HIV-1 adaptation to HLA: a window into virus-host immune interactions

Jonathan M. Carlson[1], Anh Q. Le[2], Aniqa Shahid[2], and Zabrina L. Brumme[2,3]

1. Microsoft Research, Redmond, WA, USA
2. Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada
3. British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada

Keywords

Human Immunodeficiency Virus 1 (HIV-1), Human Leukocyte Antigen (HLA) Class I, CD8+ T-lymphocyte (CTL), evolution, adaptation, immune escape

Abstract

HIV-1 develops specific mutations within its genome that allow it to escape detection by human leukocyte antigen (HLA) class I-restricted immune responses, notably those of CD8+ cytotoxic T lymphocytes (CTL). HLA thus represents a major force driving the evolution and diversification of HIV-1 within individuals and at the population level. Importantly, the study of HIV-1 adaptation to HLA also represents an opportunity to identify what qualities constitute an effective immune response, how the virus in turn adapts to these pressures, and how we may harness this information to design HIV-1 vaccines that stimulate effective cellular immunity.

CTL escape as a window into HIV-1 and host cellular immune interactions

Originating from a single zoonotic transmission event approximately a century ago [1], the HIV-1 group M pandemic strain has subsequently diversified into nine subtypes and dozens of circulating recombinant forms that differ by up to 30% in their envelope amino acid sequence [2]. HIV-1's high mutation rate and tolerance for genetic diversity represent central challenges to vaccine design. These features allow HIV-1 to rapidly and effectively adapt to changes in the environment – in particular, to evolutionary selection pressures imposed by the host immune response. On the other hand, HIV-1's extraordinary adaptability represents an opportunity to study host–viral interactions at an uncommon level of detail. Because host genetics, and thus immune responses, vary substantially among individuals, the 'optimal' HIV-1 sequence also differs among individuals. Because the immune response is itself adaptive, the optimal HIV-1 sequence within an individual also differs over time. Thus, studying immune-driven HIV-1 evolution provides a window into host immunity as well as the pathways and constraints governing viral adaptation – information that can in turn be translated into testable hypotheses and interventions.

This review focuses on mutational HIV-1 adaptation to human leukocyte antigen (HLA) class I-restricted immune responses, primarily those restricted by CD8+ T-lymphocytes (CTL). We begin by outlining historical discoveries that shaped our understanding of the specificity and reproducibility of HIV-1 immune escape. We then move on to discuss the contribution of HLA-driven escape to the evolution and diversification of HIV-1 in individuals and populations. Finally, we highlight how the study of HIV adaptation to HLA yields insights into viral/host cellular immune interactions and their clinical consequences, with the ultimate goal of translating these into vaccine concepts. The genetic origins and global diversification of HIV-1 are not covered in-depth, however we refer the reader to two excellent articles on the subject [1, 2].

HIV-1 immune escape is predictable based on host HLA

HLA class I-restricted CTL play a major role in HIV-1 control (see Glossary) via their recognition of short, virus-derived peptide epitopes presented on the infected cell surface by the highly polymorphic HLA-A, B, and C molecules (**Figure 1A**). HIV-1-specific CTL appear coincident with the dramatic decline in viremia during the acute phase of infection [3, 4] and mediate the subsequent control of viremia to setpoint levels (see Glossary) [5]. Experimental depletion of CD8+ lymphocytes in rhesus macaques renders them unable to control simian immunodeficiency virus (SIV) infection (*e.g.* [6]), whereas in humans, strong epidemiological links between carriage of specific protective HLA alleles – notably HLA-B*57 and HLA-B*27 – and slower HIV disease progression are well-established (*e.g.* [7]). Independent effects of HLA-C expression level on HIV-1 control have also been demonstrated [8]. Furthermore, the targeting of specific epitopes (typically conserved ones) or proteins (primarily $p24^{Gag}$ and to a lesser extent Pol) has been linked to control [9-11]. However, perhaps the clearest demonstration that HLA-restricted CTL exert pressure on HIV-1 *in vivo* is the virus' propensity to escape this pressure via mutation.

CTL escape in HIV-1 was first described in the early 1990s, when researchers observed that "accumulation of…mutations in T-cell antigenic targets…provides a mechanism for immune escape" [12]. Researchers also deduced the HLA-restricted nature of this phenomenon, noting

that "different HLA class I molecules select distinct HIV-derived epitopes to stimulate CTL responses…[therefore]…HLA type could have an effect on virus escape" [12]. Indeed, despite extraordinary host and viral genetic variation, we now appreciate that HIV-1 immune escape is highly predictable based on host HLA. For example, three-quarters of HLA-B*57-expressing HIV-1 subtype B infected persons will select the Gag-T242N substitution (at position three of the Gag-TW10 epitope; see Glossary) within three months of infection [13, 14], while fifty percent will subsequently select G248A at epitope position 9 [14-16], thereby conferring complete escape from B*57-restricted, TW10-specific CTL [13].

CTL escape mutations fall into three categories (**Figure 1B-D**). First, escape can act upon processes upstream of epitope–HLA binding, by introducing mutations that interfere with intracellular epitope processing [17] (**Figure 1B**). An example of an "antigen processing escape mutation" is the B*57:03-restricted Gag-A146P substitution, selected at the residue immediately upstream of Gag-IW9, which prevents N-terminal aminopeptidase-mediated trimming of the epitope [18]. Antigen processing mutations can also occur within [19] or distal to [20] the epitope. Second, escape can occur via mutations that compromise epitope-HLA binding, which typically occur at anchor sites (generally epitope positions two and/or C-terminus) (**Figure 1C**). Of all the HLA-associated polymorphisms that occur within or near optimally-described CTL epitopes, roughly 20% occur at anchor positions, a frequency that is approximately twofold higher than expected by chance [16, 21]. Moreover, these are estimated to confer an average tenfold reduction in predicted epitope-HLA binding affinity [16]. An example is the B*27-restricted Gag-R264K mutation, selected at position two of the B*27-restricted Gag-KK10 epitope, which abrogates epitope-HLA binding [22]. Third, escape can act upon downstream processes: certain mutations do not alter epitope processing or presentation, but reduce or abrogate recognition of the epitope–HLA complex by some or all of the T-cell receptors (TCR) expressed by the original selecting CTL(s) (**Figure 1D**). "TCR escape mutations" typically occur at central epitope positions. For example, the B*27-associated Gag-L268M substitution [12], selected at position 6 of Gag-KK10, retains B*27-binding ability but abrogates epitope recognition by many B*27-restricted TCR clonotypes (see Glossary) in the repertoire [23]. Of note, the above categories are not mutually exclusive – for example, anchor residue mutations may, via alterations in peptide secondary structure, affect both HLA binding and TCR recognition [24].

Statistical association studies: a powerful tool to identify HLA-associated polymorphisms in HIV-1

The discovery that CTL escape mutations were predictable based on host HLA was key to the next major advancement: the identification of HLA-associated polymorphisms in HIV-1 (see Glossary) by statistical association. These models are predicated on the idea that, if HIV-1 can readily adapt to changing selection pressures, and if the HLA-restricted CTL response is a major source of such pressures, then fitting statistical evolutionary models to large datasets of linked HLA and HIV-1 genotypes will yield insights into the sites, mutational pathways, and characteristics of HLA-mediated escape. Specifically, such analyses identify HIV-1 amino acids that are significantly over-represented among persons expressing a given HLA allele (putative escape mutations) and those that are under-represented in such persons (putative immunologically susceptible forms for the given HLA) [16, 25, 26].

The first such study identified nearly 100 HLA-associated polymorphisms in HIV-1 reverse transcriptase in a cohort of ~400 patients, suggesting that the impact of HLA on HIV-1 diversity was extensive and widespread [25]. Although the methods used in this foundational study came under some criticism [27], its broad conclusions have been confirmed by newer methods. Specifically, because HIV-1 sequences are related to one another through common descent, it is now appreciated that standard tests of association that assume independence of observations should not be applied [27]. To address this, statistical association studies now employ phylogenetic techniques to reconstruct all possible sequences (and their probabilities) at the immediate interior nodes of the tree, thereby estimating the transmitted HIV-1 sequence (**Figure 2**). This allows researchers to test the null hypothesis that observed HIV-1 polymorphisms (at the tree tips) are explained by neutral evolution along the tree, against the alternative hypothesis that these are better explained by HLA-driven pressures within the current host [28]. Such approaches are termed "phylogenetically-corrected" (see Glossary). Moreover, these analyses test for pairwise relationships between every HLA allele and HIV-1 polymorphism in the dataset, while correcting for various host and viral genetic confounders [26, 28, 29] – as such, they feature stringent corrections for multiple hypothesis testing.

Statistical association studies, largely undertaken on HIV-1 subtype-specific datasets, have yielded "immune escape maps" detailing the locations and pathways of HLA-driven escape throughout the HIV-1 proteome. These escape maps are most detailed for subtype B (*e.g.* [16]) followed by C (*e.g.* [30]) and CRF01 (AE) [31]; other subtypes remain somewhat understudied in this context.

HIV-1 adaptation to HLA: population-level analyses reveal insights into escape mechanisms and pathways

Studying HIV adaptation to HLA via statistical association has yielded valuable insights into the nature of host cellular immune responses and the virus's ability to evade these via mutation. We now review some of these insights.

Firstly, as mentioned, CTL escape is highly reproducible in context of host HLA. For example, the HLA-A*24:02-associated Y135F substitution in the HIV-1 accessory protein Nef (**Figure 3A**) is selected in over 80% of HIV-1 subtype B-infected, HLA-A*24:02-expressing persons [16]. CTL escape is also exquisitely HLA-specific. Even though many closely-related HLA alleles bind the same (or similar) HIV-1 epitopes, over 60% of HLA-associated polymorphisms are nevertheless best defined at the HLA subtype (4-digit) level (see Glossary), whereas fewer than 10% are shared across HLA supertypes [16]. For example, HLA-B*57:02, B*57:03 and B*58:01 all present the Gag-TW10 epitope [32], but drive different escape pathways within it [33]. This exquisite specificity also implies that, if an HLA-associated polymorphism is identified in HIV-1, then the restricting HLA allele must be exerting selection pressure on the virus at (or near) this location. In fact, the identification of HLA-associated polymorphisms has helped guide the discovery of novel CTL epitopes in HIV-1 (*e.g.* [34]), including those in 'cryptic' (alternative, including antisense) reading frames [35, 36].

CTL escape can also occur via multiple pathways. For example, B*08-driven escape at Nef codon 94 can occur via selection of K94E, M, N or Q [26, 37] (**Figure 3B**). In other cases, multiple HLA alleles put pressure on the same HIV-1 codon, sometimes in opposing directions. For example at Gag codon 147, escape from HLA-B*14:02 and B*15:01-mediated pressure occurs via selection of the consensus I, whereas for HLA-A*25:01, B*13:02 and B*57:01, this occurs via selection of L [16] (**Figure 3C**). Escape pathways can also be HIV-1 subtype-dependent; *e.g.* B*57:03-driven escape in Gag-KF11 differs in subtypes B and C [38].

HIV-1 adaptation to HLA: individual-level analyses reveal insights into escape dynamics

Studying immune-driven viral evolution at the individual level has revealed important insight into immune escape dynamics. It is now understood that a severe genetic bottleneck occurs at HIV-1 transmission such that 80% of heterosexual infections are initiated by a single founder virus [39, 40]. Following transmission, this virus undergoes rapid population growth and star-like diversification (representing random mutation of founder and descendant sequences), yielding a swarm of related variants that become the evolutionary substrate upon which the host's immune responses exert pressure [41]. CTL escape begins rapidly thereafter [42, 43]. The first escape mutations appear during acute-phase viremia decline [5], with the most rapid of these attaining fixation shortly thereafter [5, 44] (as early as 21 days post-infection in humans [5, 41] and 17 days in SIV-infected macaques [45]).

Surprisingly, the conceptually straightforward pathway whereby the initially selected escape mutation directly outcompetes its founder is unlikely to represent the norm [5]. More commonly, initial variants appear as a diverse pool of low-frequency TCR escape mutants that retain some ability to be targeted by existing (or *de novo*) CTL [44, 46]. The host CTL response subsequently drives the selection of a second wave of more effective escape variants, often at HLA-anchor residues, that ultimately outcompete their predecessors [44, 47]. For example, in a B*57:03-expressing individual, initial escape within Gag-TW10 occurred via a minority G248E TCR escape mutation at epitope position 9 that modestly reduced CTL recognition, which later gave way to the canonical G248A (along with T242N and V247I substitutions at positions 3 and 8) [46]. These observations are consistent with escape being driven by random mutation, such that a majority of viable mutations are likely to confer only a weak advantage to the virus. These are eventually outcompeted by more effective escape mutations when the latter happen to arise.

Despite the initial star-like shape of intra-host phylogenies, subsequent CTL-driven selective sweeps change the intra-host consensus sequence, yielding more ladder-like phylogenies [48]. One study conservatively estimated that a minimum of 30% of intra-host consensus changes in Gag/Pol and 60% in Nef were attributable to HLA pressures [14]. Another indicated that throughout the HIV-1 proteome, between 9 and 18 substitutions attain fixation in the first 6 months of infection, the majority likely attributable to CTL [41]. The rapid and dramatic impact of HLA on within-host HIV evolution following transmission has also been demonstrated via cohort-level analyses: for 15-20% of known HLA-associated polymorphisms, the statistical strength of association linking the HLA with the selected HIV-1 polymorphism is already near-maximal at ~3 months post-infection [49], supporting very rapid and reproducible escape at a substantial number of viral sites.

Escape studies reveal correlates of protective immunity

Because CTL escape mutations mark viral sites under reproducible *in vivo* selection by a particular HLA allele, analysis of these sites can help reveal why certain HLA alleles are more effective at controlling HIV-1 than others. In particular, analysis of HIV adaptation pathways supports CTL response breadth as beneficial: protective alleles exert effective immune pressure on a larger number of sites across HIV-1 (defined here as "pressure strong enough to compel the virus to reproducibly escape *in vivo*" as detected by an HLA-association at that location) [16]. The statistical strength and location of immune pressure is also key. Protective HLA alleles tend to exert strong selection pressure on mutationally constrained sites, particularly in Gag and to a lesser extent Pol [16, 50], consistent with the observation that these alleles tend to drive fitness-costly mutations ([51, 52], and following section).

Protective HLA alleles also tend to select escape mutations at HLA anchor residues more often than nonprotective ones [16]. This may be because CTL responses restricted by protective HLA alleles tend to recognize more epitope variants, perhaps due to increased clonal diversity and/or cross-reactivity of the TCR repertoire [23, 53], thus necessitating the selection of escape mutations that dramatically (or completely) abrogate epitope-HLA binding. For example, initial escape within the B*27-restricted KK10 epitope proceeds via selection of L268M which yields incomplete escape from certain TCR clonotypes present in the original selecting CTL pool [23]. In at least some B*27 patients, *de novo* and/or cross-reactive CTL clonotypes continue to respond to the L268M-containing KK10 epitope; indeed, enrichment of these TCR clonotypes in B*27 controllers versus progressors suggest they are critical to sustained B*27-mediated viremia control in these persons [23, 53]. Eventually, anchor residue escape via R264K, which is estimated to reduce epitope-HLA binding by 10-100 fold [54], confers essentially complete escape *in vivo*. But, due to the requirement that the upstream stabilizing S173A mutation be present prior to R264K selection [55], escape via this mutation may not occur until years later, possibly explaining the prolonged HIV-1 control often observed in B*27+ persons [22].

Fitness consequences of escape: an explicit link to vaccine design?

Fitness consequences of escape constitute another important consideration. Although escape confers a net advantage to HIV-1 in context of an effective immune response, many of these mutations weaken the virus, yielding a net disadvantage in the absence of immune pressure. Early evidence supporting widespread fitness impacts of escape came from *in vivo* reversion data: upon transmission to a host lacking the restricting HLA, certain escape mutations selected in the previous host will revert to the original (usually subtype consensus) residue [13, 14, 44, 56-58]. In a recent study, roughly one third of all tested Gag escape mutations significantly reduced *in vivo* replication when engineered into the HIV-1 reference strain NL4-3 ($HIV\text{-}1_{NL4\text{-}3}$), highlighting the frequency with which HIV-1 tolerates functional reductions in exchange for immune evasion, and emphasizing the potency of CTL responses as evolutionary selection pressures [52]. While Gag harbors many fitness-costly escape sites (*e.g.* [55, 59-62]), costly mutations have also been reported in Pol [63] and Nef [64], and it is likely that they also occur in other proteins that have received less attention.

Fitness costs can be fully or partially rescued by compensatory mutations – a fact that complicates their quantification and implications. For example, the B*27-associated R264K substitution (selected at position 2 in the Gag-KK10 epitope) essentially abolishes *in vitro* viral replication when engineered into HIV-1_{NL4-3} [55], but it never occurs alone *in vivo.* It is most frequently accompanied by the upstream S173A substitution that restores viral replication to near wild-type levels [55]. Gag-R264K is an extreme example, however. More often, fitness costs are subtle and/or require the accumulation of multiple escape mutations to be detectable. For example, the B*57-driven Gag-T242N mutation reduces viral fitness only modestly [62, 65], but dose-dependent replicative reductions are observed when it is present alongside other B*57-driven $p24^{Gag}$ mutations [66], each of which can be rescued via compensation.

In the absence of adequate compensation, fitness-costly mutations, even if they allow complete escape from the original immune response, would be expected to offer a smaller net advantage to the virus than fitness-neutral mutations. In fact, various lines of evidence suggest that even minor fitness differences may incur residual clinical benefit to the host. For example, HIV-1 elite controllers—rare individuals able to suppress plasma viremia to below clinical detection limits without antiretroviral therapy—frequently harbor rare, fitness-costly escape mutations that reduce *in vitro* function of HIV-1 proteins [67-69]. Also, protective HLA alleles consistently drive more fitness-costly mutations than non-protective ones [51, 52]. Biologically relevant implications of fitness costs are also supported by high estimates of heritability of clinical traits (notably plasma viral loads, see Glossary) among donor and recipient pairs [70-72]. Furthermore, low HIV-1 replication capacity is associated with reduced viral loads in early [30] and chronic [61, 73] infection, and reduced rates of CD4 decline [30, 74], while transmission of HLA-B associated Gag escape mutations to an HLA-mismatched host is linked to lower viral loads in the recipient [75, 76]. Finally, HIV-1 sequences that carry high predicted fitness costs are less likely to be transmitted (and predict lower set point viral loads upon acquisition) [77].

Escape studies yield insights into vaccine strategies

These observations have led to the idea that immune-mediated containment of HIV-1 replication to levels that slow disease progression might be achievable through the design of vaccines that focus CTL responses against viral regions where escape cannot occur (or, if this is not possible, where escape can only occur at a major fitness cost and/or requires extensive compensation, thereby prolonging control by delaying the time to escape) [78]. However, identifying epitopes where escape is costly remains a challenge (see Outstanding Questions Box). One strategy is to focus on conserved regions (e.g. [10, 79, 80]), on the assumption that escape mutations will be more costly therein. The observation that ‘protective’ HIV-1 peptides (those against which *in vitro* respondents exhibit lower median viral loads than non-respondents) are more conserved than average [10] supports effective CTL responses as those targeting viral regions where escape is costly and/or requires extensive compensation. Nevertheless, the correlation between sequence conservation and viral replicative capacity is imperfect [80]. On one hand, conservation could simply reflect a lack of selection pressure, implying that natural immune responses to those regions are absent or ineffective. On the other hand, HLA-driven escape confounds such generalizations, as strong selection can force escape despite fitness costs, causing constrained regions to exhibit sequence variability at the population level [81]. To this end, Pereyra *et al.* used a model of conservation derived from protein structures that is

independent of HLA-driven selection to link HIV control to the ability to target epitopes in structurally conserved regions, providing further evidence that limiting escape options is an attractive vaccine design strategy [11].

As most 'conserved' HIV-1 regions still harbor some level of diversity, polyvalent immunogens that capture key variation in an optimized manner (*e.g.* mosaic vaccines) offer another promising strategy [82]. Mosaics contain multiple whole- or partial- protein immunogens that maximally cover observed sequence diversity with the aim of increasing the cross-reactivity and/or diversity of the TCR repertoire, a goal that appears to be achievable, at least in nonhuman primate models [82, 83]. A complementary strategy would be to ensure that mosaic vaccines explicitly incorporate both susceptible and common HLA-associated escape forms, particularly within structurally conserved epitopes, with the goal of eliciting strong *in vivo* immune responses that would, upon infection, limit the virus' escape options to noncanonical, fitness-costly pathways. This approach is supported by the observation that increased sequence coverage by Gag-specific immune responses (defined as the overall proportion of HIV-1 variant sequences targeted by the Gag-specific TCR repertoire) is a correlate of control [84]. However, a major challenge to implementing such a strategy is learning how to stimulate effective immune responses against escaped variants. Indeed, the observations of rapid progression in B*57^{+} individuals who acquire B*57-specific escape variants [66] suggests that such mutations do not elicit useful responses in the context of natural infection. Whether this result generalizes to other alleles or is exclusively driven by escape mutations that reduce epitope presentation will be critical to the success of mosaic vaccine approaches.

Finally, it is worth emphasizing that a CTL-based vaccine will need to stimulate effective cellular responses across the range of HLA class I alleles expressed in a given host population. Though 'personalized' vaccines (involving the selection of immunogens specific for an individual's HLA profile) are theoretically possible, the high burden (and unequal distribution) of HIV-1 globally renders broad implementation of such strategies unlikely. Rather, approaches to tailor HIV-1 vaccines to specific host populations and/or geographic areas, for example by designing subtype-specific immunogens and/or polyvalent mixtures thereof [85] while maximizing epitope coverage over a broad range of HLA alleles, are being pursued (e.g. [10, 79, 80]. Moreover, although escape is typically HLA-specific, widespread HLA-epitope "promiscuity" (the tendency of one epitope to be targeted by many individuals who lack the canonical presenting HLA allele) has been observed [11, 86], suggesting that broad population-based coverage may be achievable with a limited number of epitopes.

Identifying other sources of immune-driven selection on HIV-1

The methods used to identify HLA-associated viral polymorphisms can be adapted to identify additional sources of immune selection on HIV-1. The discovery of putative escape from natural killer (NK) cell pressures illustrates this point. NK cells express various cell-surface inhibitory and activating killer cell immunoglobulin-like receptors (KIR) that interact with their HLA class I ligands in an allotype-specific manner. Engagement of activating and inhibitory KIR yields stimulatory and tolerance signals, respectively: when the former overcome the latter, NK effector functions are initiated. Though not HLA-restricted in the classical sense, the allotype-specific nature of KIR–HLA interactions means that these are modulated in part by HLA

polymorphism [87]. Similarly, though not antigen-specific in the classical sense, KIR receptor-HLA interactions are determined in part by the sequence of the bound peptide (*e.g.* [88]). As such, KIR-associated pressures could drive the selection of viral escape mutations that allow infected cells to evade NK-mediated killing, for example via HIV-1 polymorphisms that reduce recognition by activating KIR and/or that enhance recognition by inhibitory KIR.

To identify such polymorphisms, 91 linked KIR/HIV-1 sequences were analyzed via statistical association, yielding 22 candidate polymorphisms [89]. Two of these, within the HIV-1 accessory protein Vpu (71M/H), were overrepresented among KIR2DL2-expressing persons, in particular those homozygous for HLA-C group 1 alleles (the preferred KIR2DL2 ligand). Furthermore, these polymorphisms enhanced the ability of the inhibitory KIR2DL2 to bind HIV-infected cells; $KIR2DL2^+$ NK cells failed to become activated in the presence of polymorphism-containing HIV-1; and $CD4^+$ T-cells infected with polymorphism-containing HIV-1 were not inhibited by $KIR2DL2^+$ NK cells [89]. These findings suggest that *in vivo* immune pressure by an inhibitory KIR could drive escape mutations that enhance binding of inhibitory KIR to HIV-infected cells, allowing these to escape NK cell-mediated elimination (**Figure 4A**).

HLA-mediated HIV-1 adaptation has also been used as evidence to support the efficacy of vaccine-induced immunity. Ideally, vaccine-induced immune responses would protect an individual from infection upon subsequent exposure (**Figure 4B, top**), but sterilizing immunity can be difficult to achieve for genetically diverse pathogens. Instead, vaccine-induced immune responses could induce a partial barrier capable of blocking infection by viral strains similar to, but not antigenically divergent from, the vaccine strain. If so, this would reduce the probability that a sequence similar to the immunogen will establish productive infection (*acquisition sieve effect*; **Figure 4B, middle)**. Alternatively, a vaccine could focus the immune response on specific epitopes shared between the founder virus and the immunogen, driving rapid escape following infection (*postinfection sieve effect*; **Figure 4B, bottom**) [90]. Though mechanistically distinct, it is difficult to distinguish acquisition from postinfection sieve effects *in vivo*. For CTL-based vaccines, an acquisition sieve effect could theoretically act by excluding vaccine-similar HIV-1 sequences from establishing productive infection due to prompt killing of infected cells by vaccine-elicited CTL. This possibility is supported by recent evidence suggesting frequent non-productive infection of cells local to the exposure site [77] as well as longstanding observations of HIV-specific T-cell responses in highly exposed seronegative individuals [91]. However, in the absence of vaccine-induced protection, postinfection selection of escape mutations by vaccine-induced CTL may be a more likely scenario [92].

The identification of sieve effects in the failed CTL-based STEP vaccine trial [92] and antibody-mediated protection in the partially effective RV144 vaccine trial [93] supports the immunogenicity of these vaccines (even if overall vaccine effectiveness was modest or absent). Specifically, sieve effects demonstrate that vaccine-induced immune responses can reduce the relative transmissibility of some viral strains. Nevertheless, sieve effects are a challenge that will need to be overcome, as their implications are potentially profound. If *in vivo* vaccine-selected HIV-1 strains are somehow 'worse' than those in natural infection (*e.g.* if escape is more rapid [94] and/or widespread among vaccinees), average rates of disease progression could increase. Widespread vaccine-driven HIV-1 evolution could also cause shifts in HIV-1 strain distributions

at the population level. The observation that vaccine-induced CTL responses may target slightly different epitopes than those in natural infection [95] may further complicate this issue.

HLA as a major driver of global HIV-1 diversity

HLA-associated immune pressures represent potent historical and ongoing drivers of global HIV-1 diversification – indeed, a recent North American study revealed that HIV-1 sites under HLA selection have diversified to a greater extent than those not under HLA selection, supporting this concept [96]. As the pandemic continues, HIV adaptation to HLA could have profound implications. The following sections discuss the mechanisms underlying this adaptation and their potential immunologic consequences.

The specificity and reproducibility of immune escape means that HIV-1 genomes in an infected individual will harbor adaptations specific to their host's HLA alleles. By extension, HIV-1 sequences circulating in a host population will harbor adaptations to the HLA alleles expressed in that population. And, because HLA allele frequencies vary widely across populations (with some alleles all but unique to certain ethnicities), so will the HLA-associated escape pathways selected in these populations. For example, the observation that over 50% of HLA-associated polymorphisms identified by statistical association in Mexico [97], and nearly two-thirds of those identified in Japan [98], are not observed to any great extent in Canada/USA/Australia (even though all of these populations harbor HIV-1 subtype B), illustrates this phenomenon (**Figure 5**).

The specificity and reproducibility of CTL escape also means that an HLA allele's frequency in a given population will largely determine the prevalence of its associated escape mutation(s) in circulation. The B*51-associated I135X mutation in HIV-1 reverse transcriptase (RT; at the C-terminus of the B*51-TI8 epitope) provides an example. Among nine cohorts on five continents, HLA-B*51 and RT-135X prevalence exhibit a strong positive correlation [99]. Notably in Japan, where the combined prevalence of B*51 - and the related allele B*52 - exceeds 40%, the population consensus at RT codon 135 is not "I" as in other parts of the world, but rather 135T, the B*51/B*52-associated escape form [100]. This example clearly illustrates how HLA allele frequencies influence HIV-1 polymorphism frequencies in host populations.

If HLA pressures drive ongoing HIV-1 diversification, then it is likely that these pressures played a role in the initial formation of HIV-1 subtypes. Recently, a mechanism has been proposed to substantiate this. Specifically, antigen processing escape mutations that alter the presentation of numerous epitopes with various HLA restrictions may be more likely to spread in circulation, thereby creating an efficient means whereby HIV-1 can adapt to a large number of HLA alleles in a population [101]. If selection of such mutations occurs in a subtype-specific context (for example, via selection of an escape mutation next to a subtype-specific motif), then selection of such broad-spectrum antigen processing escape mutations could drive HIV-1 subtype diversification [101]. Subtype-specific immune escape pathways also provide additional insight into how HLA may drive HIV-1 diversification along subtype lineages. For example, Gag-T242N is commonly selected by B*57 in HIV-1 subtypes B, C and D but rarely in subtype A1 [102], presumably due to fitness constraints in the viral backbone.

To what extent are HIV-1 escape mutations spreading in circulation – and what are the consequences?

The spread of HIV-1 escape mutations in circulation has potentially profound immunologic consequences. This is because escape mutations, when transmitted, are likely to compromise host immunity if acquired by an individual expressing the relevant HLA. Indeed, the higher circulation prevalence of escape mutations restricted by common HLA alleles may help explain the observation that common alleles are not usually protective [7]. However, escape mutations restricted by common HLA alleles are not the only ones at risk of gradual spread. We must also consider what is happening over the long-term.

In essence, if every escape mutation reverted to consensus following transmission to an HLA-mismatched host, escape mutation frequencies would remain temporally stable (**Figure 6A)**. In this case, the risk of acquiring escaped HIV-1 would roughly depend on the frequency of the HLA in the population (as well as its selection frequency in HLA-expressing individuals). However, certain escape mutations persist upon transmission to HLA-unmatched hosts, due to lack of fitness costs or efficient compensation (*e.g.* [103]). This has led to the concern that certain escape mutations, even those restricted by protective HLA alleles, could spread [104-107] (**Figure 6B**). Gradual spread of a large number of HIV-1 immune escape mutations in circulation could in turn undermine host antiviral immunity through increased transmission of "pre-adapted" HIV-1 sequences as the pandemic progresses.

A recent comparison of historic (1979-1989) versus modern (post-2000) HIV-1 subtype B cohorts in North America sought to address this issue directly [96]. Over the past ~30 years in North America, population HIV-1 diversity increased approximately twofold, and the average background frequencies of CTL escape mutations have also doubled during this time. Despite this, immunological consequences for North America are unlikely to be imminent, due to the modest absolute magnitude of these increases. However, rates of escape mutation spread may be higher, and implications more profound, in populations with high HIV-1 prevalence, different transmission dynamics, limited HLA diversity, and/or in older epidemics. Indeed, a recent comparative analysis of HLA-associated polymorphism spread among subtype C epidemics in sub-Saharan Africa revealed higher HIV adaptation to HLA—as well as a loss of the relative protection of B*57—in the older Botswana epidemic compared to the newer South African one [108]. Studies evaluating antibody neutralization resistance of historic versus modern HIV-1 envelope sequences also support gradual drift of HIV-1 towards a more neutralization-resistant phenotype [109, 110]. Taken together, evidence suggests that HIV-1 is becoming – albeit gradually – increasingly pre-adapted to host immunity over time.

Intuitively, transmission of pre-adapted HIV-1 sequences should compromise host immunity, but direct evidence is lacking. Nevertheless, the observation that transmission of B*57-escape mutations compromises B*57-mediated viral control supports this intuition [66], as do the observations that individuals infected by partners with shared HLA alleles experience higher early setpoint viral loads [111], and that B*51 and B*57 are less protective in populations with high prevalence of HLA-associated escape mutations [99, 108]. On the other hand, a majority of escape associations occur at positions that are unlikely to affect epitope processing or HLA binding [16], suggesting that they act at the level of TCR engagement – but the

consequences of transmission of such variants remains unclear. Importantly, the mosaic vaccine approach is predicated on the assumption that viral sequence variants, including escape variants, can be immunogenic, highlighting the central importance of understanding the immunologic consequences of escape mutation transmission and spread. Similarly unresolved are the consequences of escape mutation spread on HIV-1 replication [96, 108, 112, 113] and virulence [108, 114]. If transmission of escape mutations accelerates disease progression, then average viral loads may increase over time, though such effects would likely be modulated by fitness costs of escape and other evolutionary pressures that limit increases in viral load [72, 108]. The potential for escape mutation spread to alter the natural history of HIV-1 is an area ripe for further study, as the rate and extent to which this is occurring in different host populations could have direct bearing on vaccine efficacy.

Concluding remarks

Although HIV-1 escape represents a major challenge to host antiviral control, it has provided a key insights into what constitutes an effective cellular immune response – information that has in turn yielded novel vaccine approaches. Based on these advances, it is reasonable to hope that the study of immune-driven HIV-1 escape will ultimately help us develop strategies to end this pandemic.

Box: Outstanding questions

1. How can we continue to leverage studies of HLA-driven viral adaptation to gain insights into what constitutes effective cellular immunity?

2. How can we best identify key HIV-1 epitopes and/or regions where escape is impossible, or highly costly? Similarly, how do we best assess the magnitude of selection pressure on particular HIV-1 regions, and resulting viral fitness costs, in a meaningful way?

3. Can adapted (escaped) epitopes still be immunogenic, and, if so, how should we approach incorporating such variants into vaccine design?

4. To what extent, and how quickly, are HLA-associated escape mutations accumulating in circulating HIV-1 sequences? Is this phenomenon universal, or largely population-specific – if the latter, what are the host and viral determinants of this process?

5. What are the consequences of HIV-1 escape mutation spread for host antiviral immunity, viral pathogenesis, and transmission as the pandemic progresses?

Glossary

CTL epitope nomenclature in HIV-1: HIV-1 CTL epitopes are 8-12mer peptides derived from viral proteins. They are named according to their N- and C-terminal residues followed by their length. For example, the TSTLQEQIGW epitope (Gag codons 240-249) is called Gag-TW10.

The Los Alamos HIV molecular immunology database maintains a list of the best-defined HIV-1 CTL epitopes (http://www.hiv.lanl.gov/content/immunology/tables/optimal_ctl_summary.html).

HIV-1 control: refers to the ability of certain individuals to naturally maintain lower-than-average plasma viral loads. Elite controllers, rare HIV-infected individuals able to maintain viral loads below clinical detection limits for extended periods of time without medications, provide an extreme example of this phenotype. HIV-1 control correlates with expression of certain protective HLA class I alleles, including HLA-B*57, HLA-B*27, and others.

HLA-associated HIV-1 polymorphism: HIV-1 amino acids that are statistically significantly associated with expression of a particular HLA class I allele. HLA-associated polymorphisms are identified via statistical analysis of linked HIV-1/HLA genotype datasets. There are two types: *Nonadapted* associations are HIV-1 amino acids that are under-represented among persons expressing the HLA; these represent the inferred immunologically susceptible forms particular to that HLA. *Adapted* associations are HIV-1 amino acids enriched among persons expressing the HLA; these represent the inferred escape forms particular to that HLA.

HLA class I allele resolution/classification: the first two digits of the HLA name denotes the allele group (*e.g.*, members of the HLA-A*02 group encode variants of the A*02 protein, or antigen). This is also referred to as antigen-level resolution. The second set of digits specifies the allele (e.g. HLA-A*02:01 is the first allelic member of the A*02 group). This is also referred to as allele- or subtype-level resolution. HLA alleles can also be classified into supertypes that comprise HLA alleles of different types/subtypes that share peptide binding specificities. The HLA B58 supertype, for example, comprises subtypes belonging to the B*57 and B*58 groups, among others. For more information on HLA nomenclature, visit http://hla.alleles.org/.

Phylogenetic correction: the process of correcting for the underlying evolutionary relationships between HIV-1 sequences in a dataset (*i.e.*, the inferred viral phylogeny) when identifying HLA-associated polymorphisms in HIV-1. See **Figure 2**.

T-Cell receptor (TCR) Clonotype: clonal descendants of a specific T-lymphocyte that express the unique TCR α/β chain rearrangement of the original cell.

Viral fitness: the ability of a viral strain to replicate in a given environment. Fitness is highly context-dependent. For example, HIV-1 sequences harboring B*57-associated escape mutations are more fit than wild-type in a B*57-expressing host, as these mutations allow the virus to evade host immunity. However, if these mutations hamper the virus' ability to replicate or function, wild-type HIV-1 is likely to be more fit in a non-B*57-expressing host (or *in vitro*).

Viral load: HIV-1 burden in blood plasma, quantified as HIV-1 RNA copies/ml.

Viral load setpoint: After approximately 6 months post-infection, plasma viral loads stabilize at a "setpoint" level that varies widely among patients. The median viral setpoint is roughly 30,000 (4.2 $\log_{10}$) copies/mL in untreated chronic HIV-1 infection [115].

Figures

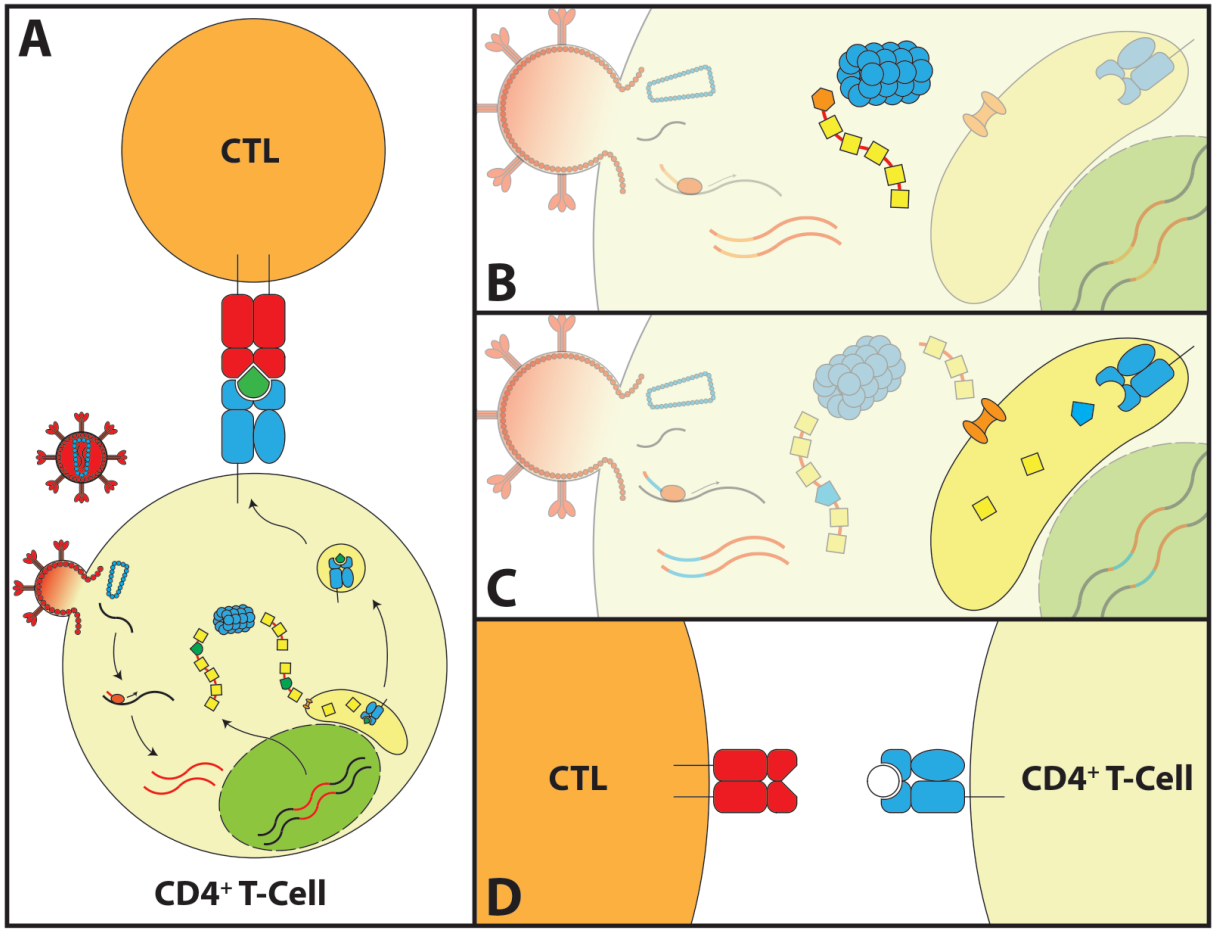

Figure 1. HIV-1 antigen presentation to CTL, and mechanisms of immune escape. **(A)** Following HIV-1 entry, reverse transcription, and integration of proviral DNA into the infected cell's genome, viral proteins (denoted by strings of geometric shapes) are produced and processed into peptide epitopes (denoted by individual geometric shapes) by the host cellular machinery. Epitopes are loaded onto HLA class I molecules for presentation at the cell surface. Recognition of the epitope–HLA complex by a T-cell receptor (TCR) expressed by a $CD8^+$ T-lymphocyte (CTL) results in CTL-mediated elimination of the infected cell. **(B)** Antigen processing escape. HIV-1 mutations can arise during reverse transcription. Here, a mutation within a viral polypeptide (denoted by the orange hexagon) abrogates the proper processing of the polypeptide by the proteasome, leading to a lack of epitope formation. **(C)** HLA binding escape. Here, a mutant polypeptide is properly processed into an epitope (denoted by a blue pentagon), but the mutant epitope cannot bind HLA. **(D)** TCR escape. Here, a mutant epitope (denoted by white circle) is processed and loaded onto HLA. However, the mutant epitope–HLA complex is not recognized by the TCR expressed by the CTL, allowing the infected cell to escape CTL recognition.

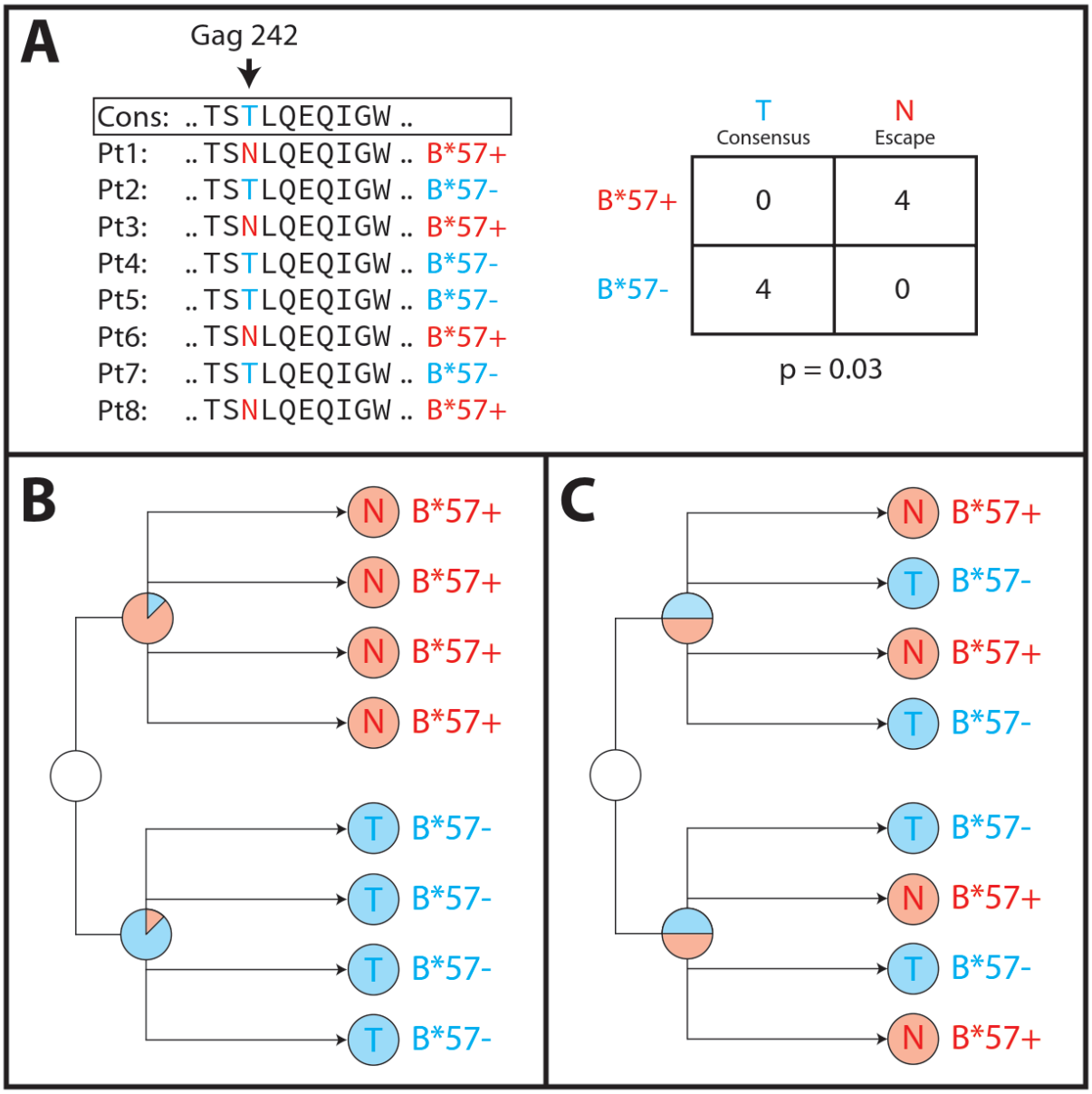

Figure 2. Identifying HLA associated polymorphisms in HIV-1 in a phylogenetically-informed manner.

(A) A hypothetical dataset containing linked HIV-1 sequences (Gag amino acids 240-249) and host HLA data (B*57 status) analyzed using Fisher's exact test. The population consensus sequence (Cons) is at the top, and patients (Pt1-8) are stratified according to their HLA-B*57 status (+ vs. -) and their HIV-1 amino acid at Gag codon 242 (consensus "T" vs. escape "N"). This yields a p-value of p=0.03, supporting a significant association between B*57 and "N" at Gag position 242. **(B)** Standard statistical approaches, however, ignore the phylogenetic structure of the data. Let us say that the cohort HIV sequences exhibited a hypothetical inferred phylogeny as shown here (where circled letters at the tree tips represent the HIV residue observed in the patients and pie charts at the internal nodes represent the probability distribution over "T" and "N" at the inferred most recent ancestor). In this case, the observed patient sequences are consistent with neutral evolution along the tree (*i.e.* the addition of patient HLA information does

not further help explain the observed results). As such, phylogenetically-corrected methods would yield a *less* significant p-value than Fisher's exact test. **(C)** In contrast, if HIV-1 sequences exhibited an inferred phylogeny as shown, consideration of patient HLA information does help explain the observed results. As such, phylogenetically-corrected methods would yield a *more* significant p-value than Fisher's exact test.

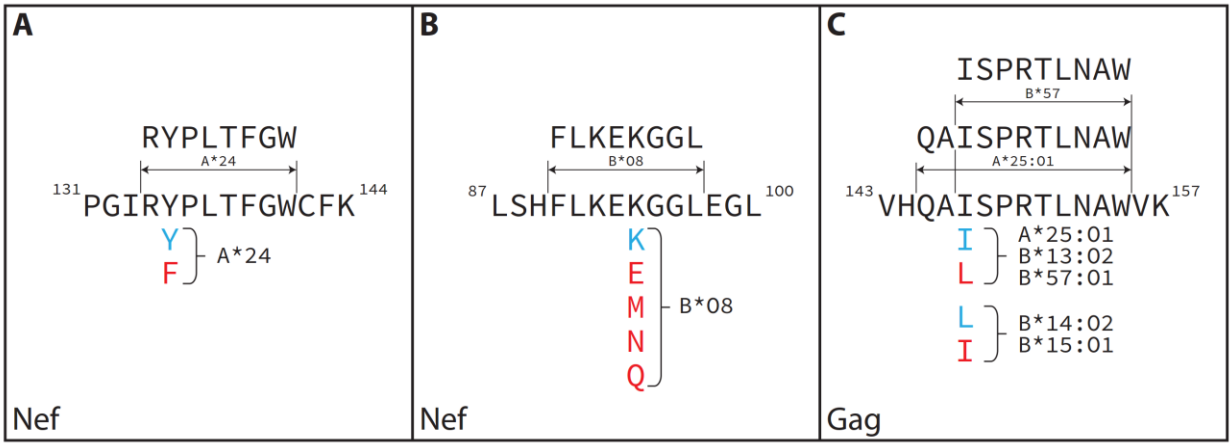

Figure 3. Examples of HLA-associated polymorphisms identified in HIV-1.
HLA associated polymorphisms identified by statistical association are commonly displayed in HIV-1 immune escape maps, which indicate their HIV-1 locations, specific amino acid residues and HLA restrictions [16]. Here, the HIV-1 subtype B consensus amino acid sequence is used as a reference, with known HLA-restricted CTL epitopes shown above and HLA-associated polymorphisms shown below. Nonadapted associations (HIV-1 amino acids that are under-represented among persons expressing the HLA) are in blue – these represent the inferred immunologically susceptible forms particular to that HLA. Adapted associations (HIV-1 amino acids enriched among persons expressing the HLA) are in red - these represent the inferred escape forms particular to that HLA. **(A)** At HIV-1 Nef codon 135, located at position 2 of the HLA-A*24-restricted RW8 epitope, Y and F represent the nonadapted (susceptible) and adapted (escaped) forms associated with HLA-A*24, respectively. **(B)** In some cases, a single HLA can drive multiple possible escape pathways at a single site. At Nef codon 94, position 5 of the B*08 restricted FL8 epitope, the B*08-associated nonadapted form is the subtype B consensus K, but there are four possible adapted forms: E, M, N, and Q. **(C)** Some HIV-1 codons come under pressure by multiple HLA alleles, sometimes in opposing directions. At Gag codon 147, the adapted form associated with HLA-A*25:01, B*13:02 and B*57:01 is L, whereas the adapted form selected by HLA-B*14:02 and B*15:01 is the subtype B consensus I.

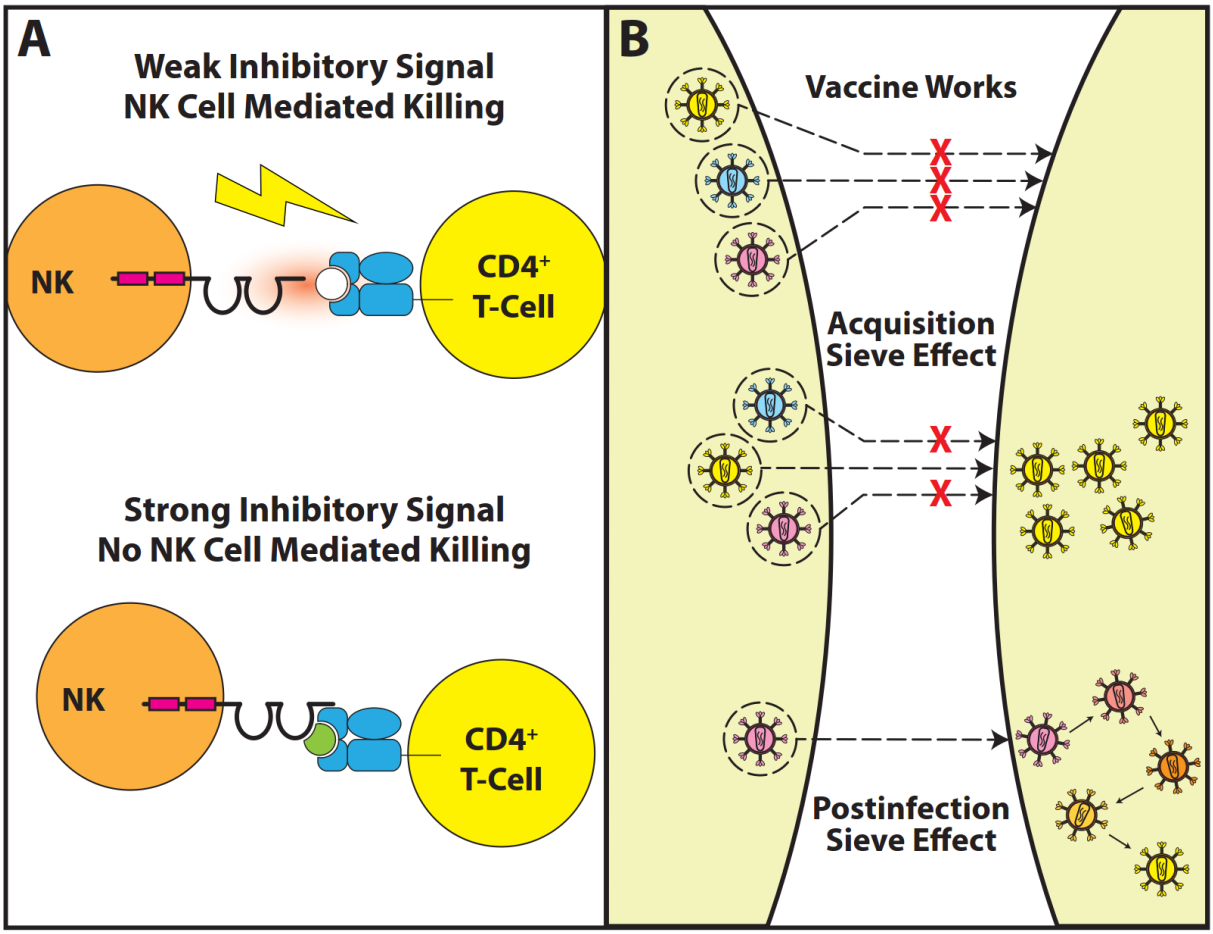

Figure 4. Other types of immune-driven selection in HIV-1: KIR and vaccine-induced escape.
(A) Mutations in HIV-1 could theoretically impair Natural Killer (NK) cell-mediated recognition of HIV-infected cells, thereby conferring escape from innate immunity. Based on a putative KIR2DL2 HIV-1 escape mutation described by Alter *et al.* [89] this figure illustrates how NK cell escape could occur. Top: weak interactions between the inhibitory NK KIR2DL2 receptor and the viral peptide/HLA-C complex on the HIV-infected cell produce weak inhibitory signals, leading to NK cell mediated elimination of the infected cell. Bottom: viral escape mutations that enhance KIR2DL2-mediated NK cell recognition of the peptide/HLA-C complex on the HIV-infected cell enhance binding of the inhibitory KIR2DL2 receptor, thus protecting the infected cell from NK cell-mediated elimination. (**B)** Sieve effects demonstrate immunogenicity in HIV-1 vaccine trials. Top: a sterilizing HIV-1 vaccine induces immune responses that block infection by any/all incoming HIV-1 strains. Middle: vaccination induces a partial barrier that blocks infection by HIV-1 strains similar to vaccine immunogen, but not those that are antigenetically divergent from it (acquisition sieve effect). Bottom: vaccination focuses immune responses on epitopes shared between founder virus and vaccine immunogen, leading to rapid *in vivo* escape (postinfection sieve effect). Although mechanistically distinct, acquisition and postinfection sieve effects are difficult to distinguish via the analysis of breakthrough HIV-1 sequences.

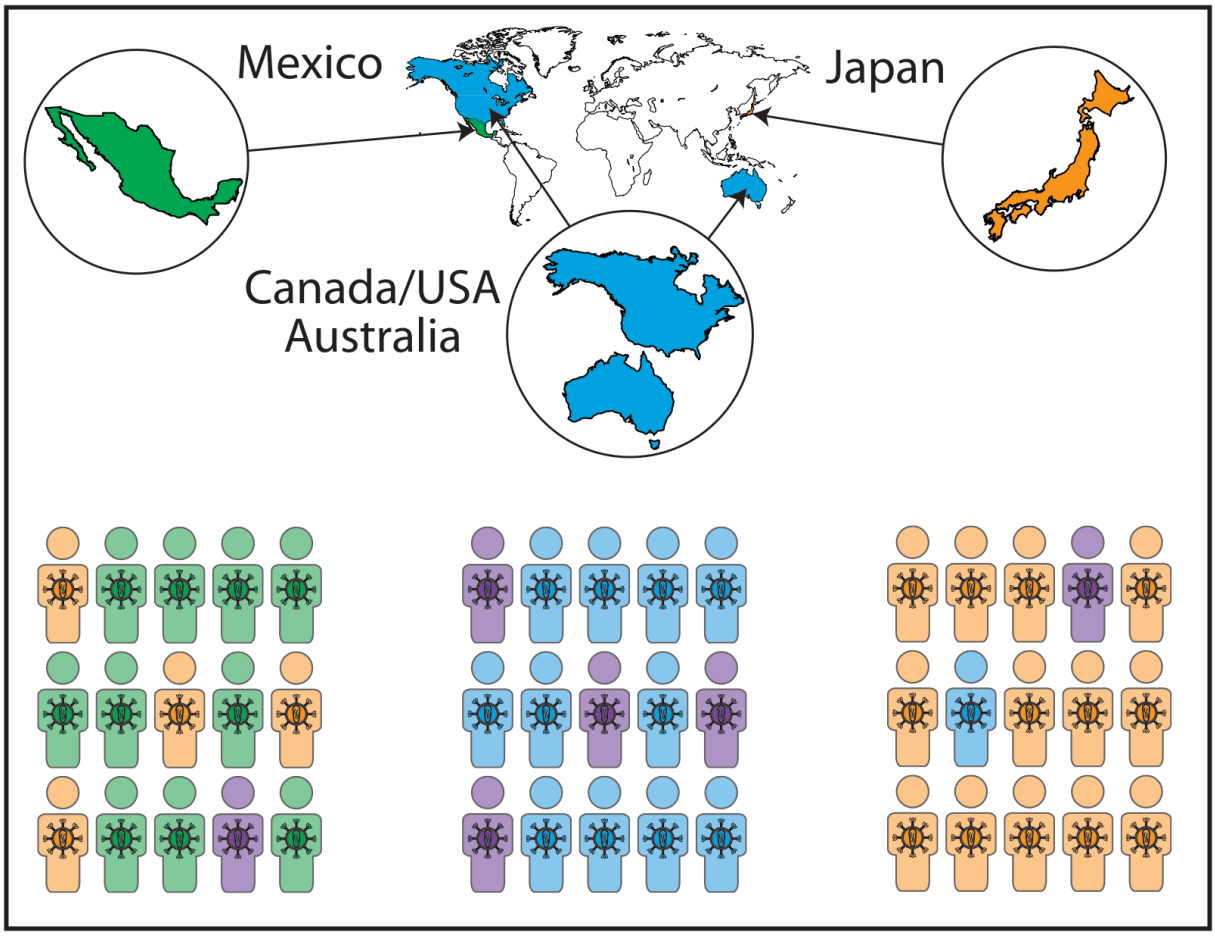

Figure 5. HIV-1 sequences in a host population will harbor adaptations to HLA alleles expressed in that population.
HIV-1 genomes in an infected individual will harbor adaptations specific to the HLA alleles expressed by that individual (denoted by matching colors of virus and host). By extension, HIV-1 sequences in a host population will harbor adaptations to HLA alleles expressed in that population. Host populations in Mexico, Canada/USA/Australia, and Japan exhibit very different HLA class I allele distributions (denoted by different host colors); as such, adaptations exhibited by HIV-1 sequences will also differ between these populations (denoted by different virus colors). HLA is therefore a major driver of global HIV-1 diversity. Note that, because reversion is neither instantaneous nor universal, escape mutations for certain HLA alleles will be found in some proportion of individuals lacking those alleles (not shown).

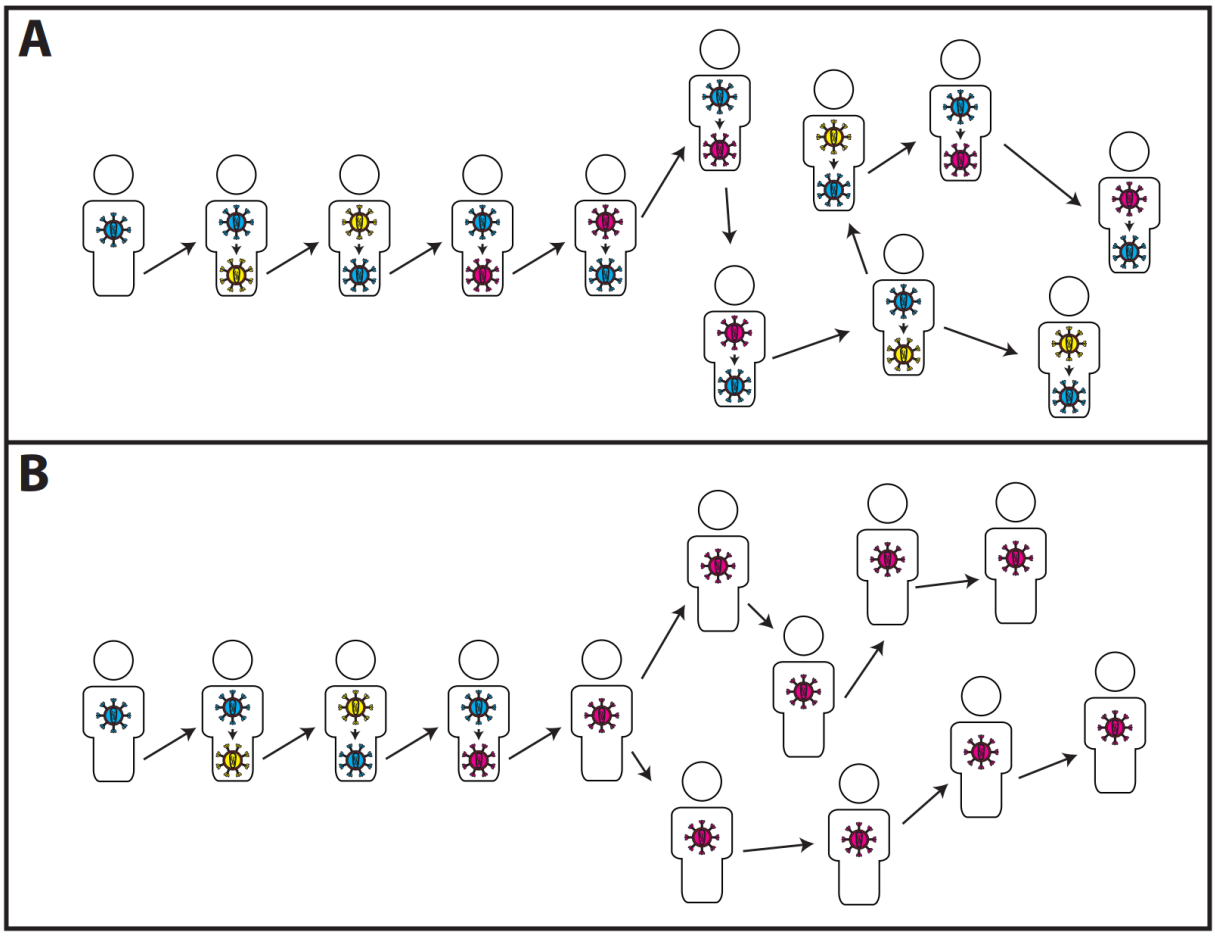

Figure 6. Reversion versus persistence of escape mutations following transmission: consequences for HIV-1 evolution at the population level.
Some escape mutations (denoted in pink and yellow) will consistently revert to consensus (denoted in blue) upon transmission to a host lacking the restricting HLA. However, certain escape mutations stably persist upon transmission to an HLA-unmatched host. Consistent reversion versus persistence of escape mutations has implications for the evolution of HIV-1 sequences over time. **(A)** If escape mutations consistently reverted to consensus following transmission to an HLA-mismatched host, the frequencies of consensus and escaped forms would remain stable over time, and the risk of acquiring the escaped form would broadly correlate with the frequency of the HLA in the population. **(B)** If an escape mutation stably persisted upon transmission to HLA-unmatched hosts, perhaps due to a lack of fitness cost or effective compensation, then the frequency of this escaped form in circulation would gradually increase as the epidemic progresses.

Acknowledgements

The authors thank Mark Brockman for critical review of the manuscript. AQL is supported by a Frederick Banting and Charles Best M.Sc. award from the Canadian Institutes of Health Research (CIHR). ZLB is supported by a New Investigator Award from the CIHR and a Scholar Award from the Michael Smith Foundation for Health Research (MSFHR). The authors declare no conflict of interest.

References

1 Faria, N.R.*, et al.* (2014) HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* 346, 56-61
2 Hemelaar, J. (2012) The origin and diversity of the HIV-1 pandemic. *Trends Mol Med* 18, 182-192
3 Koup, R.A.*, et al.* (1994) Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J Virol* 68, 4650-4655
4 Borrow, P.*, et al.* (1994) Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *J Virol* 68, 6103-6110
5 Goonetilleke, N.*, et al.* (2009) The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *J Exp Med* 206, 1253-1272
6 Matano, T.*, et al.* (1998) Administration of an anti-CD8 monoclonal antibody interferes with the clearance of chimeric simian/human immunodeficiency virus during primary infections of rhesus macaques. *J Virol* 72, 164-169
7 Carrington, M. and O'Brien, S.J. (2003) The influence of HLA genotype on AIDS. *Annu Rev Med* 54, 535-551
8 Apps, R.*, et al.* (2013) Influence of HLA-C expression level on HIV control. *Science* 340, 87-91
9 Kiepiela, P.*, et al.* (2007) CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med* 13, 46-53
10 Mothe, B.*, et al.* (2011) Definition of the viral targets of protective HIV-1-specific T cell responses. *J Transl Med* 9, 208
11 Pereyra, F.*, et al.* (2014) HIV control is mediated in part by CD8+ T-cell targeting of specific epitopes. *J Virol*
12 Phillips, R.E.*, et al.* (1991) Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* 354, 453-459
13 Leslie, A.J.*, et al.* (2004) HIV evolution: CTL escape mutation and reversion after transmission. *Nat Med* 10, 282-289
14 Brumme, Z.L.*, et al.* (2008) Marked epitope- and allele-specific differences in rates of mutation in human immunodeficiency type 1 (HIV-1) Gag, Pol, and Nef cytotoxic T-lymphocyte epitopes in acute/early HIV-1 infection. *J Virol* 82, 9216-9227
15 Feeney, M.E.*, et al.* (2005) HIV-1 viral escape in infancy followed by emergence of a variant-specific CTL response. *J Immunol* 174, 7524-7530
16 Carlson, J.M.*, et al.* (2012) Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. *J Virol* 86, 13202-13216
17 Yokomaku, Y.*, et al.* (2004) Impaired processing and presentation of cytotoxic-T-lymphocyte (CTL) epitopes are major escape mechanisms from CTL immune pressure in human immunodeficiency virus type 1 infection. *J Virol* 78, 1324-1332
18 Draenert, R.*, et al.* (2004) Immune selection for altered antigen processing leads to cytotoxic T lymphocyte escape in chronic HIV-1 infection. *J Exp Med* 199, 905-915
19 Cardinaud, S.*, et al.* (2011) CTL escape mediated by proteasomal destruction of an HIV-1 cryptic epitope. *PLoS Pathog* 7, e1002049
20 Kim, V. and Green, W.R. (1997) The role of proximal and distal sequence variations in the presentation of an immunodominant CTL epitope encoded by the ecotropic AK7 MuLV. *Virology* 236, 221-233
21 Bronke, C.*, et al.* (2013) HIV escape mutations occur preferentially at HLA-binding sites of CD8 T-cell epitopes. *AIDS* 27, 899-905
22 Goulder, P.J.*, et al.* (1997) Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat Med* 3, 212-217
23 Iglesias, M.C.*, et al.* (2011) Escape from highly effective public CD8+ T-cell clonotypes by HIV. *Blood* 118, 2138-2149
24 Reid, S.W.*, et al.* (1996) Antagonist HIV-1 Gag peptides induce structural changes in HLA B8. *J Exp Med* 184, 2279-2286
25 Moore, C.B.*, et al.* (2002) Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* 296, 1439-1443

26 Brumme, Z.L.*, et al.* (2009) HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. *PLoS ONE* 4, e6687
27 Bhattacharya, T.*, et al.* (2007) Founder effects in the assessment of HIV polymorphisms and HLA allele associations. *Science* 315, 1583-1586
28 Carlson, J.M.*, et al.* (2008) Phylogenetic dependency networks: inferring patterns of CTL escape and codon covariation in HIV-1 Gag. *PLoS Comput Biol* 4, e1000225
29 Poon, A.F.*, et al.* (2010) Phylogenetic analysis of population-based and deep sequencing data to identify coevolving sites in the nef gene of HIV-1. *Mol Biol Evol* 27, 819-832
30 Prince, J.L.*, et al.* (2012) Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis. *PLoS Pathog* 8, e1003041
31 Gesprasert, G.*, et al.* (2010) HLA-associated immune pressure on Gag protein in CRF01_AE-infected individuals and its association with plasma viral load. *PLoS ONE* 5, e11179
32 Kloverpris, H.N.*, et al.* (2012) HLA-B*57 Micropolymorphism shapes HLA allele-specific epitope immunogenicity, selection pressure, and HIV immune control. *J Virol* 86, 919-929
33 Carlson, J.M.*, et al.* (2012) Widespread Impact of HLA Restriction on Immune Control and Escape Pathways of HIV-1. *J Virol* 86, 5230-5243
34 Almeida, C.A.*, et al.* (2010) Exploiting knowledge of immune selection in HIV-1 to detect HIV-specific CD8 T-cell responses. *Vaccine* 28, 6052-6057
35 Berger, C.T.*, et al.* (2010) Viral adaptation to immune selection pressure by HLA class I-restricted CTL responses targeting epitopes in HIV frameshift sequences. *J Exp Med* 207, 61-75, S61-12
36 Bansal, A.*, et al.* (2010) CD8 T cell response and evolutionary pressure to HIV-1 cryptic epitopes derived from antisense transcription. *J Exp Med* 207, 51-59, S51-53
37 Brumme, Z.L.*, et al.* (2007) Evidence of Differential HLA Class I-Mediated Viral Evolution in Functional and Accessory/Regulatory Genes of HIV-1. *PLoS Pathog* 3, e94
38 Payne, R.P.*, et al.* (2014) Differential escape patterns within the dominant HLA-B*57:03-restricted HIV Gag epitope reflect distinct clade-specific functional constraints. *J Virol* 88, 4668-4678
39 Derdeyn, C.A.*, et al.* (2004) Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* 303, 2019-2022
40 Keele, B.F.*, et al.* (2008) Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* 105, 7552-7557
41 Herbeck, J.T.*, et al.* (2011) Demographic processes affect HIV-1 evolution in primary infection before the onset of selective processes. *J Virol* 85, 7523-7534
42 Price, D.A.*, et al.* (1997) Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proc Natl Acad Sci U S A* 94, 1890-1895
43 Borrow, P.*, et al.* (1997) Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nat Med* 3, 205-211
44 Henn, M.R.*, et al.* (2012) Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. *PLoS Pathog* 8, e1002529
45 Bimber, B.N.*, et al.* (2009) Ultradeep pyrosequencing detects complex patterns of CD8+ T-lymphocyte escape in simian immunodeficiency virus-infected macaques. *J Virol* 83, 8247-8253
46 Brackenridge, S.*, et al.* (2011) An early HIV mutation within an HLA-B*57-restricted T cell epitope abrogates binding to the killer inhibitory receptor 3DL1. *J Virol* 85, 5415-5422
47 Fischer, W.*, et al.* (2010) Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS ONE* 5, e12303
48 Lemey, P.*, et al.* (2006) HIV evolutionary dynamics within and among hosts. *AIDS Rev* 8, 125-140
49 Martin, E.*, et al.* (2014) Early immune adaptation in HIV-1 revealed by population-level approaches. *Retrovirology* 11, 64
50 Wang, Y.E.*, et al.* (2009) Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. *J Virol* 83, 1845-1855
51 Matthews, P.C.*, et al.* (2008) Central role of reverting mutations in HLA associations with human immunodeficiency virus set point. *J Virol* 82, 8548-8559

52 Boutwell, C.L*., et al.* (2013) Frequent and variable cytotoxic-T-lymphocyte escape-associated fitness costs in the human immunodeficiency virus type 1 subtype B Gag proteins. *J Virol* 87, 3952-3965
53 Chen, H*., et al.* (2012) TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nat Immunol* 13, 691-700
54 Schneidewind, A*., et al.* (2008) Structural and functional constraints limit options for cytotoxic T-lymphocyte escape in the immunodominant HLA-B27-restricted epitope in human immunodeficiency virus type 1 capsid. *J Virol* 82, 5594-5605
55 Schneidewind, A*., et al.* (2007) Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. *J Virol* 81, 12382-12393
56 Li, B*., et al.* (2007) Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. *J Virol* 81, 193-201
57 Duda, A*., et al.* (2009) HLA-associated clinical progression correlates with epitope reversion rates in early human immunodeficiency virus infection. *J Virol* 83, 1228-1239
58 Treurnicht, F.K*., et al.* (2010) Adaptive changes in HIV-1 subtype C proteins during early infection are driven by changes in HLA-associated immune pressure. *Virology* 396, 213-225
59 Matthews, P.C*., et al.* (2011) HLA-A*7401-mediated control of HIV viremia is independent of its linkage disequilibrium with HLA-B*5703. *J Immunol* 186, 5675-5686
60 Prado, J.G*., et al.* (2009) Functional consequences of human immunodeficiency virus escape from an HLA-B*13-restricted CD8+ T-cell epitope in p1 Gag protein. *J Virol* 83, 1018-1025
61 Wright, J*., et al.* (2010) Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. *Journal of Virology* 84, 10820-10831
62 Brockman, M.A*., et al.* (2007) Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. *J Virol* 81, 12608-12618
63 Brockman, M.A*., et al.* (2012) Uncommon pathways of immune escape attenuate HIV-1 integrase replication capacity. *J Virol* 86, 6913-6923
64 Ueno, T*., et al.* (2008) CTL-mediated selective pressure influences dynamic evolution and pathogenic functions of HIV-1 Nef. *J Immunol* 180, 1107-1116
65 Martinez-Picado, J*., et al.* (2006) Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. *J Virol* 80, 3617-3623
66 Crawford, H*., et al.* (2009) Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. *J Exp Med* 206, 909-921
67 Navis, M*., et al.* (2007) Viral replication capacity as a correlate of HLA B57/B5801-associated nonprogressive HIV-1 infection. *J Immunol* 179, 3133-3143
68 Miura, T*., et al.* (2009) HLA-B57/B*5801 human immunodeficiency virus type 1 elite controllers select for rare gag variants associated with reduced viral replication capacity and strong cytotoxic T-lymphotye recognition. *J Virol* 83, 2743-2755
69 Miura, T*., et al.* (2010) Impaired replication capacity of acute/early viruses in persons who become HIV controllers. *J Virol* 84, 7581-7591
70 Hecht, F.M*., et al.* (2010) HIV RNA level in early infection is predicted by viral load in the transmission source. *AIDS* 24, 941-945
71 Hollingsworth, T.D*., et al.* (2010) HIV-1 transmitting couples have similar viral load set-points in Rakai, Uganda. *PLoS Pathog* 6, e1000876
72 Fraser, C*., et al.* (2014) Virulence and pathogenesis of HIV-1 infection: an evolutionary perspective. *Science* 343, 1243727
73 Brockman, M.A*., et al.* (2010) 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* 84, 11937-11949
74 Wright, J.K*., et al.* (2011) Influence of Gag-protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. *J Virol* 85, 3996-4006
75 Chopera, D.R*., et al.* (2008) Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage. *PLoS Pathog* 4, e1000033

76 Goepfert, P.A.*, et al.* (2008) Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. *J Exp Med* 205, 1009-1017
77 Carlson, J.M.*, et al.* (2014) Selection bias at the heterosexual HIV-1 transmission bottleneck. *Science* 345, 1254031
78 Allen, T.M. and Altfeld, M. (2008) Crippling HIV one mutation at a time. *J Exp Med* 205, 1003-1007
79 Borthwick, N.*, et al.* (2013) Vaccine-elicited human T cells recognizing conserved protein regions inhibit HIV-1. *Mol Ther*
80 Rolland, M.*, et al.* (2013) HIV-1 conserved-element vaccines: relationship between sequence conservation and replicative capacity. *J Virol* 87, 5461-5467
81 Ferguson, A.L.*, et al.* (2013) Translating HIV sequences into quantitative fitness landscapes predicts viral vulnerabilities for rational immunogen design. *Immunity* 38, 606-617
82 Barouch, D.H.*, et al.* (2010) Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. *Nat Med* 16, 319-323
83 Barouch, D.H.*, et al.* (2013) Protective efficacy of a global HIV-1 mosaic vaccine against heterologous SHIV challenges in rhesus monkeys. *Cell* 155, 531-539
84 Sunshine, J.*, et al.* (2014) Increased sequence coverage through combined targeting of variant and conserved epitopes correlates with control of HIV replication. *J Virol* 88, 1354-1365
85 Rerks-Ngarm, S.*, et al.* (2009) Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* 361, 2209-2220
86 Frahm, N.*, et al.* (2007) Extensive HLA class I allele promiscuity among viral CTL epitopes. *Eur J Immunol* 37, 2419-2433
87 Thananchai, H.*, et al.* (2007) Cutting Edge: Allele-specific and peptide-dependent interactions between KIR3DL1 and HLA-A and HLA-B. *J Immunol* 178, 33-37
88 Peruzzi, M.*, et al.* (1996) A p70 killer cell inhibitory receptor specific for several HLA-B allotypes discriminates among peptides bound to HLA-B*2705. *J Exp Med* 184, 1585-1590
89 Alter, G.*, et al.* (2011) HIV-1 adaptation to NK-cell-mediated immune pressure. *Nature* 476, 96-100
90 Edlefsen, P.T.*, et al.* (2013) Sieve analysis in HIV-1 vaccine efficacy trials. *Curr Opin HIV AIDS* 8, 432-436
91 Fowke, K.R.*, et al.* (2000) HIV-1-specific cellular immune responses among HIV-1-resistant sex workers. *Immunol Cell Biol* 78, 586-595
92 Rolland, M.*, et al.* (2011) Genetic impact of vaccination on breakthrough HIV-1 sequences from the STEP trial. *Nat Med* 17, 366-371
93 Rolland, M.*, et al.* (2012) Increased HIV-1 vaccine efficacy against viruses with genetic signatures in Env V2. *Nature* 490, 417-420
94 Betts, M.R.*, et al.* (2005) Characterization of functional and phenotypic changes in anti-Gag vaccine-induced T cell responses and their role in protection after HIV-1 infection. *Proc Natl Acad Sci U S A* 102, 4512-4517
95 Hertz, T.*, et al.* (2013) HIV-1 vaccine-induced T-cell responses cluster in epitope hotspots that differ from those induced in natural infection with HIV-1. *PLoS Pathog* 9, e1003404
96 Cotton, L.A.*, et al.* (2014) Genotypic and Functional Impact of HIV-1 Adaptation to Its Host Population during the North American Epidemic. *PLoS Genet* 10, e1004295
97 Avila-Rios, S.*, et al.* (2009) Unique features of HLA-mediated HIV evolution in a Mexican cohort: a comparative study. *Retrovirology* 6, 72
98 Chikata, T.*, et al.* (2014) Host-specific adaptation of HIV-1 subtype B in the Japanese population. *J Virol* 88, 4764-4775
99 Kawashima, Y.*, et al.* (2009) Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* 458, 641-645
100 Yagita, Y.*, et al.* (2013) Distinct HIV-1 escape patterns selected by cytotoxic T cells with identical epitope specificity. *J Virol* 87, 2253-2263
101 Tenzer, S.*, et al.* (2014) HIV-1 adaptation to antigen processing results in population-level immune evasion and affects subtype diversification. *Cell Rep* 7, 448-463
102 McKinnon, L.R.*, et al.* (2009) Clade-specific evolution mediated by HLA-B*57/5801 in human immunodeficiency virus type 1 clade A1 p24. *J Virol* 83, 12636-12642
103 Schneidewind, A.*, et al.* (2009) Transmission and long-term stability of compensated CD8 escape mutations. *J Virol* 83, 3993-3997

104 Goulder, P.J.*, et al.* (2001) Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature* 412, 334-338
105 Navis, M.*, et al.* (2008) Molecular evolution of human immunodeficiency virus type 1 upon transmission between human leukocyte antigen disparate donor-recipient pairs. *PLoS ONE* 3, e2422
106 Leslie, A.*, et al.* (2005) Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. *J Exp Med* 201, 891-902
107 Fryer, H.R.*, et al.* (2010) Modelling the evolution and spread of HIV immune escape mutants. *PLoS Pathog* 6, e1001196
108 Payne, R.*, et al.* (2014) Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. *Proc Natl Acad Sci U S A*
109 Bunnik, E.M.*, et al.* (2010) Adaptation of HIV-1 envelope gp120 to humoral immunity at a population level. *Nat Med* 16, 995-997
110 Bouvin-Pley, M.*, et al.* (2013) Evidence for a Continuous Drift of the HIV-1 Species towards Higher Resistance to Neutralizing Antibodies over the Course of the Epidemic. *PLoS Pathog* 9, e1003477
111 Yue, L.*, et al.* (2013) Cumulative impact of host and viral factors on HIV-1 viral-load control during early infection. *J Virol* 87, 708-715
112 Nomura, S.*, et al.* (2013) Significant reductions in Gag-protease-mediated HIV-1 replication capacity during the course of the epidemic in Japan. *J Virol* 87, 1465-1476
113 Juarez-Molina, C.I.*, et al.* (2014) Impact of HLA selection pressure on HIV fitness at a population level in Mexico and Barbados. *J Virol* 88, 10392-10398
114 Herbeck, J.T.*, et al.* (2012) Is the virulence of HIV changing? A meta-analysis of trends in prognostic markers of HIV disease progression and transmission. *AIDS* 26, 193-205
115 Lyles, R.H.*, et al.* (2000) Natural history of human immunodeficiency virus type 1 viremia after seroconversion and proximal to AIDS in a large cohort of homosexual men. Multicenter AIDS Cohort Study. *J Infect Dis* 181, 872-880