JN, Anthony C, Doria-Rose NA, et al. Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. *Nat Med*. 2015;21:1332–1336.
71. Jia M, Lu H, Markowitz M, Cheng-Mayer C, Wu X. Development of broadly neutralizing antibodies and their mapping by monomeric gp120 in human immunodeficiency virus type 1-infected humans and simian-human immunodeficiency virus SHIVSF162P3N-infected macaques. *J Virol*. 2016;90:4017–4031.
72. Hessel AJ, Malherbe DC, Pissani F, et al. Achieving potent autologous neutralizing antibody responses against tier 2 HIV-1 viruses by strategic selection of envelope immunogens. *J Immunol*. 2016;196:3064–3078.
73. Zhou T, Lynch RM, Chen L, et al. Structural repertoire of HIV-1 neutralizing antibodies targeting the CD4 supersite in 14 donors. *Cell*. 2015;161:1280–1292.
74. Hraber P, Korber B, Wagh K, et al. Longitudinal antigenic sequences and sites from intra-host evolution (LASSIE) identifies immune-selected HIV variants. *Viruses*. 2015;7:5443–5475.
75. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med*. 2009;361:2209–2220.
76. Haynes BF, Gilbert PB, McElrath MJ, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N Engl J Med*. 2012;366:1275–1286.
77. Santra S, Tomaras GD, Warrier R, et al. Human non-neutralizing HIV-1 envelope monoclonal antibodies limit the number of founder viruses during SHIV mucosal infection in rhesus macaques. *PLoS Pathog*. 2015;11:e1005042.
78. Jechlinger W. Optimization and delivery of plasmid DNA for vaccination. *Expert Rev Vaccines*. 2006;5:803–825.
79. Okuda K, Wada Y, Shimada M. Recent developments in preclinical DNA vaccination. *Vaccines*. 2014;2:89–106.
80. Rinaldi M, Fioretti D, Lurescia S, eds. *DNA Vaccines: Methods and Protocols*, 3rd edn. New York: Springer; 2014.
81. Gregory AE, Titball R, Williamson D. Vaccine delivery using nanoparticles. *Front Cell Infect Microbiol*. 2013;3:13.
82. Peek LJ, Middaugh CR, Berkland C. Nanotechnology in vaccine delivery. *Adv Drug Deliv Rev*. 2008;60:915–928.
83. Rosenthal JA, Chen L, Baker JL, Putnam D, DeLisa MP. Pathogen-like particles: biomimetic vaccine carriers engineered at the nanoscale. *Curr Opin Biotechnol*. 2014;28:51–58.
84. Zhao L, Seth A, Wibowo N, et al. Nanoparticle vaccines. *Vaccine*. 2014;32:327–337.
85. Chen J, Kovacs JM, Peng H, et al. HIV-1 ENVELOPE. Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. *Science*. 2015;349:191–195.
86. Sanders RW, Derking R, Cupo A, et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog*. 2013;9:e1003618.
87. Liu M, Yang G, Wiehe K, et al. Polyreactivity and autoreactivity among HIV-1 antibodies. *J Virol*. 2015;89:784–798.
88. Roshorm Y, Cottingham MG, Potash MJ, Volsky DJ, Hanke T. T cells induced by recombinant chimpanzee adenovirus alone and in prime-boost regimens decrease chimeric EcoHIV/NDK challenge virus load. *Eur J Immunol*. 2012;42:3243–3255.
89. Kong L, Torrents de la Pena A, Deller MC, et al. Complete epitopes for vaccine design derived from a crystal structure of the broadly neutralizing antibodies PGT128 and 8ANC195 in complex with an HIV-1 Env trimer. *Acta Crystallogr D Biol Crystallogr*. 2015;71:2099–2108.
90. Huang J, Ofek G, Laub L, et al. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature*. 2012;491:406–412.