

I-B

Identifying HLA-Associated Polymorphisms in HIV-1

Zabrina L. Brumme^{a,b}, Art F. Y. Poon^b,
Jonathan M. Carlson^c, Bruce D. Walker^{d,e}

Abstract

Cytotoxic T-lymphocytes (CTL) eliminate HIV-1 infected cells through the recognition of antigenic peptides displayed by Human Leukocyte Antigen (HLA) class I molecules on the infected cell surface. HLA-restricted CTL responses drive HIV evolution through selection of viral sequence polymorphisms typically referred to as immune escape mutations. In this short review, we highlight recent methodological advances in our efforts to identify HLA-associated polymorphisms throughout the HIV-1 proteome, and describe how these methods have been used to discover novel epitopes, elucidate complex mutational pathways, and enrich our understanding of HIV-1 as it adapts to HLA-mediated selection pressure at the population level. Achieving a deeper understanding of the sites, pathways and consequences of immune escape is of critical relevance to the continued search for an effective CTL-based AIDS vaccine.

I-B-1 Introduction: Immune control of HIV-1 by CTL

Despite dramatic and sustained declines in HIV-related morbidity and mortality following the introduction of Highly Active Antiretroviral Therapy (HAART) in the mid-1990s [Palella *et al.*, 1998], there remains no cure for the 33.4 million people infected with HIV-1 globally, and no vaccine to protect the estimated 2.7 million individuals newly infected each year [UNAIDS & WHO, 2009]. One of the major challenges to the prevention, treatment and potential eradication of HIV is the virus's extensive capacity for adaptation and diversification [Taylor *et al.*,

2008; Gaschen *et al.*, 2002]. A severe genetic bottleneck occurs at the time of transmission such that a single founder virus initiates productive infection in over 80% of heterosexual transmissions [Salazar-Gonzalez *et al.*, 2009; Derdeyn *et al.*, 2004]. Subsequently, however, the virus's rapid and extensive replication, combined with the high error rate of HIV-1 reverse transcriptase, mutation-inducing effects of host RNA editing enzymes such as APOBEC 3G [Wood *et al.*, 2009], and frequent recombination [Neher & Leitner, 2010] result in the rapid diversification of the initial founder into a swarm or "quasispecies" [Kamp, 2003] of related viral variants within the infected individual. This diversity becomes the evolutionary substrate for selection imposed by immune [Goulder & Watkins, 2004; Wei *et al.*, 2003] and antiretroviral [Larder & Kemp, 1989] pressures, yielding continuous evolution of HIV within the host [Shankarappa *et al.*, 1999; Goonetilleke *et al.*, 2009] as well as globally [Kawashima *et al.*, 2009]. Indeed, since the divergence of HIV-1 group M from its most recent SIVcpz ancestor less than 100 years ago [Korber *et al.*, 2000; Worobey *et al.*, 2008], HIV-1 has undergone dramatic diversification, with contemporaneous strains differing by up to 35% in their envelope nucleotide sequence [Gaschen *et al.*, 2002]. Achieving a deeper understanding of how immune selection pressures have and continue to shape intra-individual and global HIV evolution is key to the search for an effective AIDS vaccine.

Over the natural course of infection, immune responses imposed by antibodies [Wei *et al.*, 2003], CD8+ T-lymphocytes [Koup, 1994] and perhaps even innate immune responses [Thananchai *et al.*, 2009] act as a major selective force driving HIV-1 evolution in a continuous, dynamic process known as "immune escape" [McMichael *et al.*, 2010]. This review will focus specifically on mutational escape from cellular immune responses mediated by CD8+ (cytotoxic) T-lymphocytes (CTL). Specifically, we will summarize both historical data and recent advances in efforts to identify immune escape sites across the HIV-1 proteome, as well as discuss the consequences of CTL escape on HIV evolution at the population level.

CTL eliminate HIV-infected cells through recognition of virus-derived peptides (or "epitopes") that are synthesized and processed intracellularly, and presented on the cell surface by Human Leukocyte Antigen (HLA) class I molecules. Due to the requirement that T-cells recognize epitopes in complex with host HLA molecules, and that

^aFaculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

^bBritish Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada

^cMicrosoft Research, Los Angeles, CA, USA

^dRagon Institute of MGH, MIT and Harvard, Boston MA, USA

^eHoward Hughes Medical Institute, Chevy Chase, MD, USA

In *HIV Molecular Immunology 2010*. Karina Yusim, Bette T. M. Korber, Christian Brander, Barton F. Haynes, Richard Koup, John P. Moore, Bruce D. Walker, and David I. Watkins, editors. Publisher: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. p. 3–16.

these HLA molecules are constrained with respect to the epitope motifs that they can present [Klein & Sato, 2000], antigen recognition (and thus the process of CTL escape), is said to be “HLA-restricted”. HIV-specific CTL represent major mediators of host antiviral control as demonstrated by observational [Borrow *et al.*, 1994; Koup *et al.*, 1994; Almeida *et al.*, 2007; Altfeld *et al.*, 2003; Migueles *et al.*, 2000], experimental [Schmitz *et al.*, 1999], and molecular epidemiologic [Carrington & O’Brien, 2003; Kaslow *et al.*, 2001; Trachtenberg *et al.*, 2003; Carrington *et al.*, 1999; Fellay *et al.*, 2007] data. The importance of CTL in immune control of HIV-1 is also demonstrated by the selection of HLA-restricted immune escape mutations in the viral genome. First described in 1991 [Phillips *et al.*, 1991], CTL escape is now understood to begin as early as the first weeks following transmission [Goonetilleke *et al.*, 2009; Borrow *et al.*, 1997; Allen *et al.*, 2000; Turnbull *et al.*, 2009], and accounts for a substantial proportion of observed substitutions in the first year of infection [Goonetilleke *et al.*, 2009; Brumme *et al.*, 2008]. Selection of escape mutations has been documented to occur throughout the disease course [Feeney *et al.*, 2004; Goulder *et al.*, 1997], even after initiation of antiretroviral therapy [Casazza *et al.*, 2005], although available data suggest that there is relatively little immune escape occurring in the late stages of infection [Koibuchi *et al.*, 2005].

I-B-2 Historic studies of CTL escape in HIV-1

Since the original report [Phillips *et al.*, 1991], a large number of escape mutations reproducibly selected in the context of specific HLA class I alleles have been identified across HIV-1 through the analysis of clinically-derived sequences, followed by experimental validation (e.g. Goulder & Watkins [2004]; Borrow *et al.* [1997]; Kelleher *et al.* [2001]; Allen *et al.* [2005]; Leslie *et al.* [2004]). We now know, for example, that three-quarters of persons expressing the “protective” HLA-B*57 allele will select the T242N mutation in Gag (codon three of the TW10 epitope in p24^{Gag}) within the first weeks to months following HIV-1 infection [Brumme *et al.*, 2008; Leslie *et al.*, 2004], whereas in B*27-expressing individuals, the first mutation arising in the immunodominant p24^{Gag} KK10 epitope is typically the L268M, followed by R264K often years later [Goulder & Watkins, 2004; Kelleher *et al.*, 2001; Schneidewind *et al.*, 2007]. We also recognize that although many escape mutations occur within the boundaries of the CTL epitope (thus impairing HLA class I presentation and/or T-cell receptor interactions with the peptide-HLA complex), other mechanisms of escape also occur. For example, mutations within [Yokomaku *et al.*, 2004] or outside [Draenert *et al.*, 2004] epitope bound-

aries may impair intracellular peptide processing, while secondary or compensatory mutations proximal or distal to the epitope site [Schneidewind *et al.*, 2007; Brockman *et al.*, 2007] also occur. Furthermore, we now understand that in some [Leslie *et al.*, 2004; Friedrich *et al.*, 2004; Li *et al.*, 2007; Duda *et al.*, 2009] although not all [Kawashima *et al.*, 2009; Leslie *et al.*, 2005; Schneidewind *et al.*, 2009; Cornelissen *et al.*, 2009] cases, escape mutations revert upon transmission to a new host who lacks the restricting HLA class I allele. The existing global diversity of HIV-1, therefore, has arisen in large part as a result of the tens of millions of cycles of transmission bottlenecks, followed by diversification, reversion and escape, that have occurred since the beginning of the epidemic. Indeed, HLA has been shown to be a strong selection force for both intra- and interclade HIV evolution at a population level [Matthews *et al.*, 2009].

Despite the virus’s seemingly endless capacity for mutation and diversification, a turning point in our understanding of immune escape was the realization that HIV-1 evolution follows generally predictable patterns as a consequence of HLA allele-specific selection pressure as well as inherent viral constraints [Allen *et al.*, 2005]. The most striking example of this comes from genetically identical adult twins infected at the same time with the same virus through injection drug use [Draenert *et al.*, 2006]. In this case, the immunodominance patterns and the earliest mutations that arose were the same in both individuals [Draenert *et al.*, 2006].

I-B-3 Mapping CTL escape sites across the HIV-1 proteome

Equally important to the history of research in CTL escape was the development of novel statistical methods and their application to large clinically-derived HIV-1 sequence datasets to identify sites of HLA-mediated immune selection in a systematic way [Bhattacharya *et al.*, 2007; Carlson *et al.*, 2008; Rousseau *et al.*, 2008; Kosakovsky Pond *et al.*, 2008]. Recent efforts to map escape mutations across the entire length of the viral proteome have examined HIV-1 sequences from large cohorts of chronically infected, untreated individuals to identify HLA-associated polymorphisms in the context of all HLA class I alleles expressed in human populations. These efforts have allowed construction of proteome-wide maps of escape in HIV-1 [Rousseau *et al.*, 2008; Brumme *et al.*, 2009; John *et al.*, 2010; Wang *et al.*, 2009] and an appreciation of the extent to which immune selection pressures have shaped the evolution of this virus. Indeed, the use of polyclonal and/or mosaic immunogens incorporating knowledge of common polymorphisms and escape variants has been proposed as strategy to address the challenge of HIV-1 diversity in vaccine design [Bhattacharya

et al., 2007; Brander *et al.*, 2006; Korber *et al.*, 2009].

The first study to apply statistical approaches to identify escape mutations using population-level data by Moore *et al.* [2002] was published in 2002. In an analysis of more than 400 clinically-derived HIV-1 sequences spanning codons 20–227 of Reverse Transcriptase, Moore *et al.* [2002] identified 64 positive associations between expression of a specific HLA class I allele and a specific polymorphism in RT, and 25 negative associations between an HLA class I allele and change from consensus in multivariate logistic regression models. That such a large number of HLA-associated polymorphisms were identified in this relatively short region underscored the dramatic impact of CTL selection on HIV evolution and provided evidence for population-level HIV adaptation to host immune pressures for the first time.

Since the publication of this study [Moore *et al.*, 2002], the importance of taking into account the underlying evolutionary relationships between HIV-1 sequences when attempting to identify HLA-associated polymorphisms has been recognized and addressed through the development of phylogenetically-informed methods to distinguish sites of HLA-associated immune selection from those more likely to be explained by viral lineage (or “founder”) effects [Bhattacharya *et al.*, 2007; Carlson *et al.*, 2008; Rousseau *et al.*, 2008]. Phylogenetically-informed methods recognize the fact that HIV-1 sequences (or any sequences, for that matter) are related to one another through descent from a common ancestor, some more distantly than others. These methods take this evolutionary information into account when attempting to identify sites of immune selection in HIV, and are thus able to distinguish immune-mediated polymorphisms from those that are more likely explained by neutral evolution from a shared ancestor. Phylogenetically-informed methods continue to be applied to the analysis of population-based datasets (e.g. Carlson *et al.* [2008]; Rousseau *et al.* [2008]; Brumme *et al.* [2009]; John *et al.* [2010]; Wang *et al.* [2009]; Poon *et al.* [2007]; Brumme *et al.* [2007]; Avila-Rios *et al.* [2009]), yielding a comprehensive picture of the extent to which immune selection shapes HIV-1 evolution globally. Detailed knowledge of escape mutations across the entire HIV-1 subtype B [John *et al.*, 2010; Wang *et al.*, 2009] and C [Rousseau *et al.*, 2008] proteomes are now available, while analyses of even larger population-based cohorts of 1000 or more patients have yielded detailed maps of the HIV-1 proteins most relevant to CTL-based vaccine design strategies (e.g. Gag, Pol and Nef) [Carlson *et al.*, 2008; Brumme *et al.*, 2009]. Not only have these studies uncovered hundreds of sites across the HIV-1 proteome under active selection by HLA, but they have also systematically identified specific amino acids representing the adapted (“escaped”) and nonadapted (“susceptible”) forms associated with their restricting HLA alleles. These HLA-associated

polymorphisms have been organized into HIV proteome-wide “escape maps”, revealing not only the sites and specific pathways of immune escape, but also the tremendous density thereof (for an example see Figure I-B-1).

Indeed, up to 40% of amino acids in certain HIV proteins (e.g. Nef) exhibit evidence for active, contemporaneous immune selection pressures [Brumme *et al.*, 2009]. In general, a hierarchy of “escape mutation density” has been observed in HIV-1 proteins, typically reported as accessory/regulatory > Gag > Pol > Env, although these results may vary from study to study due to cohort size, HLA frequency distribution, HIV subtypes and analytic methods used [Rousseau *et al.*, 2008; John *et al.*, 2010; Wang *et al.*, 2009]. Methods to identify viral sites under potential immune selection using HIV-1 sequences in the absence of HLA class I data have also been developed. For example, [Poon *et al.*, 2007] have developed methods to screen large datasets of clinically-derived “bulk” HIV-1 sequences for the presence of ambiguous (mixed) nucleotide bases encoding nonsynonymous substitutions, thus identifying sites under active selection in circulating sequences, while Neher & Leitner [2010] have developed novel methods to estimate recombination rate and selection strength in inpatient HIV evolution using longitudinal clonal envelope sequences.

I-B-4 Recent improvements in tools for CTL escape mapping and data visualization

Methods for phylogenetically informed identification of CTL escape mutations continue to improve and expand. For example, in addition to accounting for the evolutionary relationships between HIV sequences, recent strategies also correct for the strong confounding effects of linkage disequilibrium between HLA class I alleles [Carlson *et al.*, 2008]. Because HLA class I alleles are situated in proximity on human chromosome 6, they tend to be inherited together and are thus said to be in “linkage disequilibrium” with one another; an individual expressing HLA-B*57, for example, will also likely express the linked allele Cw*06. An analysis that does not take HLA linkage disequilibrium into account would likely identify a B*57-driven escape mutation as also being associated with Cw*06, even though the latter allele is not responsible for its selection. While older strategies largely employed manual, post-hoc adjustment for linkage disequilibrium, meaning that researchers would comb through the lists of associations and assign the “correct” allele by hand (e.g. Bhattacharya *et al.* [2007]; Brumme *et al.* [2007]), newer strategies allow the direct identification of the allele driving the association [Carlson *et al.*, 2008].

Similarly, strategies to detect and account for amino acid co-evolution in HIV-1 have also been developed



Figure I-B-1: Detail of an Immune Escape map for HIV-1 subtype B p24^{Gag} (codons 201-300 only). “Adapted” (escape form) amino acids are red while “non-adapted” (susceptible form or wild-type) amino acids are blue. UPPERCASE letters distinguish polymorphisms that survive correction for HIV codon covariation (“direct” associations), while lowercase letters distinguish polymorphisms that do not survive correction for codon covariation (“indirect” associations). Polymorphisms associated with the same HLA allele that occur in proximity to one another are grouped together in yellow boxes. Optimally-defined CTL epitopes containing HLA-associated polymorphisms are indicated above the consensus sequence. Shaded vertical bars separate blocks of 10 amino acids. This map was derived from the recent analysis of more than 1500 HIV-infected individuals from three cohorts in Canada, the USA and Australia [Brumme *et al.*, 2009] and reports all HLA-associated polymorphisms with $q \leq 0.05$.

[Carlson *et al.*, 2008; Poon *et al.*, 2010, 2008]. Amino acid co-evolution occurs when an amino acid substitution at one site preferentially occurs in the context of a substitution at another. These secondary sites may represent co-varying or compensatory mutations associated with the primary site. From an analytical perspective, failure to account for amino acid co-evolution could result in both the primary and compensatory mutations being identified as correlated with the restricting HLA allele, when in fact only the former is directly selected.

The fact that new strategies feature corrections for amino acid co-evolution allows the discrimination of viral polymorphisms that are directly attributable to immune selection from those that may represent compensatory and/or secondary mutations [Carlson *et al.*, 2008; Poon *et al.*, 2007]. This is relevant to many applications, particularly the study of the consequences of immune escape on viral replication capacity and/or HIV protein function. It has been hypothesized that long-term “protective” effects of certain HLA class I alleles (e.g. B*57) on HIV disease progression are due in part to the selection of key escape mutations that compromise viral replication capacity [Schneidewind *et al.*, 2007; Brockman *et al.*, 2007; Martinez-Picado *et al.*, 2006; Troyer *et al.*, 2009], thus resulting in lower viral loads and a more benign disease course [Altfeld & Allen, 2006; Quiñones-Mateu *et al.*, 2000]. However, these fitness costs may be rescued through selection of compensatory mutations that fully or partially restore protein function. Compensatory mutations associated with a number of well-defined CTL escape mutations (for example, those accompanying the B*57-associated T242N and B*27-associated R264K mutations [Kelleher *et al.*, 2001; Schneidewind *et al.*, 2007; Brockman *et al.*, 2007]) were historically identified by visual inspection of HIV sequences followed by experimental validation using site-directed mutants (e.g.

Schneidewind *et al.* [2007, 2008]). However thanks to analytical methods such as those described in the following sections, candidate compensatory mutations can now be systematically identified for experimental validation, thus greatly aiding the characterization of the relative consequences of immune escape on HIV fitness.

Similar to the escape maps that display the primary HLA-associated polymorphisms, graphical tools have also been developed to visualize putative secondary/compensatory pathways on a protein-wide basis, in both two-dimensional as well as 3D structural models. One tool for the identification and interactive display of escape mutations and covarying residues in HIV-1 has been developed by Carlson *et al.* [2008], and is available at <http://research.microsoft.com/en-us/um/redmond/projects/MSCompBio/PhyloViewer/>.

To illustrate the functionality of this tool, we analyzed a dataset of 803 clinically-derived p24^{Gag} sequences with linked HLA class I data from British Columbia, Canada, and focused on primary and secondary polymorphisms identified as being associated with HLA B*27. For example, amino acids in p24^{Gag} that covary with the B*27-associated R264K mutation are displayed in a two-dimensional phylogenetic dependency network (Figure I-B-2), as well as a 3-dimensional structural model of the N-terminus of p24^{Gag} (Figure I-B-3).

This “divide-and-conquer” approach, which is emerging as the *de facto* standard in understanding CTL escape in HIV-1, overcomes the tremendous complexity of finding associations among a large number of HLA alleles and HIV codons by performing pairwise analyses with statistical corrections to adjust for linkage, epistasis, and multiple comparisons. Another direction of methodological development is to analyze all of the variables simultaneously using a full Bayesian network treatment where statistical associations are compactly represented in a directed

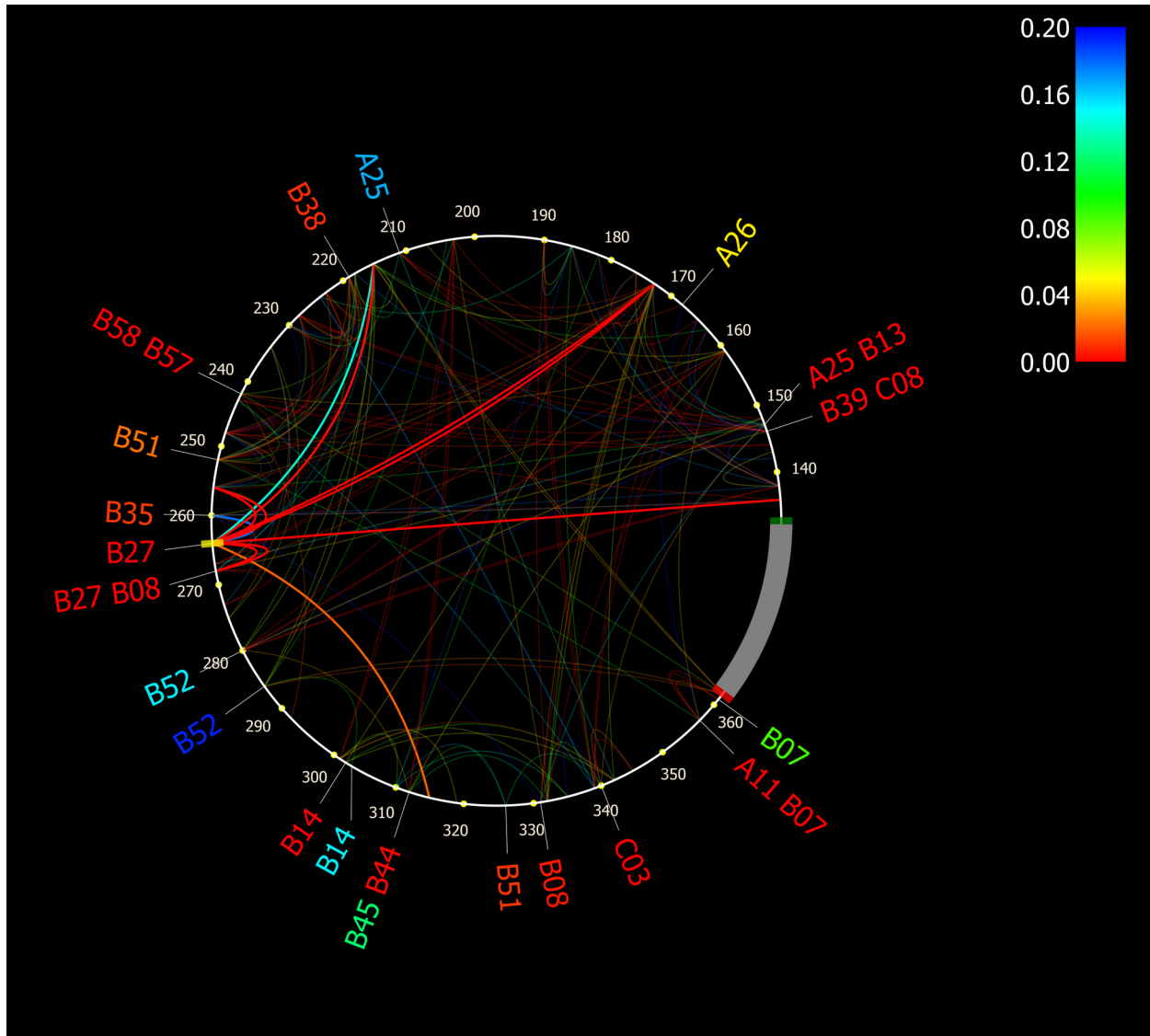


Figure I-B-2: Phylogenetic Dependency Network of p24^{Gag} sites covarying with codon 264. The amino acid sequence of p24^{Gag} is drawn counterclockwise, with the N-terminus at 3 o'clock. HLA-associated polymorphisms with $q \leq 0.2$ are identified at their respective positions along the circle's circumference, while covarying amino acids (also $q \leq 0.2$) are joined together by arcs within the circle. Colors indicate the statistical strength of the association. This figure highlights the specific codons covarying with codon 264, residue 2 of the well-characterized B*27-restricted KK10 epitope in p24^{Gag}. Codon 264 (yellow dash, 9-o'clock position) is joined to its respective covarying sites (p24^{Gag} codons 136, 173, 215, 256, 260, 268 and 315) via brightly-colored arcs.

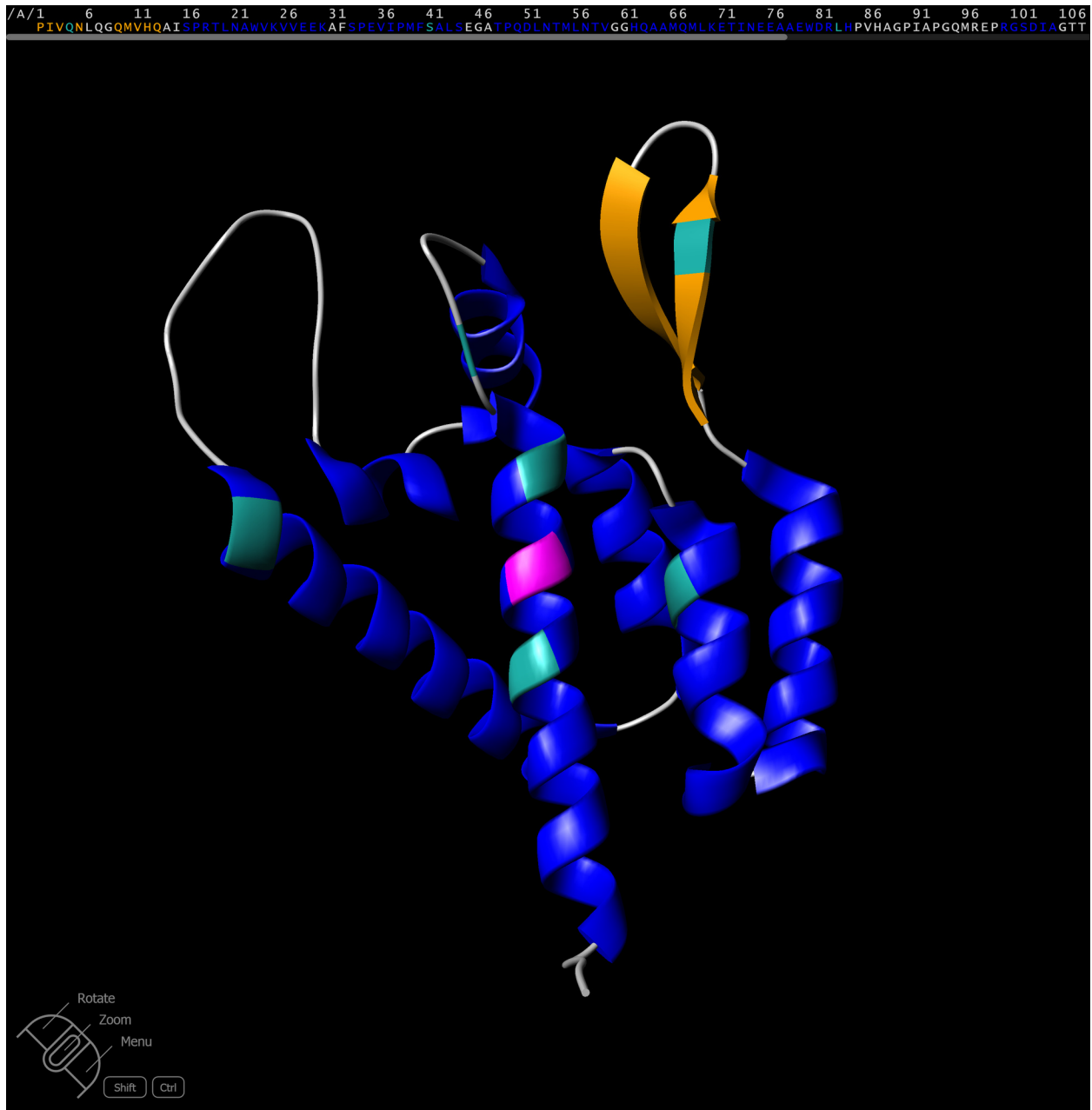


Figure I-B-3: Corresponding three-dimensional model of p24^{Gag} sites covarying with codon 264. The ribbon diagram is a 3D model of the N-terminus of p24^{Gag}. Gag codon 264, residue 2 of the well-characterized B*27-restricted KK10 epitope in p24^{Gag} is shown in pink. Other codons in p24^{Gag} observed to covary with specific codon 264 variants at $q < 0.2$ (identified in this analysis as Gag codons 136, 173, 215, 256, 260, 268) are shown in green. Codon 315 is not included in the 3D structure. The program used to construct Figures I-B-2 and I-B-3 is described in Carlson *et al.* [2008] and is available at <http://research.microsoft.com/en-us/um/redmond/projects/MSCompBio/PhyloDViewer/>

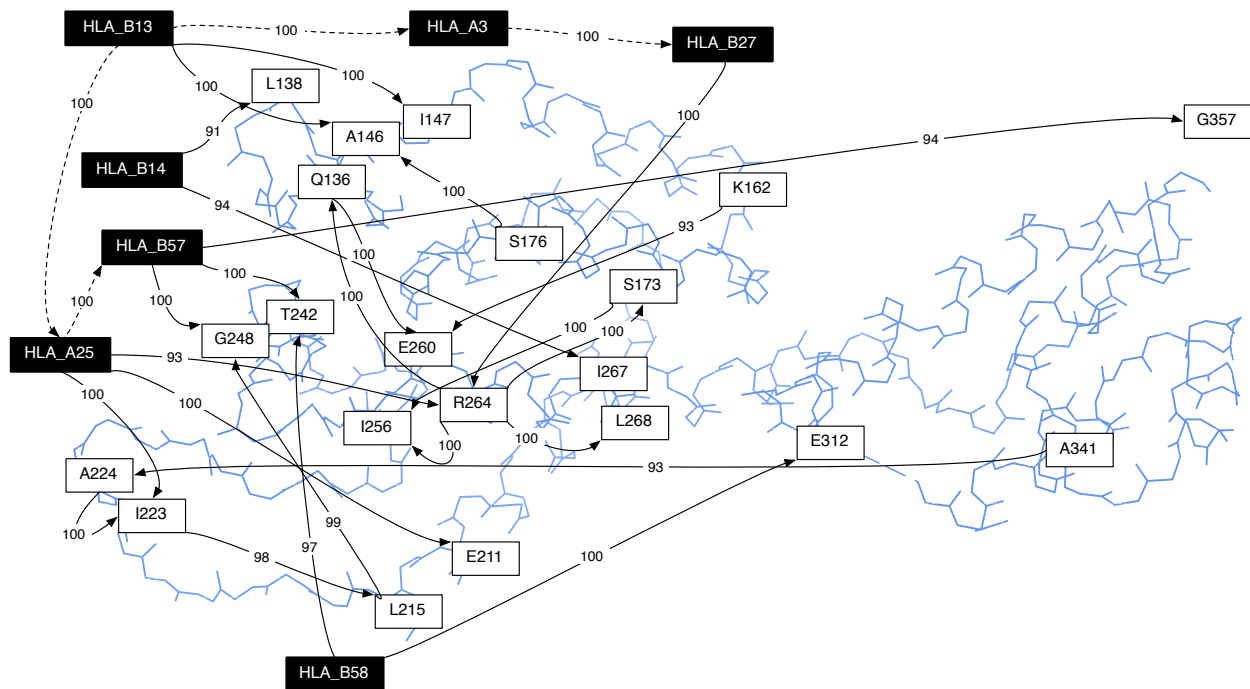


Figure I-B-4: Consensus Bayesian network of HLA alleles and HIV-1 p24 codons. Analysis was restricted to the following HLA alleles only: A*03, A*25, B*13, B*14, B*27, B57 and B*58. Nodes in the network correspond to binary variables encoding the presence or absence of an HLA allele (filled boxes) or nonsynonymous substitutions at a codon (open boxes) as identified by the population consensus residue and numbered from the start of the Gag polypeptide. Codons are mapped onto a structural model of an HIV-1 p24 monomer (Protein Data Bank ID 3H4E). Arrows (directed edges) from variable X to Y indicate that Y is conditionally dependent on X . A dotted edge between HLA alleles represents a putative linkage relationship in the study population, whereas a solid edge indicates sequence co-evolution at those codon sites. 20 codon-to-codon edges that were not part of an HLA-associated network were omitted for clarity. A solid edge from an HLA allele to a codon indicates a putative association due to selection for CTL escape or reversion. Each edge is labelled with the frequency it appears in a sample of 100 graphs, taken at regular intervals from a Markov chain Monte Carlo sample run for 106 steps with the first half discarded as burn-in. Graphs were constrained to have a maximum of 2 parents per node. This analysis was carried out in HyPhy [Kosakovsky Pond *et al.*, 2005] using a modified implementation of the Spidermonkey algorithm [Poon *et al.*, 2008].

acyclic graph [Poon *et al.*, 2008]. In this approach, viral lineage effects are accounted for by mapping nonsynonymous substitutions to the phylogeny using maximum likelihood reconstruction and counting imputed substitutions along the branches. HLA alleles are then assigned to terminal branches and handled as missing data on internal branches (with imputation by Gibbs sampling). A Markov chain Monte Carlo (MCMC) algorithm is employed to sample from the posterior probability distribution of possible graphs. An example of the application of Bayesian networks to modeling CTL escape, undertaken on the same dataset of 803 clinically-derived p24^{Gag} sequences with linked HLA class I data from British Columbia, and focusing on HLA-associated and HIV covariation pathways in p24^{Gag}, is shown in Figure I-B-4, which represents the graph with the highest sampled posterior probability. This strategy incurs an enormous computational cost that necessitates simplifying assumptions such as limiting the number of “parental” variables that

a given variable can be conditioned upon (typically no more than 2). However, the resulting Bayesian network confers the benefit of simultaneously evaluating all associations (CTL escape, linkage, and epistasis) in the same probabilistic framework, which could also facilitate extensions to predictive models relating CTL escape to clinical parameters such as viral load. Methods for carrying out this analysis are publically released in the open-source software package HyPhy (Hypothesis testing using Phylogenies <http://hyphy.org>) [Kosakovsky Pond *et al.*, 2005] and as a user-friendly web application (Spidermonkey <http://www.datamonkey.org>) [Poon *et al.*, 2008].

Regardless of the specific analytical methods used, the pathways of HLA-associated immune escape and amino acid covariation are clearly identifiable. For example, both methods highlighted in the current review identify the well-known HLA-B*27-associated R264K/L268M substitutions within the KK10 epitope in p24^{Gag}, as well

as the upstream S173A compensatory mutation that rescues the fitness costs associated with the R264K [Schneidewind *et al.*, 2007]. Numerous novel pathways are also identified. Indeed, the systematic identification of primary sites of immune escape along with their secondary/covarying mutation pathways is helping to address fundamental knowledge gaps in HIV pathogenesis.

I-B-5 Using these tools to address knowledge gaps in HIV immunobiology and evolution

Quantifying incidence, prevalence and consequences of escape in individuals and populations

Just as the systematic and ongoing identification of drug resistance mutations [Johnson *et al.*, 2009] has been of paramount importance to the treatment and clinical monitoring of HIV infection [Hirsch *et al.*, 2008], the identification of immune escape pathways is similarly relevant to HIV immunobiology research and vaccine design. The availability of a proteome-wide reference list of HLA-associated polymorphisms has allowed researchers to characterize the kinetics of CTL escape in acute/early infection [Brumme *et al.*, 2008; Duda *et al.*, 2009], to quantify their prevalence in different HIV-infected populations [Miura *et al.*, 2009b], to investigate correlations between their presence/absence and HIV clinical parameters such as pVL and/or CD4 count [Brumme *et al.*, 2008; Rousseau *et al.*, 2008; Matthews *et al.*, 2008] and most recently to estimate CTL-mediated evolution following administration of a therapeutic HIV vaccine [Li *et al.*, 2010]. As already mentioned, incorporation of common HLA-associated polymorphisms and escape variants into polyvalent and/or mosaic vaccine immunogens has been proposed as a potential strategy to address the challenge of HIV-1 diversity in vaccine design [Bhattacharya *et al.*, 2007; Brander *et al.*, 2006; Korber *et al.*, 2009].

Discovering novel epitopes

Escape maps can reveal the locations of novel CTL epitopes, which can then be validated experimentally [Bhattacharya *et al.*, 2007]. Most recently, these techniques have been used to scan for the presence of escape mutations in alternative (“cryptic”) HIV reading frames that were silent in the standard reading frame [Bansal *et al.*, 2010; Berger *et al.*, 2010]. Berger *et al.* [2010] identified a novel HLA-A*03-restricted frameshift epitope in an alternative Pol (Integrase) reading frame, likely generated through the incorrect use of a Leucine start codon upstream of the epitope, which may contribute to *in vivo* HIV immune control in A*03+ individuals. Bansal *et al.*

[2010] identified numerous epitopes encoded within anti-sense HIV reading frames, some of which escaped relatively rapidly following infection, suggesting strong immune selection. Taken together with accumulating data from SIV models [Maness *et al.*, 2009, 2010], results suggest that that cryptic epitopes may represent important and untapped sources of T cell epitopes for vaccine immunogens.

Comparative studies across populations, HLA groups and HIV-1 subtypes

The ability to identify HLA-associated polymorphisms in HIV-1 has also allowed comparative studies of escape across populations. In a study of North and Central American (Mexican) populations, Avila-Rios *et al.* [2009] reported unique HIV-1 escape patterns in Mexico, predominantly associated with HLA-B*39, a group that contains alleles which are highly specific for Amerindian origin populations. This may be a much broader phenomenon affecting many common 2-digit HLA class I allele groups containing 4-digit subtypes which have diverged functionally across different racial/ethnic groups through human evolution [John *et al.*, 2010]. Comparative studies also allow universal escape pathways to be distinguished from subtype-specific escape pathways; for example, the B*57-associated T242N escape mutation in Gag is commonly selected in subtypes B [Leslie *et al.*, 2004; Brumme *et al.*, 2009; John *et al.*, 2010; Wang *et al.*, 2009], C [Carlson *et al.*, 2008] and D [McKinnon *et al.*, 2009], but is rarely selected in subtype A1 [McKinnon *et al.*, 2009].

The extent to which HLA-associated selection pressures shape HIV evolution at the population level can be examined by comparing the prevalence of escape mutations throughout the epidemic’s history [Schellens *et al.*, 2009] as well as in different contemporaneous cohorts worldwide [Kawashima *et al.*, 2009; Avila-Rios *et al.*, 2009]. In an analysis of more than 2,800 HIV-infected individuals from 9 cohorts spanning 5 continents, Kawashima *et al.* demonstrated that the frequency of the B*51-associated T135X mutation in HIV RT (occurring at the C-terminus of the B*51-TI8 epitope) correlated strongly with HLA-B*51 prevalence in the cohort [Kawashima *et al.*, 2009]. For example, in Kumamoto, Japan, where HLA-B*51 frequency exceeds 20%, the prevalence of T135X approaches 75%, strongly supporting population-level HIV adaptation to HLA and suggesting that the “protective” effects of certain HLA class I alleles may eventually be compromised as a result of the fixation of “escaped” forms in circulating viral sequences [Kawashima *et al.*, 2009]. Indeed, the B*27-associated R264K mutation in p24^{Gag}, when accompanied by the upstream S173A compensatory mutation appears to remain stable upon transmission [Schneidewind *et al.*, 2009; Cornelissen *et al.*, 2009], and B*57-associated mutations

within the p24^{Gag} KF11 appear to be similarly stabilized in the presence of the upstream compensatory mutation S165N [Crawford *et al.*, 2007]. Taken together, these data suggest that HLA alleles currently associated with relative control of HIV infection (e.g. HLA-B*57, HLA-B*27 and in some cohorts, HLA-B*51), may cease to be associated with beneficial effects as HLA-specific escape mutations accumulate in contemporary circulating sequences [Leslie *et al.*, 2005]. Indeed, developing a deeper understanding of the multiple interacting factors that influence the probability and rates of mutational escape and reversion as HIV is transmitted from person to person—including fitness costs of escape [Davenport *et al.*, 2008], presence of compensatory mutations [Schneidewind *et al.*, 2009; Cornelissen *et al.*, 2009], diversity of initial transmitted inoculum [Loh *et al.*, 2009], residues where one individual's HLA-restricted escape mutation represents the susceptible form for another HLA allele, and vice versa [Bhattacharya *et al.*, 2007; Frahm *et al.*, 2007], and other host/viral factors, is of paramount importance to vaccine design.

I-B-6 Application of next-generation sequencing technologies to CTL escape

Our understanding of immune escape dynamics has recently been advanced through use of new high-sensitivity sequencing technologies, combined with analyses of extremely early events in acute HIV-1 infection. Longitudinal Single Genome Amplification (SGA) analyses of acute-phase HIV sequences [Salazar-Gonzalez *et al.*, 2009; Goonetilleke *et al.*, 2009] have revealed that the earliest CTL responses (and corresponding escape events) arise not against the well-characterized immunodominant epitopes [Altfeld *et al.*, 2006], but rather against previously uncharacterized epitopes, and in some cases can be detected during the decline of viremia from peak levels. These results indicate that the first HIV-specific CTL responses arise earlier and may be more potent in controlling acute-phase viremia than previously recognized [Salazar-Gonzalez *et al.*, 2009]. Indeed, the identification of unique epitopes and escape mutations through detailed intra-patient analysis underscores an important limitation of identifying HLA-associated polymorphisms at the population level; although the latter will identify escape mutations that transcend the complexities of host and viral genetic diversity, each individual is also likely to develop unique escape mutations within known or novel CTL epitopes in context of their transmitted founder HIV sequence and T-cell repertoire. Although these “personalized” mutations may not be frequent or reproducible enough to reach statistical significance in population-based studies, they may impact an individual's disease

course as strongly as the selection of well-known escape variants [Salazar-Gonzalez *et al.*, 2009; Miura *et al.*, 2009a].

Most recently, researchers have begun to apply next-generation “ultradeep” sequencing technologies to investigate the dynamics of HIV evolution and escape during acute [Henn *et al.*, 2009; Bimber *et al.*, 2009] and chronic [Poon *et al.*, 2010] infection with extremely high sensitivities. Consistent with the results of SGA, ultradeep HIV sequencing in acute infection has revealed an extremely rapid, dynamic and complex pattern of escape mutations, often occurring simultaneously at multiple sites in the viral proteome, and in many cases accompanied by the selection of compensatory mutations [Henn *et al.*, 2009; Bimber *et al.*, 2009]. Ultradeep sequencing of chronically-infected individuals has helped illuminate novel pathways of intra-individual escape and compensatory mutations that may not necessarily be detectable through population level analysis of “bulk” HIV sequence data [Poon *et al.*, 2010], thereby enriching and complementing existing data on patterns of immune-mediated codon covariation in HIV [Carlson *et al.*, 2008].

I-B-7 Conclusion

The application of phylogenetically informed methods to population-based datasets to identify viral sites under contemporaneous immune selection represents a major advancement in the study of T-cell escape in HIV-1. Indeed, such studies indicate that there are considerable constraints on HIV evolution. Ultimately, the integration of population-level analyses combined with detailed intra-patient analyses and impact of mutations on viral fitness will bring us closer to the goal of achieving a comprehensive understanding of the sites, pathways, kinetics and implications of immune escape at both the individual and population level. This knowledge, in turn, will help us overcome the complex challenges of HIV-1 diversity and evolution in vaccine design.

I-B-8 References

- [Allen *et al.*, 2005] T. M. Allen, M. Altfeld, S. C. Geer, E. T. Kalife, C. Moore, K. M. O'Sullivan, I. DeSouza, M. E. Feeney, R. L. Eldridge, E. L. Maier, D. E. Kaufmann, M. P. Lahaie, L. Reyor, G. Tanzi, M. N. Johnston, C. Brander, R. Draenert, J. K. Rockstroh, H. Jessen, E. S. Rosenberg, S. A. Mallal, & B. D. Walker, 2005. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. *J Virol* **79**(21):13239–13249. On p. 4.
- [Allen *et al.*, 2000] T. M. Allen, D. H. O'Connor, P. Jing, J. L. Dzuris, B. R. Mothé, T. U. Vogel, E. Dunphy, M. E. Liebl, C. Emerson, N. Wilson, K. J. Kunstman, X. Wang, D. B. Allison, A. L. Hughes, R. C. Desrosiers, J. D. Altman, S. M. Wolinsky, A. Sette, & D. I. Watkins, 2000. Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia. *Nature* **407**(6802):386–390. On p. 4.

- [Almeida *et al.*, 2007] J. R. Almeida, D. A. Price, L. Papagno, Z. A. Arkoub, D. Sauce, E. Bornstein, T. E. Asher, A. Samri, A. Schnuriger, I. Theodorou, D. Costagliola, C. Rouzioux, H. Agut, A.-G. Marcelin, D. Douek, B. Autran, & V. Appay, 2007. Superior control of HIV-1 replication by CD8+ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. *J Exp Med* **204**(10):2473–2485. On p. 4.
- [Altfeld *et al.*, 2003] M. Altfeld, M. M. Addo, E. S. Rosenberg, F. M. Hecht, P. K. Lee, M. Vogel, X. G. Yu, R. Draenert, M. N. Johnston, D. Strick, T. M. Allen, M. E. Feeney, J. O. Kahn, R. P. Sekaly, J. A. Levy, J. K. Rockstroh, P. J. Goulder, & B. D. Walker, 2003. Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. *AIDS* **17**(18):2581–2591. On p. 4.
- [Altfeld & Allen, 2006] M. Altfeld & T. M. Allen, 2006. Hitting HIV where it hurts: An alternative approach to HIV vaccine design. *Trends Immunol* **27**(11):504–510. On p. 6.
- [Altfeld *et al.*, 2006] M. Altfeld, E. T. Kalife, Y. Qi, H. Streeck, M. Lichterfeld, M. N. Johnston, N. Burgett, M. E. Swartz, A. Yang, G. Alter, X. G. Yu, A. Meier, J. K. Rockstroh, T. M. Allen, H. Jessen, E. S. Rosenberg, M. Carrington, & B. D. Walker, 2006. HLA alleles associated with delayed progression to AIDS contribute strongly to the initial CD8+ T cell response against HIV-1. *PLoS Med* **3**(10):e403. On p. 11.
- [Avila-Rios *et al.*, 2009] S. Avila-Rios, C. E. Ormsby, J. M. Carlson, H. Valenzuela-Ponce, J. Blanco-Heredia, D. Garrido-Rodriguez, C. Garcia-Morales, D. Heckerman, Z. L. Brumme, S. Mallal, M. John, E. Espinosa, & G. Reyes-Teran, 2009. Unique features of HLA-mediated HIV evolution in a Mexican cohort: A comparative study. *Retrovirology* **6**:72. On pp. 5 & 10.
- [Bansal *et al.*, 2010] A. Bansal, J. Carlson, J. Yan, O. T. Akinsiku, M. Schaefer, S. Sabbaj, A. Bet, D. N. Levy, S. Heath, J. Tang, R. A. Kaslow, B. D. Walker, T. Ndung'u, P. J. Goulder, D. Heckerman, E. Hunter, & P. A. Goepfert, 2010. CD8 T cell response and evolutionary pressure to HIV-1 cryptic epitopes derived from antisense transcription. *J Exp Med* **207**(1):51–59. On p. 10.
- [Berger *et al.*, 2010] C. T. Berger, J. M. Carlson, C. J. Brumme, K. L. Hartman, Z. L. Brumme, L. M. Henry, P. C. Rosato, A. Piechocka-Trocha, M. A. Brockman, P. R. Harrigan, D. Heckerman, D. E. Kaufmann, & C. Brander, 2010. Viral adaptation to immune selection pressure by HLA class I-restricted CTL responses targeting epitopes in HIV frameshift sequences. *J Exp Med* **207**(1):61–75. On p. 10.
- [Bhattacharya *et al.*, 2007] T. Bhattacharya, M. Daniels, D. Heckerman, B. Foley, N. Frahm, C. Kadie, J. Carlson, K. Yusim, B. McMahon, B. Gaschen, S. Mallal, J. I. Mullins, D. C. Nickle, J. Herbeck, C. Rousseau, G. H. Learn, T. Miura, C. Brander, B. Walker, & B. Korber, 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. *Science* **315**(5818):1583–1586. On pp. 4, 5, 10 & 11.
- [Bimber *et al.*, 2009] B. N. Bimber, B. J. Burwitz, S. O'Connor, A. Detmer, E. Gostick, S. M. Lank, D. A. Price, A. Hughes, & D. O'Connor, 2009. Ultradeep pyrosequencing detects complex patterns of CD8+ T-lymphocyte escape in simian immunodeficiency virus-infected macaques. *J Virol* **83**(16):8247–8253. On p. 11.
- [Borrow *et al.*, 1994] P. Borrow, H. Lewicki, B. H. Hahn, G. M. Shaw, & M. B. A. Oldstone, 1994. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *J Virol* **68**(9):6103–6110. On p. 4.
- [Borrow *et al.*, 1997] P. Borrow, H. Lewicki, X. Wei, M. S. Horwitz, N. Peffer, H. Meyers, J. A. Nelson, J. E. Gairin, B. H. Hahn, M. B. Oldstone, & G. M. Shaw, 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**(2):205–11. On p. 4.
- [Brander *et al.*, 2006] C. Brander, N. Frahm, & B. D. Walker, 2006. The challenges of host and viral diversity in HIV vaccine design. *Curr Opin Immunol* **18**(4):430–437. On pp. 5 & 10.
- [Brockman *et al.*, 2007] M. A. Brockman, A. Schneidewind, M. Lahaie, A. Schmidt, T. Miura, I. DeSouza, F. Ryvkin, C. A. Derdeyn, S. Allen, E. Hunter, J. Mulenga, P. A. Goepfert, B. D. Walker, & T. M. Allen, 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**(22):12608–12618. On pp. 4 & 6.
- [Brumme *et al.*, 2008] Z. L. Brumme, C. J. Brumme, J. Carlson, H. Streeck, M. John, Q. Eichbaum, B. L. Block, B. Baker, C. Kadie, M. Markowitz, H. Jessen, A. D. Kelleher, E. Rosenberg, J. Kaldor, Y. Yuki, M. Carrington, T. M. Allen, S. Mallal, M. Altfeld, D. Heckerman, & B. D. Walker, 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**(18):9216–9227. On pp. 4 & 10.
- [Brumme *et al.*, 2007] Z. L. Brumme, C. J. Brumme, D. Heckerman, B. T. Korber, M. Daniels, J. Carlson, C. Kadie, T. Bhattacharya, C. Chui, J. Szinger, T. Mo, R. S. Hogg, J. S. G. Montaner, N. Frahm, C. Brander, B. D. Walker, & P. R. Harrigan, 2007. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. *PLoS Pathog* **3**(7):e94. On p. 5.
- [Brumme *et al.*, 2009] Z. L. Brumme, M. John, J. M. Carlson, C. J. Brumme, D. Chan, M. A. Brockman, L. C. Swenson, I. Tao, S. Szeto, P. Rosato, J. Sela, C. M. Kadie, N. Frahm, C. Brander, D. W. Haas, S. A. Riddler, R. Haubrich, B. D. Walker, P. R. Harrigan, D. Heckerman, & S. Mallal, 2009. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. *PLoS One* **4**(8):e6687. On pp. 4, 5, 6 & 10.
- [Carlson *et al.*, 2008] J. M. Carlson, Z. L. Brumme, C. M. Rousseau, C. J. Brumme, P. Matthews, C. Kadie, J. I. Mullins, B. D. Walker, P. R. Harrigan, P. J. R. Goulder, & D. Heckerman, 2008. Phylogenetic dependency networks: Inferring patterns of CTL escape and codon covariation in HIV-1 gag. *PLoS Comput Biol* **4**(11):e1000225. On pp. 4, 5, 6, 8, 10 & 11.
- [Carrington *et al.*, 1999] M. Carrington, G. W. Nelson, M. P. Martin, T. Kissner, D. Vlahov, J. J. Goedert, R. Kaslow, S. Buchbinder, K. Hoots, & S. J. O'Brien, 1999. HLA and HIV-1: Heterozygote advantage and B*35-Cw*04 disadvantage. *Science* **283**(5408):1748–1752. On p. 4.
- [Carrington & O'Brien, 2003] M. Carrington & S. J. O'Brien, 2003. The influence of HLA genotype on AIDS. *Annu Rev Med* **54**:535–551. On p. 4.
- [Casazza *et al.*, 2005] J. P. Casazza, M. R. Betts, B. J. Hill, J. M. Brechley, D. A. Price, D. C. Douek, & R. A. Koup, 2005. Immunologic pressure within class I-restricted cognate human immunodeficiency virus epitopes during highly active antiretroviral therapy. *J Virol* **79**(6):3653–3663. On p. 4.
- [Cornelissen *et al.*, 2009] M. Cornelissen, F. M. Hoogland, N. K. T. Back, S. Jurriaans, F. Zorgdrager, M. Bakker, K. Brinkman, M. Prins, & A. C. van der Kuyl, 2009. Multiple transmissions of a stable human leucocyte antigen-B27 cytotoxic T-cell-escape strain of HIV-1 in The Netherlands. *AIDS* **23**(12):1495–1500. On pp. 4, 10 & 11.
- [Crawford *et al.*, 2007] H. Crawford, J. G. Prado, A. Leslie, S. Hué, I. Honeyborne, S. Reddy, M. van der Stok, Z. Mncube, C. Brander, C. Rousseau, J. I. Mullins, R. Kaslow, P. Goepfert, S. Allen, E. Hunter, J. Mulenga, P. Kiepiela, B. D. Walker, & P. J. R. Goulder, 2007. Compensatory mutation partially restores fitness and delays reversion of escape mutation within the immunodominant HLA-B*5703-restricted Gag epitope in chronic human immunodeficiency virus type 1 infection. *J Virol* **81**(15):8346–8351. On p. 11.

- [Davenport *et al.*, 2008] M. P. Davenport, L. Loh, J. Petracic, & S. J. Kent, 2008. Rates of HIV immune escape and reversion: Implications for vaccination. *Trends Microbiol* **16**(12):561–566. On p. 11.
- [Derdeyn *et al.*, 2004] C. A. Derdeyn, J. M. Decker, F. Bibollet-Ruche, J. L. Mokili, M. Muldoon, S. A. Denham, M. L. Heil, F. Kasolo, R. Munsonda, B. H. Hahn, G. M. Shaw, B. T. Korber, S. Allen, & E. Hunter, 2004. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* **303**(5666):2019–2022. On p. 3.
- [Draenert *et al.*, 2006] R. Draenert, T. M. Allen, Y. Liu, T. Wrin, C. Chappey, C. L. Verrill, G. Sirera, R. L. Eldridge, M. P. Lahaie, L. Ruiz, B. Clotet, C. J. Petropoulos, B. D. Walker, & J. Martinez-Picado, 2006. Constraints on HIV-1 evolution and immunodominance revealed in monozygotic adult twins infected with the same virus. *J Exp Med* **203**(3):529–539. On p. 4.
- [Draenert *et al.*, 2004] R. Draenert, S. Le Gall, K. J. Pfafferott, A. J. Leslie, P. Chetty, C. Brander, E. C. Holmes, S.-C. Chang, M. E. Feeney, M. M. Addo, L. Ruiz, D. Ramduth, P. Jeena, M. Altfeld, S. Thomas, Y. Tang, C. L. Verrill, C. Dixon, J. G. Prado, J. Kiepiela, J. Martinez-Picado, B. D. Walker, & P. J. R. Goulder, 2004. Immune selection for altered antigen processing leads to cytotoxic T lymphocyte escape in chronic HIV-1 infection. *J Exp Med* **199**(7):905–915. On p. 4.
- [Duda *et al.*, 2009] A. Duda, L. Lee-Turner, J. Fox, N. Robinson, S. Dustan, S. Kaye, H. Fryer, M. Carrington, M. McClure, A. R. Mclean, S. Fidler, J. Weber, R. E. Phillips, A. J. Frater, & SPARTAC Trial Investigators, 2009. HLA-associated clinical progression correlates with epitope reversion rates in early human immunodeficiency virus infection. *J Virol* **83**(3):1228–1239. On pp. 4 & 10.
- [Feeney *et al.*, 2004] M. E. Feeney, Y. Tang, K. A. Roosevelt, A. J. Leslie, K. McIntosh, N. Karthas, B. D. Walker, & P. J. R. Goulder, 2004. Immune escape precedes breakthrough human immunodeficiency virus type 1 viremia and broadening of the cytotoxic T-lymphocyte response in an HLA-B27-positive long-term-nonprogressing child. *J Virol* **78**(16):8927–8930. On p. 4.
- [Fellay *et al.*, 2007] J. Fellay, K. V. Shianna, D. Ge, S. Colombo, B. Ledergerber, M. Weale, K. Zhang, C. Gumbs, A. Castagna, A. Cosarizza, A. Cozzi-Lepri, A. De Luca, P. Easterbrook, P. Francioli, S. Mallal, J. Martinez-Picado, J. M. Miro, N. Obel, J. P. Smith, J. Wyniger, P. Descombes, S. E. Antonarakis, N. L. Letvin, A. J. McMichael, B. F. Haynes, A. Telenti, & D. B. Goldstein, 2007. A whole-genome association study of major determinants for host control of HIV-1. *Science* **317**(5840):944–947. On p. 4.
- [Frahm *et al.*, 2007] N. Frahm, D. E. Kaufmann, K. Yusim, M. Muldoon, C. Kesmir, C. H. Linde, W. Fischer, T. M. Allen, B. Li, B. H. McMahon, K. L. Faircloth, H. S. Hewitt, E. W. Mackey, T. Miura, A. Khatri, S. Wolinsky, A. McMichael, R. K. Funkhouser, B. D. Walker, C. Brander, & B. T. Korber, 2007. Increased sequence diversity coverage improves detection of HIV-specific T cell responses. *J Immunol* **179**(10):6638–50. On p. 11.
- [Friedrich *et al.*, 2004] T. C. Friedrich, E. J. Dodds, L. J. Yant, L. Vojnov, R. Rudersdorf, C. Cullen, D. T. Evans, R. C. Desrosiers, B. R. Mothé, J. Sidney, A. Sette, K. Kunstman, S. Wolinsky, M. Piatak, J. Lifson, A. L. Hughes, N. Wilson, D. H. O'Connor, & D. I. Watkins, 2004. Reversion of CTL escape-variant immunodeficiency viruses in vivo. *Nat Med* **10**(3):275–281. On p. 4.
- [Gaschen *et al.*, 2002] B. Gaschen, J. Taylor, K. Yusim, B. Foley, F. Gao, D. Lang, V. Novitsky, B. Haynes, B. H. Hahn, T. Bhattacharya, & B. Korber, 2002. Diversity considerations in HIV-1 vaccine selection. *Science* **296**(5577):2354–2360. On p. 3.
- [Goonetilleke *et al.*, 2009] N. Goonetilleke, M. K. P. Liu, J. F. Salazar-Gonzalez, G. Ferrari, E. Giorgi, V. V. Ganusov, B. F. Keele, G. H. Learn, E. L. Turnbull, M. G. Salazar, K. J. Weinhold, S. Moore, CHAVI Clinical Core B, N. Letvin, B. F. Haynes, M. S. Cohen, P. Hraber, T. Bhattacharya, P. Borrow, A. S. Perelson, B. H. Hahn, G. M. Shaw, B. T. Korber, & A. J. McMichael, 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**(6):1253–1272. On pp. 3, 4 & 11.
- [Goulder *et al.*, 1997] P. J. Goulder, R. E. Phillips, R. A. Colbert, S. McAdam, G. Ogg, M. A. Nowak, P. Giangrande, G. Luzzi, B. Morgan, A. Edwards, A. J. McMichael, & S. Rowland-Jones, 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat Med* **3**(2):212–217. On p. 4.
- [Goulder & Watkins, 2004] P. J. R. Goulder & D. I. Watkins, 2004. HIV and SIV CTL escape: Implications for vaccine design. *Nat Rev Immunol* **4**(8):630–640. On pp. 3 & 4.
- [Henn *et al.*, 2009] M. R. Henn, C. Boutwell, N. Lennon, K. Power, C. Malboeuf, P. Charlebois, A. Gladden, J. Levin, M. Casali, L. Phillips, A. Berlin, A. Berical, R. Erlich, S. Anderson, H. Streeck, M. Kemper, E. Ryan, Y. Wang, L. Green, K. Axten, Z. Brumme, C. Brumme, C. Russ, E. Rosenberg, H. Jessen, M. Altfeld, C. Nusbaum, B. Walker, B. Birren, & T. M. Allen, 2009. Ultra-deep sequencing of full-length HIV-1 genomes identifies rapid viral evolution during acute infection. *Retrovirology* **6**(S3):P400. On p. 11.
- [Hirsch *et al.*, 2008] M. S. Hirsch, H. F. Günthard, J. M. Schapiro, F. Brun-Vézinet, B. Clotet, S. M. Hammer, V. A. Johnson, D. R. Kuritzkes, J. W. Mellors, D. Pillay, P. G. Yeni, D. M. Jacobsen, & D. D. Richman, 2008. Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society—USA panel. *Clin Infect Dis* **47**(2):266–285. On p. 10.
- [John *et al.*, 2010] M. John, D. Heckerman, I. James, L. P. Park, J. M. Carlson, A. Chopra, S. Gaudieri, D. Nolan, D. W. Haas, S. A. Riddler, R. Haubrich, & S. Mallal, 2010. Adaptive interactions between HLA and HIV-1: Highly divergent selection imposed by HLA class I molecules with common supertype motifs. *J Immunol* **184**(8):4368–4377. On pp. 4, 5 & 10.
- [Johnson *et al.*, 2009] V. A. Johnson, F. Brun-Vézinet, B. Clotet, H. F. Günthard, D. R. Kuritzkes, D. Pillay, J. M. Schapiro, & D. D. Richman, 2009. Update of the drug resistance mutations in HIV-1: December 2009. *Top HIV Med* **17**(5):138–145. On p. 10.
- [Kamp, 2003] C. Kamp, 2003. A quasispecies approach to viral evolution in the context of an adaptive immune system. *Microbes Infect* **5**(15):1397–1405. On p. 3.
- [Kaslow *et al.*, 2001] R. A. Kaslow, C. Rivers, J. Tang, T. J. Bender, P. A. Goepfert, R. El Habib, K. Weinhold, M. J. Mulligan, & NIAID AIDS vaccine evaluation group, 2001. Polymorphisms in HLA class I genes associated with both favorable prognosis of human immunodeficiency virus (HIV) type 1 infection and positive cytotoxic T-lymphocyte responses to ALVAC-HIV recombinant canarypox vaccines. *J Virol* **75**(18):8681–8689. On p. 4.
- [Kawashima *et al.*, 2009] Y. Kawashima, K. Pfafferott, J. Frater, P. Matthews, R. Payne, M. Addo, H. Gatanaga, M. Fujiwara, A. Hachiya, H. Koizumi, N. Kuse, S. Oka, A. Duda, A. Prendergast, H. Crawford, A. Leslie, Z. Brumme, C. Brumme, T. Allen, C. Brander, R. Kaslow, J. Tang, E. Hunter, S. Allen, J. Mulenga, S. Branch, T. Roach, M. John, S. Mallal, A. Ogwu, R. Shapiro, J. G. Prado, S. Fidler, J. Weber, O. G. Pybus, P. Klenerman, T. Ndung'u, R. Phillips, D. Heckerman, P. R. Harrigan, B. D. Walker, M. Takiguchi, & P. Goulder, 2009. Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* **458**(7238):641–645. On pp. 3, 4 & 10.
- [Kelleher *et al.*, 2001] A. D. Kelleher, C. Long, E. C. Holmes, R. L. Allen, J. Wilson, C. Conlon, C. Workman, S. Shaunak, K. Olson, P. Goulder, C. Brander, G. Ogg, J. S. Sullivan, W. Dyer, I. Jones, A. J.

- McMichael, S. Rowland-Jones, & R. E. Phillips, 2001. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. *J Exp Med* **193**(3):375–386. On pp. 4 & 6.
- [Klein & Sato, 2000] J. Klein & A. Sato, 2000. The HLA system—first of two parts. *N Engl J Med* **343**(10):702–709. On p. 4.
- [Koibuchi *et al.*, 2005] T. Koibuchi, T. M. Allen, M. Lichterfeld, S. K. Mui, K. M. O’Sullivan, A. Trocha, S. A. Kalams, R. P. Johnson, & B. D. Walker, 2005. Limited sequence evolution within persistently targeted CD8 epitopes in chronic human immunodeficiency virus type 1 infection. *J Virol* **79**(13):8171–8181. On p. 4.
- [Korber *et al.*, 2000] B. Korber, M. Muldoon, J. Theiler, F. Gao, R. Gupta, A. Lapedes, B. H. Hahn, S. Wolinsky, & T. Bhattacharya, 2000. Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**(5472):1789–1796. On p. 3.
- [Korber *et al.*, 2009] B. T. Korber, N. L. Letvin, & B. F. Haynes, 2009. T-cell vaccine strategies for human immunodeficiency virus, the virus with a thousand faces. *J Virol* **83**(17):8300–8314. On pp. 5 & 10.
- [Kosakovsky Pond *et al.*, 2005] S. L. Kosakovsky Pond, S. D. W. Frost, & S. V. Muse, 2005. HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* **21**(5):676–679. On p. 9.
- [Kosakovsky Pond *et al.*, 2008] S. L. Kosakovsky Pond, A. F. Y. Poon, S. Zárate, D. M. Smith, S. J. Little, S. K. Pillai, R. J. Ellis, J. K. Wong, A. J. Leigh Brown, D. D. Richman, & S. D. W. Frost, 2008. Estimating selection pressures on HIV-1 using phylogenetic likelihood models. *Stat Med* **27**(23):4779–4789. On p. 4.
- [Koup, 1994] R. A. Koup, 1994. Virus escape from CTL recognition. *J Exp Med* **180**(3):779–782. On p. 3.
- [Koup *et al.*, 1994] R. A. Koup, J. T. Safrit, Y. Cao, C. A. Andrews, G. McLeod, W. Borkowsky, C. Farthing, & D. D. Ho, 1994. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J Virol*, **68**(7):4650–4655. On p. 4.
- [Larder & Kemp, 1989] B. A. Larder & S. D. Kemp, 1989. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* **246**(4934):1155–1158. On p. 3.
- [Leslie *et al.*, 2005] A. Leslie, D. Kavanagh, I. Honeyborne, K. Pfafferott, C. Edwards, T. Pillay, L. Hilton, C. Thobakgale, D. Ramduth, R. Draenert, S. Le Gall, G. Luzzi, A. Edwards, C. Brander, A. K. Sewell, S. Moore, J. Mullins, C. Moore, S. Mallal, N. Bhardwaj, K. Yusim, R. Phillips, P. Klenerman, B. Korber, P. Kiepiela, B. Walker, & P. Goulder, 2005. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. *J Exp Med* **201**(6):891–902. On pp. 4 & 11.
- [Leslie *et al.*, 2004] A. J. Leslie, K. J. Pfafferott, P. Chetty, R. Draenert, M. M. Addo, M. Feeney, Y. Tang, E. C. Holmes, T. Allen, J. G. Prado, M. Altfeld, C. Brander, C. Dixon, D. Ramduth, P. Jeena, S. A. Thomas, A. St. John, T. A. Roach, B. Kupfer, G. Luzzi, A. Edwards, G. Taylor, H. Lyall, G. Tudor-Williams, V. Novelli, J. Martinez-Picado, P. Kiepiela, B. D. Walker, & P. J. R. Goulder, 2004. HIV evolution: CTL escape mutation and reversion after transmission. *Nat Med* **10**(3):282–289. On pp. 4 & 10.
- [Li *et al.*, 2007] B. Li, A. D. Gladden, M. Altfeld, J. M. Kaldor, D. A. Cooper, A. D. Kelleher, & T. M. Allen, 2007. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. *J Virol* **81**(1):193–201. On p. 4.
- [Li *et al.*, 2010] J. Li, C. Brumme, Z. Brumme, H. Wang, J. Spritzler, M. Robertson, M. Lederman, B. Walker, R. Schooley, D. Kuritzkes, & ACTG 5197 Study Team, 2010. Factors associated with viral rebound in HIV+ subjects receiving a therapeutic HIV-1 gag vaccine. In *17th Conference on Retroviruses and Opportunistic Infections*. San Francisco. On p. 10.
- [Loh *et al.*, 2009] L. Loh, J. C. Reece, C. S. Fernandez, S. Alcantara, R. Center, J. Howard, D. F. J. Purcell, M. Balamurali, J. Petracic, M. P. Davenport, & S. J. Kent, 2009. Complexity of the inoculum determines the rate of reversion of SIV Gag CD8 T cell mutant virus and outcome of infection. *PLoS Pathog* **5**(4):e1000378. On p. 11.
- [Maness *et al.*, 2009] N. J. Maness, J. B. Sacha, S. M. Piaskowski, K. L. Weisgrau, E. G. Rakasz, G. E. May, M. B. Buechler, A. D. Walsh, N. A. Wilson, & D. I. Watkins, 2009. Novel translation products from simian immunodeficiency virus SIVmac239 Env-encoding mRNA contain both Rev and cryptic T-cell epitopes. *J Virol* **83**(19):10280–10285. On p. 10.
- [Maness *et al.*, 2010] N. J. Maness, N. A. Wilson, J. S. Reed, S. M. Piaskowski, J. B. Sacha, A. D. Walsh, E. Thoryk, G. J. Heidecker, M. P. Citron, X. Liang, A. J. Bett, D. R. Casimiro, & D. I. Watkins, 2010. Robust, vaccine-induced CD8+ T lymphocyte response against an out-of-frame epitope. *J Immunol* **184**(1):67–72. On p. 10.
- [Martinez-Picado *et al.*, 2006] J. Martinez-Picado, J. G. Prado, E. E. Fry, K. Pfafferott, A. Leslie, S. Chetty, C. Thobakgale, I. Honeyborne, H. Crawford, P. Matthews, T. Pillay, C. Rousseau, J. I. Mullins, C. Brander, B. D. Walker, D. I. Stuart, P. Kiepiela, & P. Goulder, 2006. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. *J Virol* **80**(7):3617–3623. On p. 6.
- [Matthews *et al.*, 2009] P. C. Matthews, A. J. Leslie, A. Katzourakis, H. Crawford, R. Payne, A. Prendergast, K. Power, A. D. Kelleher, P. Klenerman, J. Carlson, D. Heckerman, T. Ndung’u, B. D. Walker, T. M. Allen, O. G. Pybus, & P. J. R. Goulder, 2009. HLA footprints on human immunodeficiency virus type 1 are associated with inter-clade polymorphisms and intra-clade phylogenetic clustering. *J Virol* **83**(9):4605–4615. On p. 4.
- [Matthews *et al.*, 2008] P. C. Matthews, A. Prendergast, A. Leslie, H. Crawford, R. Payne, C. Rousseau, M. Rolland, I. Honeyborne, J. Carlson, C. Kadie, C. Brander, K. Bishop, N. Mlotshwa, J. I. Mullins, H. Coovadia, T. Ndung’u, B. D. Walker, D. Heckerman, & P. J. R. Goulder, 2008. Central role of reverting mutations in HLA associations with human immunodeficiency virus set point. *J Virol* **82**(17):8548–8559. On p. 10.
- [McKinnon *et al.*, 2009] L. R. McKinnon, R. Capina, H. Peters, M. Mendoza, J. Kimani, C. Wachihi, A. Kariri, M. Kimani, M. Richmond, S. K. Kiazky, K. R. Fowke, W. Jaoko, M. Luo, T. B. Ball, & F. A. Plummer, 2009. Clade-specific evolution mediated by HLA-B*57/5801 in human immunodeficiency virus type 1 clade A1 p24. *J Virol* **83**(23):12636–12642. On p. 10.
- [McMichael *et al.*, 2010] A. J. McMichael, P. Borrow, G. D. Tomaras, N. Goonetilleke, & B. F. Haynes, 2010. The immune response during acute HIV-1 infection: Clues for vaccine development. *Nat Rev Immunol* **10**(1):11–23. On p. 3.
- [Migueles *et al.*, 2000] S. A. Migueles, M. S. Sabbaghian, W. L. Shupert, M. P. Bettinotti, F. M. Marincola, L. Martino, C. W. Hallahan, S. M. Selig, D. Schwartz, J. Sullivan, & M. Connors, 2000. HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. *Proc Natl Acad Sci USA* **97**(6):2709–2714. On p. 4.
- [Miura *et al.*, 2009a] T. Miura, M. A. Brockman, A. Schneidewind, M. Lobritz, F. Pereyra, A. Rathod, B. L. Block, Z. L. Brumme, C. J. Brumme, B. Baker, A. C. Rothchild, B. Li, A. Trocha, E. Cutrell,

- N. Frahm, C. Brander, I. Toth, E. J. Arts, T. M. Allen, & B. D. Walker, 2009a. 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-lymphocyte recognition. *J Virol* **83**(6):2743–2755. On p. 11.
- [Miura *et al.*, 2009b] T. Miura, C. J. Brumme, M. A. Brockman, Z. L. Brumme, F. Pereyra, B. L. Block, A. Trocha, M. John, S. Mallal, P. R. Harrigan, & B. D. Walker, 2009b. HLA-associated viral mutations are common in human immunodeficiency virus type 1 elite controllers. *J Virol* **83**(7):3407–3412. On p. 10.
- [Moore *et al.*, 2002] C. B. Moore, M. John, I. R. James, F. T. Christiansen, C. S. Witt, & S. A. Mallal, 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* **296**(5572):1439–1443. On p. 5.
- [Neher & Leitner, 2010] R. A. Neher & T. Leitner, 2010. Recombination rate and selection strength in HIV intra-patient evolution. *PLoS Comput Biol* **6**(1):e1000660. On pp. 3 & 5.
- [Palella *et al.*, 1998] F. J. Palella, K. M. Delaney, A. C. Moorman, M. O. Loveless, J. Fuhrer, G. A. Satten, D. J. Aschman, & S. D. Holmberg, 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N Engl J Med* **338**(13):853–860. On p. 3.
- [Phillips *et al.*, 1991] R. E. Phillips, S. Rowland-Jones, D. F. Nixon, F. M. Gotch, J. P. Edwards, A. O. Ogunlesi, J. G. Elvin, J. A. Rothbard, C. R. M. Bangham, C. R. R. Rizza, & A. J. McMichael, 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* **354**(6353):453–459. On p. 4.
- [Poon *et al.*, 2007] A. F. Y. Poon, S. L. Kosakovsky Pond, P. Bennett, D. D. Richman, A. J. Leigh Brown, & S. D. W. Frost, 2007. Adaptation to human populations is revealed by within-host polymorphisms in HIV-1 and hepatitis C virus. *PLoS Pathog* **3**(3):e45. On pp. 5 & 6.
- [Poon *et al.*, 2008] A. F. Y. Poon, F. I. Lewis, S. D. W. Frost, & S. L. Kosakovsky Pond, 2008. Spidermonkey: Rapid detection of co-evolving sites using Bayesian graphical models. *Bioinformatics* **24**(17):1949–1950. On pp. 6 & 9.
- [Poon *et al.*, 2010] A. F. Y. Poon, L. C. Swenson, W. W. Y. Dong, W. Deng, S. L. Kosakovsky Pond, Z. L. Brumme, J. I. Mullins, D. D. Richman, P. R. Harrigan, & S. D. W. Frost, 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**(4):819–32. On pp. 6 & 11.
- [Quiñones-Mateu *et al.*, 2000] M. E. Quiñones-Mateu, S. C. Ball, A. J. Marozsan, V. S. Torre, J. L. Albright, G. Vanham, G. van Der Groen, R. L. Colebunders, & E. J. Arts, 2000. A dual infection/competition assay shows a correlation between ex vivo human immunodeficiency virus type 1 fitness and disease progression. *J Virol* **74**(19):9222–9233. On p. 6.
- [Rousseau *et al.*, 2008] C. M. Rousseau, M. G. Daniels, J. M. Carlson, C. Kadie, H. Crawford, A. Prendergast, P. Matthews, R. Payne, M. Roland, D. N. Raugi, B. S. Maust, G. H. Learn, D. C. Nickle, H. Coovadia, T. Ndung'u, N. Frahm, C. Brander, B. D. Walker, P. J. R. Goulder, T. Bhattacharya, D. E. Heckerman, B. T. Korber, & J. I. Mullins, 2008. HLA class I-driven evolution of human immunodeficiency virus type 1 subtype C proteome: Immune escape and viral load. *J Virol* **82**(13):6434–6446. On pp. 4, 5 & 10.
- [Salazar-Gonzalez *et al.*, 2009] J. F. Salazar-Gonzalez, M. G. Salazar, B. F. Keele, G. H. Learn, E. E. Giorgi, H. Li, J. M. Decker, S. Wang, J. Baalwa, M. H. Kraus, N. F. Parrish, K. S. Shaw, M. B. Guffey, K. J. Bar, K. L. Davis, C. Ochsenbauer-Jambor, J. C. Kappes, M. S. Saag, M. S. Cohen, J. Mulenga, C. A. Derdeyn, S. Allen, E. Hunter, M. Markowitz, P. Hraber, A. S. Perelson, T. Bhattacharya, B. F. Haynes, B. T. Korber, B. H. Hahn, & G. M. Shaw, 2009. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J Exp Med* **206**(6):1273–1289. On pp. 3 & 11.
- [Schellens *et al.*, 2009] I. M. Schellens, M. Navis, H. W. van Deutekom, B. Boeser-Nunnink, B. Berkhout, N. Kootstra, F. Miedema, C. Kesmir, H. Schuitemaker, D. van Baarle, & J. A. Borghans, 2009. Adaptation of HIV-1 to the human immune system at the population level is driven by protective HLA-B alleles. *Retrovirology* **6**(S3):O41. On p. 10.
- [Schmitz *et al.*, 1999] J. E. Schmitz, M. J. Kuroda, S. Santra, V. G. Sasseville, M. A. Simon, M. A. Lifton, P. Racz, K. Tenner-Racz, M. Dale-sandro, B. J. Scallan, J. Ghayeb, M. A. Forman, D. C. Montefiori, E. P. Rieber, N. L. Letvin, & K. A. Reimann, 1999. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science* **283**(5403):857–860. On p. 4.
- [Schneidewind *et al.*, 2008] A. Schneidewind, M. A. Brockman, J. Sidney, Y. E. Wang, H. Chen, T. J. Suscovich, B. Li, R. I. Adam, R. L. All-gaier, B. R. Mothé, T. Kuntzen, C. Oniangue-Ndza, A. Trocha, X. G. Yu, C. Brander, A. Sette, B. D. Walker, & T. M. Allen, 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**(11):5594–5605. On p. 6.
- [Schneidewind *et al.*, 2007] A. Schneidewind, M. A. Brockman, R. Yang, R. I. Adam, B. Li, S. Le Gall, C. R. Rinaldo, S. L. Craggs, R. L. Allgaier, K. A. Power, T. Kuntzen, C.-S. Tung, M. X. LaBute, S. M. Mueller, T. Harrer, A. J. McMichael, P. J. R. Goulder, C. Aiken, C. Brander, A. D. Kelleher, & T. M. Allen, 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**(22):12382–12393. On pp. 4, 6 & 10.
- [Schneidewind *et al.*, 2009] A. Schneidewind, Z. L. Brumme, C. J. Brumme, K. A. Power, L. L. Reyor, K. O'Sullivan, A. Gladden, U. Hempel, T. Kuntzen, Y. E. Wang, C. Oniangue-Ndza, H. Jessen, M. Markowitz, E. S. Rosenberg, R.-P. Sékaly, A. D. Kelleher, B. D. Walker, & T. M. Allen, 2009. Transmission and long-term stability of compensated CD8 escape mutations. *J Virol* **83**(8):3993–3997. On pp. 4, 10 & 11.
- [Shankarappa *et al.*, 1999] R. Shankarappa, J. B. Margolick, S. J. Gange, A. G. Rodrigo, D. Upchurch, H. Farzadegan, P. Gupta, C. R. Rinaldo, G. H. Learn, X. He, X.-L. Huang, & J. I. Mullins, 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* **73**(12):10489–10502. On p. 3.
- [Taylor *et al.*, 2008] B. S. Taylor, M. E. Sobieszczyk, F. E. McCutchan, & S. M. Hammer, 2008. The challenge of HIV-1 subtype diversity. *N Engl J Med* **358**(15):1590–1602. On p. 3.
- [Thananchai *et al.*, 2009] H. Thananchai, T. Makadzange, K. Maenaka, K. Kuroki, Y. Peng, C. Conlon, S. Rowland-Jones, & T. Dong, 2009. Reciprocal recognition of an HLA-Cw4-restricted HIV-1 gp120 epitope by CD8+ T cells and NK cells. *AIDS* **23**(2):189–193. On p. 3.
- [Trachtenberg *et al.*, 2003] E. Trachtenberg, B. Korber, C. Sollars, T. B. Kepler, P. T. Hraber, E. Hayes, R. Funkhouser, M. Fugate, J. Theiler, Y. S. Hsu, K. Kunstman, S. Wu, J. Phair, H. Erlich, & S. Wolinsky, 2003. Advantage of rare HLA supertype in HIV disease progression. *Nat Med* **9**(7):928–835. On p. 4.
- [Troyer *et al.*, 2009] R. M. Troyer, J. McNevin, Y. Liu, S. C. Zhang, R. W. Krizan, A. Abraha, D. M. Tebit, H. Zhao, S. Avila, M. A. Lobritz, M. J. McElrath, S. Le Gall, J. I. Mullins, & E. J. Arts, 2009. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS Pathog* **5**(4):e1000365. On p. 6.

[Turnbull *et al.*, 2009] E. L. Turnbull, M. Wong, S. Wang, X. Wei, N. A. Jones, K. E. Conrod, D. Aldam, J. Turner, P. Pellegrino, B. F. Keele, I. Williams, G. M. Shaw, & P. Borrow, 2009. Kinetics of expansion of epitope-specific T cell responses during primary HIV-1 infection. *J Immunol* **182**(11):7131–7145. On p. 4.

[UNAIDS & WHO, 2009] UNAIDS & WHO, 2009. 2009 AIDS epidemic update. On p. 3.

[Wang *et al.*, 2009] Y. E. Wang, B. Li, J. M. Carlson, H. Streeck, A. D. Gladden, R. Goodman, A. Schneidewind, K. A. Power, I. Toth, N. Frahm, G. Alter, C. Brander, M. Carrington, B. D. Walker, M. Altfeld, D. Heckerman, & T. M. Allen, 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**(4):1845–1855. On pp. 4, 5 & 10.

[Wei *et al.*, 2003] X. Wei, J. M. Decker, S. Wang, H. Hui, J. C. Kappes, X. Wu, J. F. Salazar-Gonzalez, M. G. Salazar, J. M. Kilby, M. S. Saag, N. L. Komarova, M. A. Nowak, B. H. Hahn, P. D. Kwong, & G. M. Shaw, 2003. Antibody neutralization and escape by HIV-1. *Nature* **422**(6929):307–312. On p. 3.

[Wood *et al.*, 2009] N. Wood, T. Bhattacharya, B. F. Keele, E. Giorgi, M. Liu, B. Gaschen, M. Daniels, G. Ferrari, B. F. Haynes, A. McMichael, G. M. Shaw, B. H. Hahn, B. Korber, & C. Seoighe, 2009. HIV evolution in early infection: Selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS Pathog* **5**(5):e1000414. On p. 3.

[Worobey *et al.*, 2008] M. Worobey, M. Gemmel, D. E. Teuwen, T. Haselkorn, K. Kunstman, M. Bunce, J.-J. Muyembe, J.-M. M. Kabongo, R. M. Kalengayi, E. Van Marck, M. T. P. Gilbert, & S. M. Wolinsky, 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* **455**(7213):661–664. On p. 3.

[Yokomaku *et al.*, 2004] Y. Yokomaku, H. Miura, H. Tomiyama, A. Kawana-Tachikawa, M. Takiguchi, A. Kojima, Y. Nagai, A. Iwamoto, Z. Matsuda, & K. Ariyoshi, 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**(3):1324–1332. On p. 4.

Acknowledgements

The authors wish to thank Richard Harrigan, Simon Malal and Mina John for data access, Chanson Brumme for assistance with data retrieval and formatting, and Mark Brockman and Mina John for critical review of the manuscript. ZLB is supported by a New Investigator Award from the Canadian Institutes for Health Research (CIHR). AFYP is supported by CIHR Fellowships Award 200802HFE.