2003; Bright et al., 2004; Woollard et al., 2008; Weatherford et al., 2009). A recent report indicated that marmosets exhibit susceptible and partially resistant phenotypes upon infection with GBV-B (Weatherford et al., 2009). Consistent with this finding, the present results also showed that the marmosets appeared to exhibit two phenotypes (**Figure 1B**). Importantly, the long-term persistent GBV-B infection was established in the marmosets with lower viral loads during the initial weeks p.i. (**Figure 1B**; Cj05-002 and Cj05-004). This suggests that the mild viral growth in the marmosets with a "partially resistant" phenotype is critical for the establishment of the chronic infection. Of note, the viral growth was undetectable until week 6 p.i. in Cj05-002, owing to unexpected interferon responses that were induced by administration of an anti-luciferase small interfering RNA in a cationic liposome formulation 2 days before GBV-B infection (Yokota et al., 2007). Irrespective of the partial suppression of the viral growth, humoral immune responses were delayed and consequently the individual developed chronic infection. Taken together, it is reasonable to assume that the viral persistence in marmosets may be closely associated with inefficient antiviral immune responses that are elicited at the periods of the lower viral loads. Previously, we and others employed relatively higher amounts of GBV-B for challenge in tamarins and marmosets. This could result in greater viral loads in the acute phase than those in humans and chimpanzees infected with HCV, followed by induction of efficient protective immunity and acute clearance. To clarify the mechanisms by which chronic GBV-B infection is established, further characterization of the differences in innate and acquired antiviral immunity between individuals with acute clearance and chronic infection will be needed.

Accumulating evidence suggests that escape mutations occurring during the course of chronic HCV infection may lead to evasion of humoral and cellular antiviral immunity (Bowen and Walker, 2005a,b; Burke and Cox, 2010). Consistent with these observations, we found that GBV-B acquired multiple back or sequential non-synonymous mutations (e.g., G250V > A, S731L > S, E2346G > E in Cj05-002; and V254A > V, I285V > I, L495S > L, T735A > T, F2135L > F > S in Cj05-004) in the chronically infected marmosets. Highly selective non-synonymous mutations were identified especially in E1, but such mutations were rarely observed in core (**Figures 4** and **5**). Moreover, the non-synonymous mutations in the E1 and NS3 regions occurred throughout the observation periods in Cj05-004 with chronic GBV-B infection, which had not been identified previously (Simons et al., 1995; Bukh et al., 1999; Sbardellati et al., 2001; Martin et al., 2003; Nam et al., 2004; Kyuregyan et al., 2005; Weatherford et al., 2009; Takikawa et al., 2010). Together with the finding that the rates of both synonymous and non-synonymous mutations were similar to those observed in cases of HCV (Ogata et al., 1991; Fernandez et al., 2004), these results strongly suggest that efficient and selective evasion from immune pressures may result in long-term persistent GBV-B infection and subsequent chronic hepatitis. Further analyses on the functional significance of the non-synonymous mutations will clarify this possibility.

It is surprising that in Cj05-004, the antibody titer to NS3 was observed to steadily increase after week 10 p.i. irrespective of the scarce viral loads over 1 year p.i., including the bipartite

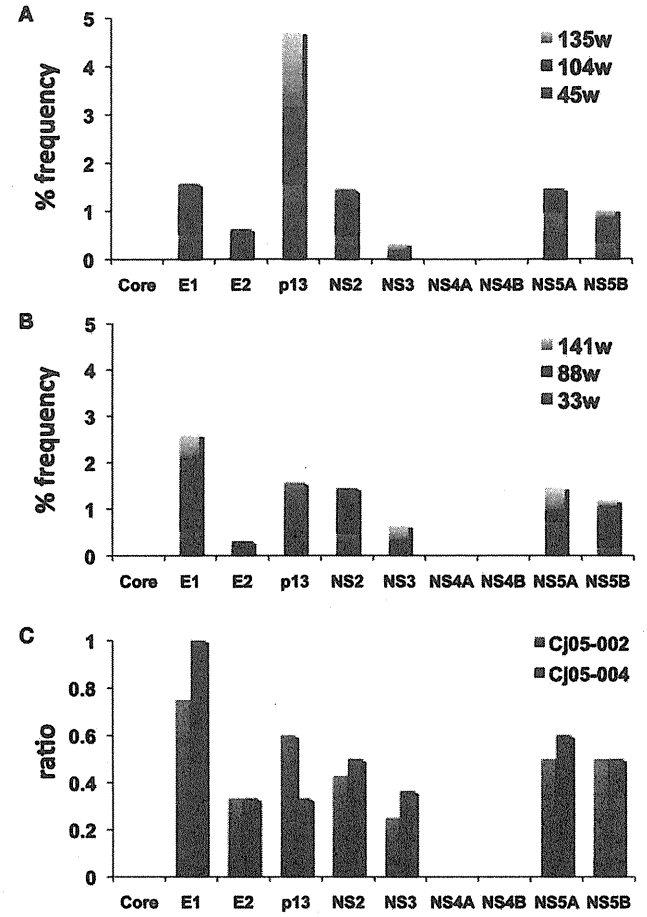


FIGURE 5 | The frequency of the non-synonymous mutations in the viral genome sequences amplified from plasma of the marmosets persistently infected with GBV-B. (A,B) The accumulated frequency of the non-synonymous mutations per total amino acid numbers of each viral protein at multiple time points are illustrated [**(A)** at weeks 45, 104, and 135 in Cj05-002; **(B)** weeks 33, 88, and 141 in Cj05-004]. To compare the frequency between the two marmosets, the data at week 229 p.i. in Cj05-004 are omitted. **(C)** The ratio of the non-synonymous mutations per total numbers of nucleotide mutations identified at weeks 45, 104, and 135 in Cj05-002 and at weeks 33, 88, and 141 in Cj05-004 in each viral gene. To compare the frequency between the two marmosets, the data at week 229 p.i. in Cj05-004 are omitted.

periods of weeks 4–26 and 34–58 p.i. when the virus was undetectable (**Figure 1**). Considering that three spikes of ALT levels were observed during these periods, our results suggest that antigenic stimulation by the lower level of viral growth in the liver, which was below detectable levels in blood, may induce the antibody and cytotoxic T-cell responses. In addition, during longitudinal analyses of monkeys experimentally infected with GBV-B, it is important to comprehensively evaluate multiple parameters, including viral loads, serum enzymes, and antibodies against core and NS3 proteins, to define whether virus-infected monkeys that produce no detectable viremia for a period of time have cleared the virus or are experiencing a latent period of chronic infection.

ACKNOWLEDGMENTS

We wish to thank T. Kurosawa, M. Fujita, and T. Ikoma for their helpful assistance and the members of Corporation for Production and Research of Laboratory Primates for technical assistance. Financial support: This work was supported by grants from the Ministry of Health, Labor, and Welfare of Japan (to Hirofumi Akari).

REFERENCES

Akari, H., Iwasaki, Y., Yoshida, T., and Iijima, S. (2009). Non-human primate surrogate model of hepatitis C virus infection. *Microbiol. Immunol.* 53, 53–57.

Beames, B., Chavez, D., Guerra, B., Notvall, L., Brasky, K. M., and Lanford, R. E. (2000). Development of a primary tamarin hepatocyte culture system for GB virus-B: a surrogate model for hepatitis C virus. *J. Virol.* 74, 11764–11772.

Beames, B., Chavez, D., and Lanford, R. E. (2001). GB virus B as a model for hepatitis C virus. *ILAR J.* 42, 152–160.

Boonstra, A., van der Laan, L. J., Vanwolleghem, T., and Janssen, H. L. (2009). Experimental models for hepatitis C viral infection. *Hepatology* 50, 1646–1655.

Bowen, D. G., and Walker, C. M. (2005a). Mutational escape from CD8+ T cell immunity: HCV evolution, from chimpanzees to man. *J. Exp. Med.* 201, 1709–1714.

Bowen, D. G., and Walker, C. M. (2005b). Adaptive immune responses in acute and chronic hepatitis C virus infection. *Nature* 436, 946–952.

Bright, H., Carroll, A. R., Watts, P. A., and Fenton, R. J. (2004). Development of a GB virus B marmoset model and its validation with a novel series of hepatitis C virus NS3 protease inhibitors. *J. Virol.* 78, 2062–2071.

Bukh, J. (2004). A critical role for chimpanzee model in the study of hepatitis C. *Hepatology* 39, 1469–1475.

Bukh, J., Apgar, C. L., and Yanagi, M. (1999). Toward a surrogate model for hepatitis C virus: An infectious molecular clone of the GB virus-B hepatitis agent. *Virology* 262, 470–478.

Bukh, J., Thimme, R., Meunier, J. C., Faulk, K., Spangenberg, H. C., Chang, K. M., Satterfield, W., Chisari, F. V., and Purcell, R. H. (2008). Previously infected chimpanzees are not consistently protected against reinfection or persistent infection after reexposure to the identical hepatitis C virus strain. *J. Virol.* 82, 8183–8195.

Burke, K. P., and Cox, A. L. (2010). Hepatitis C virus evasion of adaptive immune responses: a model for viral persistence. *Immunol. Res.* 47, 216–227.

Chisari, F. V. (2005). Unscrambling hepatitis C virus-host interactions. *Nature* 436, 930–932.

Feld, J. J., and Hoofnagle, J. H. (2005). Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature* 436, 967–972.

Fernandez, J., Taylor, D., Morhardt, D. R., Mihalik, K., Puig, M., Rice, C. M., Feinstone, S. M., and Major, M. E. (2004). Long-term persistence of infection in chimpanzees inoculated with an infectious hepatitis C virus clone is associated with a decrease in the viral amino acid substitution rate and low levels of heterogeneity. *J. Virol.* 78, 9782–9789.

Hoofnagle, J. H. (1997). Hepatitis C: the clinical spectrum of disease. *Hepatology* 26, 15S–20S.

Ishii, K., Iijima, S., Kimura, N., Lee, Y. J., Ageyama, N., Yagi, S., Yamaguchi, K., Maki, N., Mori, K., Yoshizaki, S, Machida, S., Suzuki, T., Iwata, N., Sata, T., Terao, K., Miyamura, T., and Akari, H. (2007). GBV-B as a pleiotropic virus: distribution of GBV-B in extrahepatic tissues in vivo. *Microbes Infect.* 9, 515–521.

Jacob, J. R., Lin, K. C., Tennant, B. C., and Mansfield, K. G. (2004). GB virus B infection of the common marmoset (*Callithrix jacchus*) and associated liver pathology. *J. Gen. Virol.* 85, 2525–2533.

Kyuregyan, K. K., Poleschuk, V. F., Zamyatina, N. A., Isaeva, O. V., Michailov, M. I., Ross, S., Bukh, J., Roggendorf, M., and Viazov, S. (2005). Acute GB virus B infection of marmosets is accompanied by mutations in the NS5A protein. *Virus Res.* 114, 154–157.

Lanford, R. E., Chavez, D., Notvall, L., and Brasky, K. M. (2003). Comparison of tamarins and marmosets as hosts for GBV-B infections and the effect of immunosuppression on duration of viremia. *Virology* 311, 72–80.

Lavanchy, D. (2009). The global burden of hepatitis C. *Liver Int.* 29, 74–81.

Martin, A., Bodola, F., Sanger, D. V., Goettge, K., Popov, V., Rijnbrand, R., Lanford, R. E., and Lemon, S. M. (2003). Chronic hepatitis associated with GB virus B persistence in a tamarin after intrahepatic inoculation of synthetic viral RNA. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9962–9967.

Melnikova, I. (2008). Hepatitis C therapies. *Nat. Rev. Immunol.* 5, 799–800.

Miyanari, Y., Atsuzawa, K., Usuda, N., Watashi, K., Hishiki, T., Zayas, M., Bartenschlager, R., Wakita, T., Hijikata, M., and Shimotohno, K. (2007). The lipid droplet is an important organelle for hepatitis C virus production. *Nat. Cell Biol.* 9, 1089–1097.

Muerhoff, A. S., Leary, T. P., Simons, J. N., Pilot-Matias, T. J., Dawason, G. J., Erker, J. C., Chalmers, M. L., Schlauder, G. G., Desai, S. M., and Mushahwer, I. K. (1995). Genomic organization of GB viruses A and B: two new members of the Flaviviridae associated with GB agent hepatitis. *J. Virol.* 69, 5621–5630.

Nam, J. H., Faulk, K., Engle, R. E., Govindarajan, S., St. Claire, M., and Bukh, J. (2004). In vivo analysis of the 3′ untranslated region of GB virus B after in vitro mutagenesis of an infections cDNA clone: persistent infection in a transfected tamarin. *J. Virol.* 78, 9389–9399.

Ogata, N., Alter, H. J., Miller, R. H., and Purcell, R. H. (1991). Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proc. Natl. Acad. Sci. U.S.A.* 88, 3392–3396.

Ohba, K., Mizokami, M., Lau, J. Y., Orito, E., Ikeo, K., and Gojobori, T. (1996). Evolutionary relationship of hepatitis C, pesti-, flavi-, plantviruses, and newly discovered GB hepatitis agents. *FEBS Lett.* 378, 232–234.

Perrillo, R. P. (1997). The role of liver biopsy in hepatitis C. *Hepatology* 26, 57S–61S.

Rehermann, B., and Nascimbeni, M. (2005). Immunology of hepatitis B virus and hepatitis C virus infection. *Nat. Rev. Immunol.* 5, 215–229.

Sbardellati, A., Scarselli, E., Verschoor, E., De Tomassi, A., Lazzaro, D., and Traboni, C. (2001). Generation of infectious and transmissible virions from a GB virus B full-length consensus clone in tamarins. *J. Gen. Virol.* 82, 2437–2448.

Seeff, L. B., and Hoofnagle, J. H. (2002). National Institutes of Health Consensus Development Conference: management of hepatitis C: 2002. *Hepatology* 36, S1–S2.

Simons, J. N., Pilot-Matias, T. J., Leary, T. P., Dawason, G. J., Desai, S. M., Schlauder, G. G., Muerhoff, A. S., Erker, J. C., Buijk, S. L., Chalmers, M. L., Van Sant, C. L., and Mushahwar, I. K. (1995). Identification of two *Flavivirus*-like genomes in the GB hepatitis agent. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3401–3405.

Takikawa, S., Engle, R. E., Faulk, K. N., Emerson, S. U., Purcell, R. H., and Bukh, J. (2010). Molecular evolution of GB virus B hepatitis virus during acute resolving and persistent infections in experimentally infected tamarins. *J. Gen. Virol.* 91, 727–733.

Weatherford, T., Chavez, D., Brasky, K. M., and Lanford, R. E. (2009). The marmoset model of GB virus B infections: adaptation to host phenotypic variation. *J. Virol.* 83, 5806–5814.

Woollard, D. J., Haqshenas, G., Dong, X., Pratt, B. F., Kent, S. J., and Gowans, E. J. (2008). Virus-specific T-cell immunity correlates with control of GB virus B infection in marmosets. *J. Virol.* 82, 3054–3060.

Yokota, T., Iijima, S., Kubodera, T., Ishii, K., Katakai, Y., Ageyama, N., Chen, Y., Lee, Y. J., Unno, T., Nishina, K., Iwasaki, Y., Maki, N., Mizusawa, H., and Akari, H. (2007). Efficient regulation of viral replication by siRNA in a non-human primate surrogate model for hepatitis C. *Biochem. Biophys. Res. Commun.* 361, 294–300.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 October 2011; paper pending published: 31 October 2011; accepted: 15 November 2011; published online: 07 December 2011.

Citation: Iwasaki Y, Mori K-i, Ishii K, Maki N, Iijima S, Yoshida T, Okabayashi S, Katakai Y, Lee Y-J, Saito A, Fukai H, Kimura N, Ageyama N, Yoshizaki S, Suzuki T, Yasutomi Y, Miyamura T, Kannagi M and Akari H (2011) Long-term persistent GBV-B infection and development of a chronic and progressive hepatitis C-like disease in marmosets. Front. Microbio. 2:240. doi: 10.3389/fmicb.2011.00240

This article was submitted to Frontiers in Virology, a specialty of Frontiers in Microbiology.

APPENDIX

MATERIALS AND METHODS

Liver samples obtained by necropsy from the GBV-B-infected marmosets were histopathologically analyzed as described in Section "Materials and Methods." Elastica–van Gieson staining was performed to evaluate fibrosis according to a standard laboratory protocol. To detect CD3 and CD20 antigens, liver samples were fixed in 10% neutral buffered formalin and embedded in paraffin wax. Sections were deparaffinized by pretreatment with 0.5% periodic acid and then subjected to antigen retrieval with citric acid buffer and heating in an autoclave for 10 min at 121°C. Sections were then incubated free floating in the monoclonal antibody solution for CD20 (DAKO) and CD3 (DAKO) overnight at 4°C. Following brief washes with buffer, the sections were sequentially incubated with biotinylated goat anti-mouse IgG (1:400), followed by streptavidin–biotin–horseradish peroxidase complex (sABC kit; DAKO, Denmark). Immunoreactive elements were visualized by treating the sections with 3,3′-diaminobenzidine tetroxide (Dojin Kagaku, Japan). The sections were then counterstained with hematoxylin.

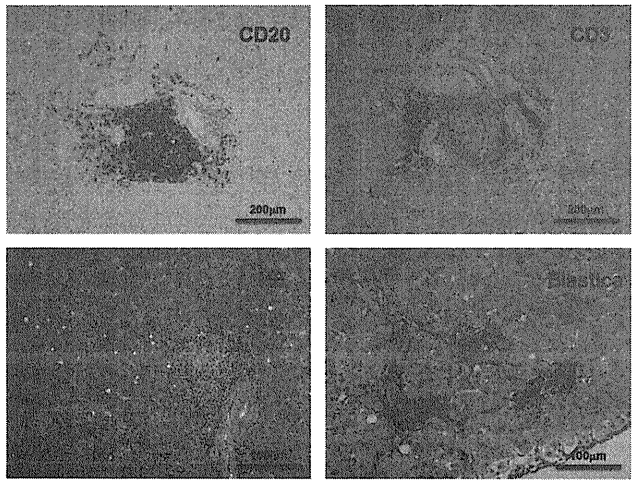


FIGURE A1 | Histopathological and immunohistochemical analyses of the liver from Cj05-004 at week 229 p.i.

OPEN ACCESS Freely available online

PLoS one

Unique CRF01_AE Gag CTL Epitopes Associated with Lower HIV-Viral Load and Delayed Disease Progression in a Cohort of HIV-Infected Thais

Masahiko Mori[1,2], Busarawan Sriwanthana[3], Nuanjun Wichukchinda[3], Chetsada Boonthimat[3], Naho Tsuchiya[1,4], Toshiyuki Miura[5], Panita Pathipvanich[6], Koya Ariyoshi[1,4]*, Pathom Sawanpanyalert[3]

1 Department of Clinical Medicine, Institute of Tropical Medicine (NEKKEN), Nagasaki University, Nagasaki, Japan, **2** Japan Foundation for AIDS Prevention (JFAP), Tokyo, Japan, **3** Thai National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand, **4** Global COE Program, Nagasaki University, Nagasaki, Japan, **5** Advanced Clinical Research Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan, **6** Day Care Center, Lampang Hospital, Lampang, Thailand

Abstract

Cytotoxic T Lymphocytes (CTLs) play a central role in controlling HIV-replication. Although numerous CTL epitopes have been described, most are in subtype B or C infection. Little is known about CTL responses in CRF01_AE infection. Gag CTL responses were investigated in a cohort of 137 treatment-naïve HIV-1 infected Thai patients with high CD4+ T cell counts, using gIFN Enzyme-Linked Immunospot (ELISpot) assays with 15-mer overlapping peptides (OLPs) derived from locally dominant CRF01_AE Gag sequences. 44 OLPs were recognized in 112 (81.8%) individuals. Both the breadth and magnitude of the CTL response, particularly against the p24 region, positively correlated with CD4+ T cell count and inversely correlated with HIV viral load. The breadth of OLP response was also associated with slower progression to antiretroviral therapy initiation. Statistical analysis and single peptide ELISpot assay identified at least 17 significant associations between reactive OLP and HLA in 12 OLP regions; 6 OLP-HLA associations (35.3%) were not compatible with previously reported CTL epitopes, suggesting that these contained new CTL Gag epitopes. A substantial proportion of CTL epitopes in CRF01_AE infection differ from subtype B or C. However, the pattern of protective CTL responses is similar; Gag CTL responses, particularly against p24, control viral replication and slow clinical progression.

Citation: Mori M, Sriwanthana B, Wichukchinda N, Boonthimat C, Tsuchiya N, et al. (2011) Unique CRF01_AE Gag CTL Epitopes Associated with Lower HIV-Viral Load and Delayed Disease Progression in a Cohort of HIV-Infected Thais. PLoS ONE 6(8): e22680. doi:10.1371/journal.pone.0022680

Editor: Johan K. Sandberg, Karolinska Institutet, Sweden

Received April 26, 2011; **Accepted** June 28, 2011; **Published** August 3, 2011

Funding: This study was mainly supported by the Japan Ministry of Health, Labour and Welfare; the Japan Foundation for AIDS Prevention and gCOE program, Nagasaki University; and partly by the Ministry of Public Health, Thailand. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: kari@nagasaki-u.ac.jp

Introduction

Cytotoxic T-Lymphocytes (CTLs) are an important component of the adaptive immune system which mediate control of HIV replication during acute infection and consequent viral set point [1]. Numerous CTL epitopes have been reported across the HIV proteome. However, the influence of CTL on clinical outcome varies, as their recognition of viral antigen is restricted by highly polymorphic class I Human Leukocyte Antigen (HLA) molecules [2,3]. Furthermore, the tremendous degree of viral diversity increases this complexity; to date, 13 prototype HIV clades and 43 circulating recombinant forms (CRF) have been described [4]. Some epitopes have been reported in a single clade; others have been reported in multiple clades (cross-clade) [5,6]. No reported epitope to date universally covers all HIV subtypes, or overcomes the global variation in HLA allele distribution (CTL Epitopes. Los Alamos National Lab. http://www.hiv.lanl.gov/).

Gag CTL responses, but not other CTL responses, have consistently been reported to have a significant association with viral control and clinical outcome [7]. However these findings were derived mainly from African or Caucasian populations infected with subtype C or B HIV, respectively; data from Asian populations infected with subtypes circulating in south-east Asia, such as CRF01_AE, have not yet been reported. To determine whether a similar association exists in south-east Asian subtypes, CTL epitope information is essential. However, as of April 2011, only 26 of 420 known Gag epitopes have been reported in CRF01_AE infection. Recently, the first successful phase III HIV vaccine trial was reported from Thailand [8], although its efficacy was marginal. For the development of a more effective vaccine, we believe it is crucial to accurately understand the influence of sequence variation amongst HIV subtypes, and HLA diversity amongst ethnic groups. To provide more information about CTL epitopes in CRF01_AE infection, we investigated cellular immune responses to Gag overlapping peptides in an HIV-1 CRF01_AE-infected Thai population and evaluated their impact on clinical outcome.

Methods

Subjects

This study was approved by the Thai Ministry of Public Health Ethics Committee and was conducted according to set guidelines for research. Written informed consent was obtained after

explaining the purpose and expected consequences of the study. Patients were eligible for inclusion if they were chronically HIV-infected and antiretroviral-naïve, with a CD4+ T cell count >200 cells/ul. A total of 137 HIV-1 CRF01_AE infected individuals were recruited at a government referral hospital in Thailand from October 2003 to May 2009. Study subjects were requested to visit the clinic every 3 months and CTL responses were evaluated every 6 months. The study endpoint was initiation of antiretroviral therapy, when their CD4+ T cell count declined below 200 cells/ul.

Synthetic HIV-1 Gag overlapping peptides

Fifteen-mer overlapping peptides (OLPs) of locally dominant CRF01_AE Gag sequences were designed based on 125 *gag* clonal sequences derived from 45 CRF01_AE infected individuals attending the clinic. All deduced amino-acid sequence data were aligned and the most frequent 15-mer amino-acid sequence was used as the dominant sequence.

Peptides were synthesized by Sigma Genosys (Hokkaido, Japan) with a high purity of >90% as determined by high-pressure liquid chromatography. In total, 98 peptides were synthesized and 20 pools were made by mixing 10 peptides per pool in a 10×10 matrix design so that a single responsible peptide could be identified by detecting the common peptide between two reactive pools, as described previously [9–11]. When more than one peptide was recognized, we further confirmed the responsible peptide recognition by individually testing candidate peptides, which were suspected by the matrix method.

ELISpot assay

1×10^5 fresh PBMC/well were plated onto multiScreen plates (MAHA54510; Millipore) that had been coated overnight at 4°C with 50 µl of anti-gIFN capture Ab 1-D1-K (2 µg/ml; Mabtech, Ohio, USA). Peptides were added directly to wells at a final concentration of 1 µM in 50 µl of R10 and incubated at 37°C in 5% CO_2 for 24 hrs. PBMC were stimulated with either medium alone for negative control, 10 µg/ml phytohemagglutinin (PHA; Sigma-Aldrich) for positive control or peptide (1 µM final concentration) for 24 hrs at 37°C. Plates were washed extensively with wash buffer (PBS/Tween20 0.001%), followed by incubation with biotinylated anti-human gIFN mAb (0.5 µg/ml; clone 7-B6-1; Mabtech) in PBS/10% FBS for 2 hrs at 37°C. Following six further washes with wash buffer, 2 µg/ml streptavidin HRP (Mabtech) was added to wells with 1 hr incubation at room temperature. Spots were visualized using BCIP/NBT substrate (Chemicon, Australia) and were counted using an Automated Enzyme-Linked Immunospot (ELISpot) Reader System with KS 4.3 software by an independent scientist in a blinded fashion. Each assay was undertaken in triplicate. Spot forming units (SFU) were counted and expressed as SFU per million PBMCs, using the average result from triplicate wells followed by subtraction of the negative control values. A response was defined as positive if it was three times higher than the negative control and greater than 150 SFU/1×10^6 PBMC. The breadth of response was defined as the total number of peptides recognized by each subject. The magnitude of response for an individual was defined as the sum of all positive peptide responses (in SFU/1×10^6 PBMC). To avoid overestimation of breadth or magnitude, two adjacent positive overlapping peptides were counted as one response, using the higher of the two responses.

HLA class I typing

Genomic DNA was extracted from buffy coat using the QIAamp DNA blood Mini Kit (Qiagen, Hilden, Germany) and 4-digit HLA class I typing for A, B and Cw loci was undertaken by bead-based array hybridization (WAKFlow HLA typing kit, Wakunaga Pharmaceutical, Hiroshima, Japan) according to manufacturer's instructions at a commercial laboratory (Kyoto HLA Laboratory, Kyoto, Japan).

Statistical analysis

Statistical analysis was performed using EXCEL 2007 and SPSS. We first selected viral loads (VL) in the lowest (=q1) and highest (=q4) quartiles (n=34 for each) and compared the number of individuals with positive ELISpot responses to p17, p24 and p15 proteins, using Fisher's exact test to compare groups. We then analyzed the association between breadth and clinical outcome (CD4+ T cell count and VL), using the Kruskal-Wallis test, and between magnitude and clinical outcome (CD4+ T cell count and VL) using Spearman's correlation test. We also performed a longitudinal analysis of the effect of breadth on Highly Active Anti-Retroviral Therapy (HAART) initiation, using the log rank test and Cox regression. For this analysis, the first individual was enrolled on 6 July 2000 and the last individual on 4 September 2007, with a censoring date of 31 May 2009. Analysis of OLP-HLA associations was undertaken using Fisher's exact test with 95% confidential intervals (CI). To have enough statistical power, we analyzed OLP-HLA associations when OLPs were recognized by 3 or more individuals with relevant HLA alleles and at least in one individual, the OLP recognition was confirmed by single peptide ELISpot experiments.

Results

Individuals' background, including HLA distribution

Of 137 individuals recruited, 107 were female and 30 were male. Median age was 31 years (range 16–56), CD4+ T cell count 461 cells/ul (range 204–1,191), and VL 4.22 log copies/ml (range 2.60–5.88). No individual had any HIV-related symptoms at the time of enrollment. In total, 87 variations of HLA alleles were found: 23 variations in HLA_A, 46 in HLA_B and 18 in HLA_Cw in four digits (Table S1). Median duration of follow-up was 22 months (range 0–60) and ELISpot experiments were repeated median 4 times (range 1–11) per individual. The peptide recognition pattern was confirmed to be consistent on at least two occasions for all except 24 individuals, in whom ELISpot assays were undertaken only once. During the follow-up period, the peptide recognition pattern did not change in any individual.

Gag OLP recognition and clinical outcome

Among 137 individuals, 112 (81.8%) recognized at least one OLP. Of 98 OLPs, 44 (44.9%) were recognized by at least one individual (Figure 1A): 12 peptides in p17, 26 in p24 and 6 in p15. The second half of p24 (HXB2 261–360; OLP 52–69), was the most highly targeted protein region; the first half of p17 (HXB2 5–60; OLP 1–9) was the second most highly targeted region. 14 OLPs were recognized in one individual and the other 30 OLPs were recognized in more than one individual. The most frequently recognized peptides were all located in the second half of p24: OLP 54 (HXB2 271–285), was recognized by 27 individuals; OLP 59 (HXB2 296–310) by 23 individuals; and OLP 66 (HXB2 331–345) by 22 individuals.

To further elucidate the peptide recognition pattern that best contributes to viral control, we next compared ELISpot responses between two extreme VL groups: the lowest quartile (=q1) (median VL 3.27 log copies/ml (range 2.60–3.71)) and the highest quartile (=q4) (median 5.09 log copies/ml (range 4.76–5.88)) (Figure 1B). Median CD4+ T cell count was 515 cells/ul (range

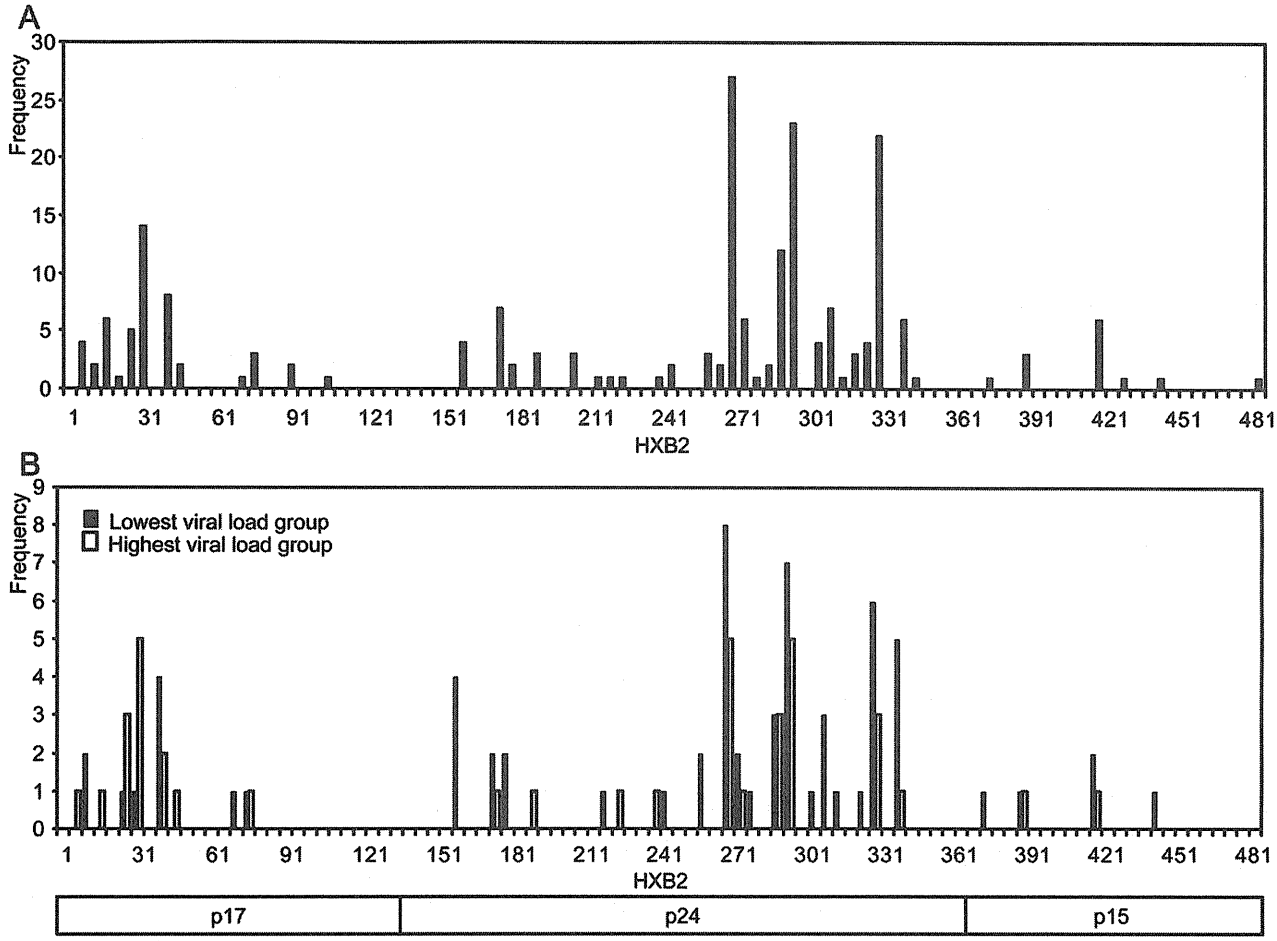


Figure 1. Pattern of CRF01_AE Gag CTL responses. Frequencies of overlapping peptide (OLP) responses in 112 individuals are shown (A); Frequencies of OLP responses in the lowest viral load group (lowest quartile, n = 34) and the highest viral load group (highest quartile, n = 34) were compared (B).
doi:10.1371/journal.pone.0022680.g001

243–1,057) in q1 and 429 cells/ul (range 204–856) in q4 ($p = 0.022$). Interestingly, individuals in q1 more frequently recognized p24 peptides than those in q4 (29/34 vs 18/34, respectively; $p = 0.0018$, Fisher's exact test), whereas individuals in q4 tended to recognize p17 peptides more frequently (9/34 vs 12/34, respectively; $p = 0.6$), although this difference was not significant.

ELISpot breadth, magnitude and clinical outcome

We next investigated the relationship between breadth and clinical outcome. The CD4+ T cell count was significantly higher in individuals with a greater breadth of response, with median CD4+ T cell count of 409 cells/ul (range 204–995), 455 cells/ul (range 243–793), 495 cells/ul (range 264–1,087) and 538 cells/ul (range 303–1,191) in individuals with 0, 1, 2 and ≥3 responses, respectively ($p = 0.018$ by Kruskal-Wallis test) (Figure 2A left). VL was significantly lower in individuals with a greater breadth of response, with median VL of 4.83 log copies/ml (range 2.60–5.88), 4.21 log copies/ml (range 2.60–5.83), 4.26 log copies/ml (range 2.76–5.71) and 3.82 log copies/ml (range 2.60–5.04) in individuals with 0, 1, 2 and ≥3 responses, respectively ($p = 0.0015$) (Figure 2A right). In a site-specific analysis, we did not find any significant association with CD4+ T cell count in any sites (Figure 2B). Interestingly, we found a significant association with VL only in p24 (4.57 log copies/ml (range 2.60–5.88), 4.21 log copies/ml (range 2.60–5.80), 4.17 log copies/ml (range 2.60–5.23) and 3.37 log copies/ml (range 2.60–4.14) in individuals with 0, 1, 2 and ≥3 responses, respectively; $p = 0.00028$) but not in other sites (Figure 2C).

We also found that magnitude of ELISpot response was positively correlated with CD4+ T cell count ($p = 0.0032$ by Spearman's correlation test $y = 0.031x+453$ $R^2 = 0.080$) and inversely correlated with VL ($p = 0.0084$ $y = -0.0001x+4.41$ $R^2 = 0.055$) (Figure 3A). In a detailed site-specific analysis, magnitude in p24 had a significant correlation with clinical outcome both in CD4+ T cell count ($p = 0.048$ $y = 0.013x+493$ $R^2 = 0.010$) (Figure 3B) and VL ($p = 0.0018$ $y = -0.0001x+4.39$ $R^2 = 0.065$) (Figure 3C), but not in other sites.

We next investigated the effect of breadth on clinical progression using the initiation of antiretroviral therapy as the end-point. During the follow-up period, 66/137 (48.2%) individuals started antiretroviral therapy. Intriguingly, we found that individuals with a wider breadth of CTL response were less likely to start antiretroviral therapy than those with a narrower breadth of response (Figure 4A, $p = 0.001$ by log rank test): 18/25 (72.0%), 13/34 (38.2%), 30/57 (52.6%) and 5/21(23.8%) individuals with 0, 1, 2 and ≥3 responses, respectively, initiating antiretroviral therapy. These data imply that strong CTL responses delay

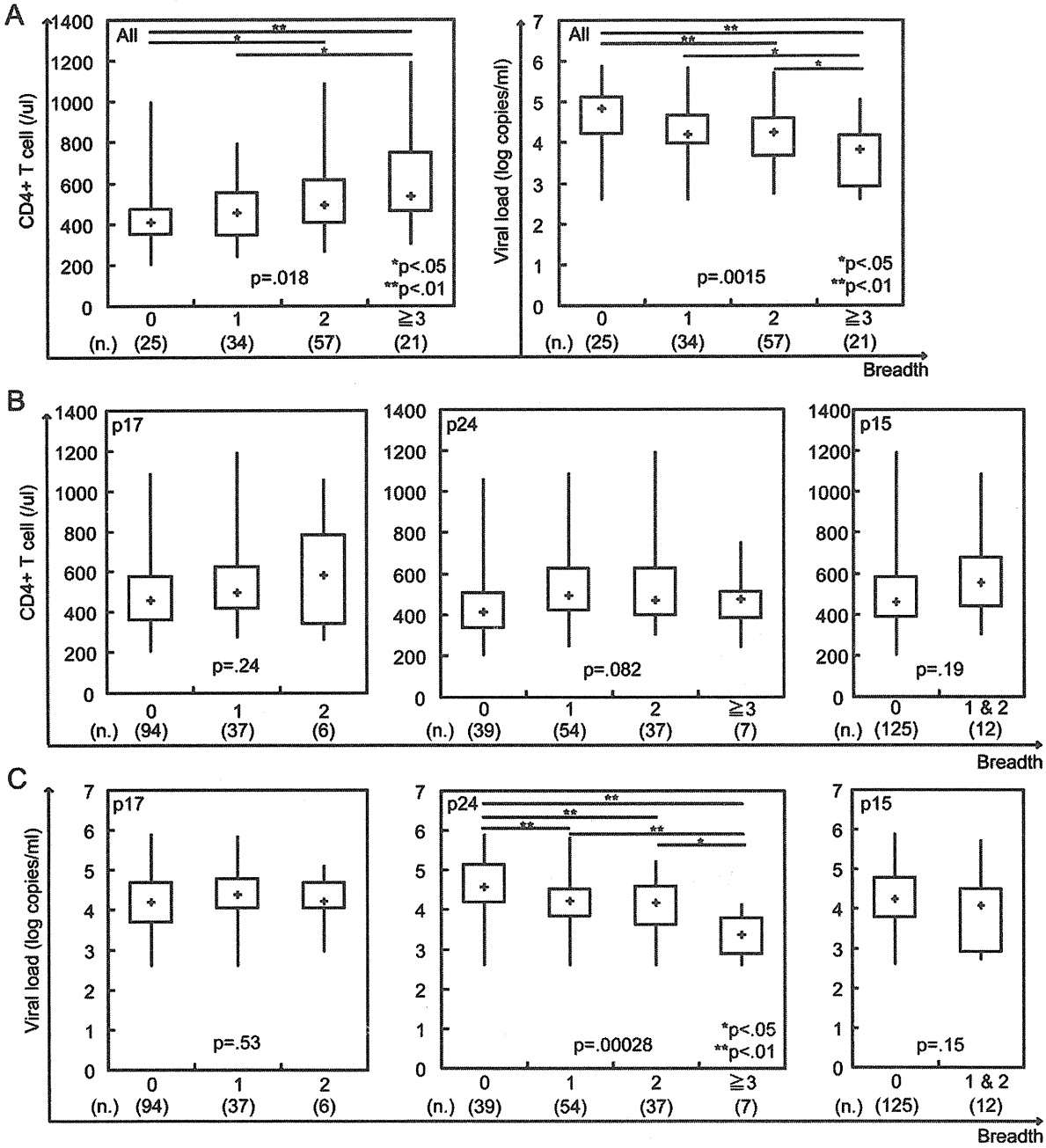


Figure 2. ELISpot breadth is associated with CD4+ T cell count and viral load. The associations between ELISpot breadth (the number of reacting OLP) and CD4+ T cell count or viral load were analyzed using the Kruskal-Wallis test (A). The p17, p24 or p15 site-specific ELISpot breadth was also compared with CD4+ T cell count (B) and viral load (C); * and ** showed a significant difference of $p<0.05$(*) and $p<0.01$(**) by Mann-Whitney u-test.

doi:10.1371/journal.pone.0022680.g002

clinical progression by slowing the decline in CD4+ T cell count. In a detailed site-specific analysis, individuals with a p24 response, but not other responses, were significantly less likely to start antiretroviral therapy than individuals without a p24 response ($p = 0.001$). However, the breadth of p24 response did not seem to correlate with clinical progression (Figure 4B).

Multivariate analysis of the relationship between CTL response and initiation of antiretroviral therapy, using Cox proportional hazards model, showed that the association between breadth of CTL response and initiation of HAART was independent of the baseline CD4+ T cell count (>350 cells/ul or not) and VL (<4.0 log copies/ml, 4.0–4.9 log copies/ml and ≥5.0 log copies/ml): adjusted Hazard Ratio (aHR) for individuals making ≥3 OLP responses was 0.23 ($p = 0.005$ with 95% CI of 0.08–0.64).

Detection of reactive OLP-HLA association

Associations between OLP responses and HLA were statistically analyzed. In total, 14 peptides (4 in p17, 9 in p24 and 1 in p15) with 31 OLP-HLA associations were identified (Table S2). 13 associations were found both with HLA-B and Cw alleles each and

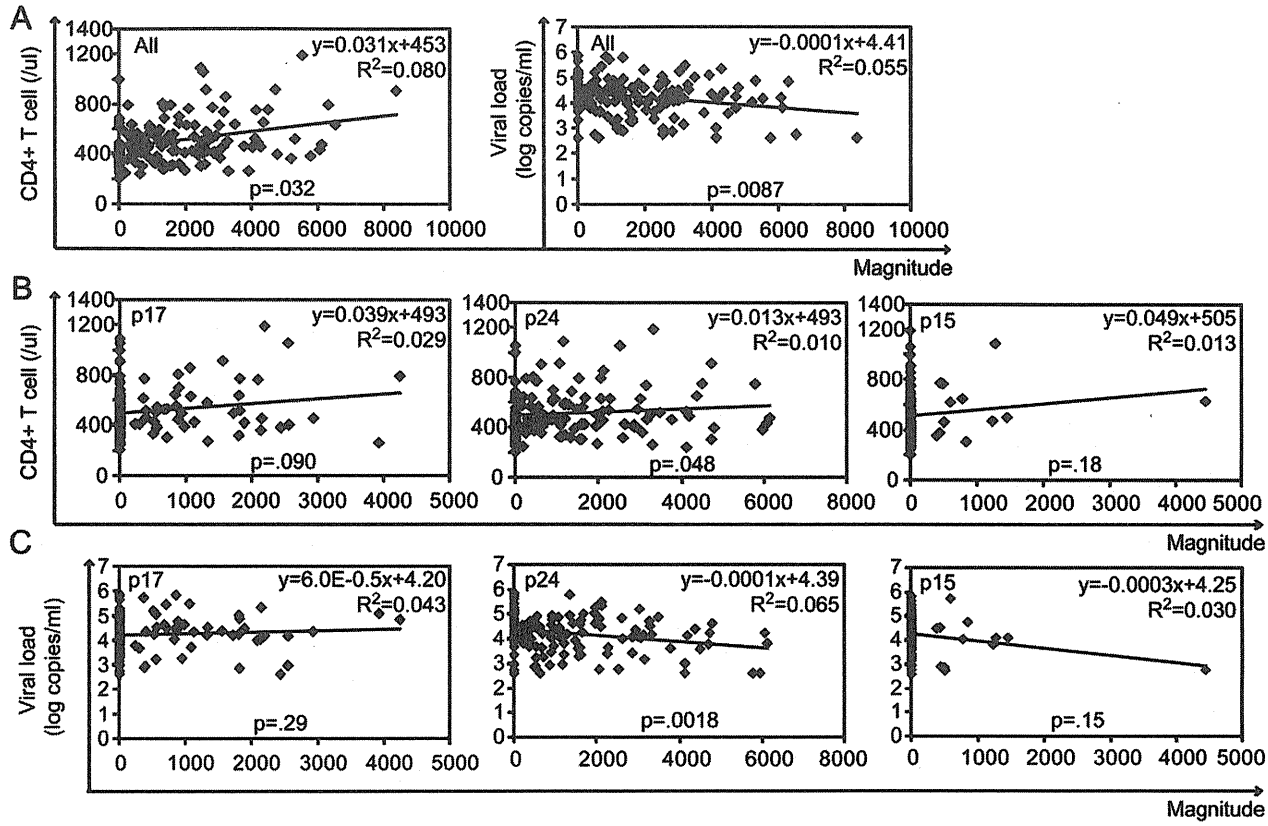


Figure 3. ELISpot magnitude is associated with CD4+ T cell count and viral load. The associations between ELISpot magnitude (total SFU per 1.0 M PBMC) and CD4+ T cell count or viral load were analyzed by Spearman's correlation (A). The p17, p24 or p15 site-specific response was also compared with CD4+ T cell count (B) and viral load (C).
doi:10.1371/journal.pone.0022680.g003

5 were found with HLA-A alleles. Two adjacent OLPs shared the same responsible HLA allele: HLA_A*0207, B*4601 and Cw*0102 in OLP 54–55, and B*4601 in OLP 58–59, suggesting that CTL epitopes reside in the overlapping region of these peptides. Some of the OLP-HLA associations may not reflect genuine CTL epitopes. 10 OLP responses were associated with two or more responsible HLA alleles. Of these, 9 OLP responses were associated with a pair of HLA alleles in linkage disequilibrium (LD), which were identified using the Los Alamos database (HLA Linkage Disequilibrium. Los Alamos National Lab. http://www.hiv.lanl.gov/). Among the 10 OLP responses, 7 included reported epitopes in either one of the HLA alleles. OLP 54, 55 and 59 responses were also associated with HLA alleles that have haplotype associations: HLA_A*0207-B*4601-Cw*0102. In total, 11 OLP-HLA associations were compatible with previously reported CTL epitopes: 4 epitopes were already reported as cross-clade epitopes including CRF01_AE or subtype A and the remaining 7 epitopes were reported in other subtypes but neither in subtype A nor CRF01_AE. Consequently, we identified at least 17 OLP-HLA associations in 12 OLP regions; 6 OLP-HLA associations (35.3%) were not compatible with previously reported CTL epitopes, suggesting that these are likely to contain unique CRF01_AE Gag CTL epitopes.

Discussion

This is the first study to investigate Gag CTL epitopes and their effect on clinical outcome in a systematic way in a CRF01_AE-infected Asian cohort. In this study, which tested optimal OLPs in a well-described cohort, we succeeded in predicting a number of unique CRF01_AE Gag epitope and novel cross-clade epitope candidates. Although one third of CTL epitope candidates in CRF01_AE infection were not compatible with previously reported CTL epitopes in other subtypes, both cross-sectional and longitudinal analysis showed the pattern of protective CTL responses was similar to previous studies; specifically, that a Gag CTL response, particularly against p24, was associated with better control of viral replication and slower clinical progression [7,11–15]. These findings are also compatible with our previous study in which an association with clinical outcome was found only for the number of HLA-associated mutations in p24 but not in other sites [16]. Both studies imply that immune pressure on p24 Gag influences the clinical outcome in CRF01_AE infected Asian individuals. Several papers have discussed the advantages of CTL immune pressure against p24 for viral control , which include selection of escape mutations that lead to viral fitness cost [17,18], sequence stability compared with other viral particles [4,19,20], the abundance of Gag protein in incoming virions [21], and more rapid antigen presentation of Gag epitopes following viral infection [18].

While our findings showed the clear-cut relationship between ELISpot breadth and clinical parameters, the slopes of the trend lines between ELISpot magnitude and clinical parameters were rather shallow. Furthermore ELISpot magnitude did not correlate with onset of HAART initiation. These findings are consistent with a recently published study that breadth of the CTL response rather than magnitude associated best with clinical outcome [22].

In this study, we could not detect any OLP-HLA associations in HLA_B*57, which is well-known as one of the most protective

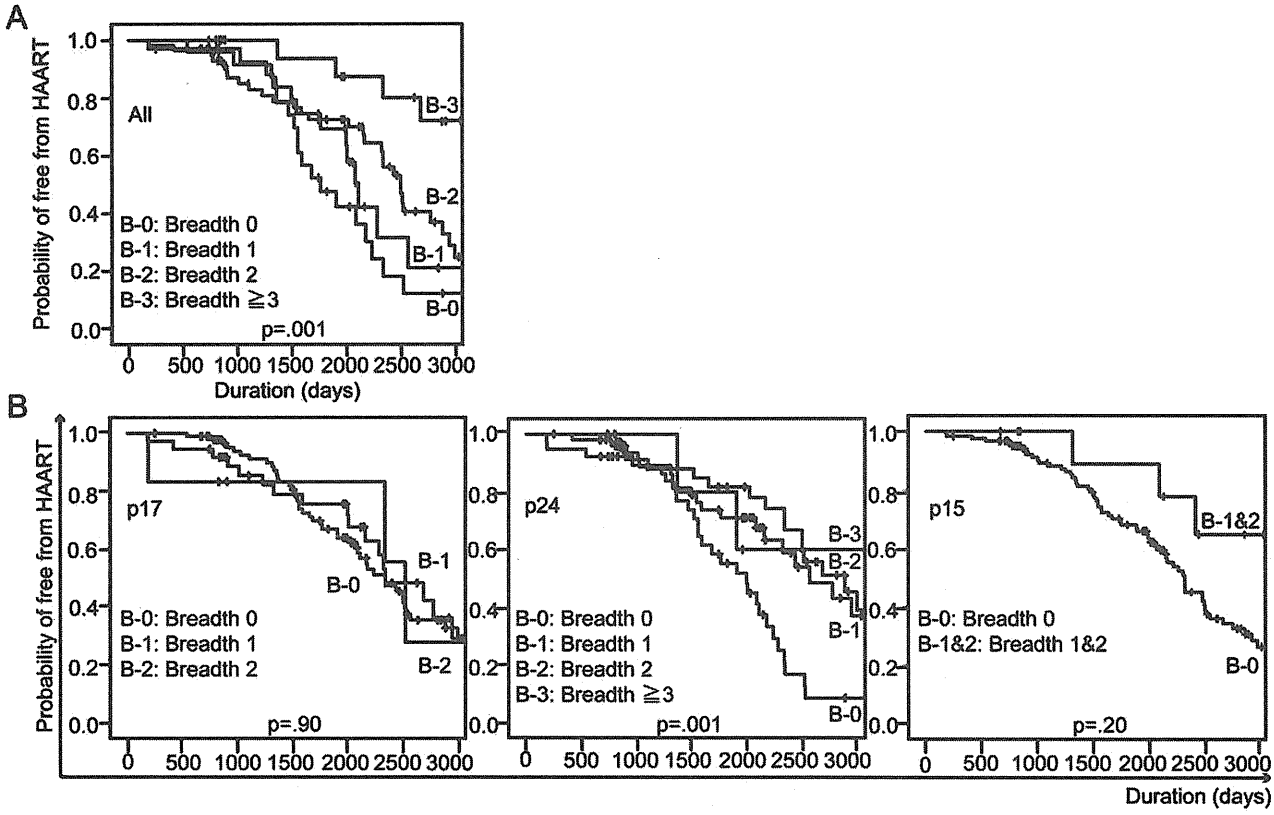


Figure 4. ELISpot breadth is related to delayed initiation of antiretroviral therapy. The impact of ELISpot breadth on antiretroviral therapy initiation was evaluated by Kaplan-Meier analysis, using the log rank test (A). The effect of p17, p24 or p15 site-specific ELISpot breadth was also analyzed (B).
doi:10.1371/journal.pone.0022680.g004

alleles for viral control [2,3,23]. Three individuals expressed B*5701; however, none had any response to OLP 47, which contains the TW10 (TSTLQEQIGW) epitope [24]. We have previously found in our cohort that all B*57 patients had the T242N escape mutation [16]. This suggests that the virus circulating in B*57 individuals lacks the wild-type TW10 sequence *in vivo* and no longer stimulates TW10 CTL cells [25].

In this study, OLP-HLA associations were predicted by statistical analysis. Thus these associations are not necessarily a reflection of new CTL epitopes with responsible HLA alleles. We excluded LD associations, including haplotypes and adjacent OLP responses with the same HLA allele association, in which CTL epitopes presumably reside in the overlapping region of these peptides. The most immunodominant OLP, number 54 (NKIVRMYSPVSILDI), was associated with three HLA alleles: A*0207, B*4601 and Cw*0102. "RMYSPVSIL" was previously identified as an A*0207-restricted CTL epitope [26]. All three responsible HLA alleles were found to be in LD. However, the association with B*4601 and Cw*0102 was much stronger than for A*0207 (odds ratio 29.4 in B*4601 and 104 in Cw*0102 vs 5.5 in A*0207) and further analysis including by ^{51}Cr release assay is warranted.

From this study, we have substantially increased information about CTL epitopes in CRF01_AE infection, reporting at least 6 unique CRF01_AE CTL epitope and 7 novel cross-clade epitope candidates. CRF01_AE is a recombinant HIV-1 with Gag derived from subtype A [4], from which CTL epitope information is limited, compared to subtypes B or C. We anticipate that if a more detailed epitope mapping study were to be conducted in subtype A-infected populations, there would be a large number of epitopes cross-recognized between CRF01_AE and subtype A.

Although details of OLP-HLA associations are substantially different between subtypes, interestingly we found a similarity in the immunodominant regions between subtypes. Our data showed that the second half of p24 was the most immunodominant regions, followed by the first half of p17 regions. This finding is consistent with previous reports [13,15,27]. We were concerned that the compatibility between OLP sequences and circulating Gag sequences may vary depending on the conservativeness and influence on the pattern of Gag CTL responses. However, the proportion of gag clones that were completely matched to the amino-acid sequence of OLPs was not associated with the frequency of OLP responses (data not shown).

Cross-clade CTL responses are said to be influenced by the viral sequence variability between subtypes, especially the sequence at anchor positions of the HLA binding motif [4,28–31]. Among the 7 newly identified cross-clade epitope candidates, 6 shared the same sequences with reported epitopes at both the B and F pockets. We also compared sequence compatibility at the anchor positions of the best-defined 12 epitopes, not identified in our study. 11 out of 12 also had compatible sequences at anchor positions, implying that sequence compatibility at anchor positions per se does not predict cross-clade reactivity. Other factors should be considered, such as sequences at flanking regions affecting peptide cleavage by the proteasome [32,33] and epitope-HLA complex recognition by T cell receptors (TCRs) [34,35].

This study has a number of limitations. First, we focused on Gag CTL immune responses and did not investigate whole viral proteins. However, since this type of analysis requires a large number of cells, and the volume of blood that we were able to take was rather limited, we decided to focus on Gag responses, as Gag is known to be the most important viral target. Instead of testing a large number of OLPs individually, we undertook experiments in triplicate, using a matrix system, to improve reliability. However, it would have been ideal if we had obtained enough volume of blood to confirm all responses using the individual peptides. Second, we detected OLP-HLA associations by a statistical method and not by the standard HLA-restriction analysis. This approach is easily influenced by sample size and the impact of LD. Thus our study does not provide direct evidence. Third, we have not yet confirmed these OLP responses with CTL using the ^{51}Cr release assay. However, ELISpot assays are now widely accepted as a technique for mapping CTL epitopes [36]. Fourth, these data are based on single cytokine release of gIFN; we did not evaluate multi-functionality of CTL with other cytokines such as IL2 or TNFα [37].

However, our data indicate the existence of a substantial number of unique CTL epitopes in CRF01_AE infection; it is therefore worth conducting a systematic analysis of CTL epitopes when vaccine trials are undertaken in different populations infected with different subtypes.

Supporting Information

Table S1 HLA allele frequencies in the study population.
(XLS)

Table S2 Gag overlapping peptide responses and their HLA allele associations.
(XLS)

Acknowledgments

We would like to thank Ms Bongkod Jitjuk, Ms Phattaraporn Orataiwun, Ms Suthira Kasemsuk, Ms Sripai Saneewong-na-Ayuthaya, Ms Katkaew Thamachai, Ms Anongnard Suyasarojna, Ms Nutira Boonna, and Mr Praphan Wongnamnong for their excellent technical assistance at Lampang Hospital.

Author Contributions

Conceived and designed the experiments: BS PS KA. Performed the experiments: NW CB MM. Analyzed the data: MM NT. Contributed reagents/materials/analysis tools: BS PP KA. Wrote the paper: MM KA. Clinical evaluation and patient recruitment: PP. Critical review: TM.

References

1. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF (2010) The immune response during acute HIV-1 infection: clues for vaccine development. Nat Rev Immunol 0: 11–23.
2. Kaslow RA, Carrington M, Apple R, Park L, Muñoz A, et al. (1996) Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. Nat Med 2: 405–11.
3. Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, et al. (2004) Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature 432: 769–75.
4. Buonaguro L, Tornesello ML, Buonaguro FM (2007) Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. J Virol 81: 10209–19.
5. Dorrell L, Willcox BE, Jones EY, Gillespie G, Njai H, et al. (2001) Cytotoxic T lymphocytes recognize structurally diverse, clade-specific and cross-reactive peptides in human immunodeficiency virus type-1 gag through HLA-B53. Eur J Immunol 31: 1747–56.
6. Ferrari G, Humphrey W, McElrath MJ, Excler JL, Duliege AM, et al. (1997) Clade B-based HIV-1 vaccines elicit cross-clade cytotoxic T lymphocyte reactivities in uninfected volunteers. Proc Natl Acad Sci USA 94: 1396–401.
7. Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, et al. (2007) CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat Med 13: 46–53.
8. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, et al. (2009) Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med 361: 2209–20.
9. Addo MM, Yu XG, Rathod A, Cohen D, Eldridge RL, et al. (2003) Comprehensive epitope analysis of human immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. J Virol 77: 2081–92.
10. Kern F, Surel IP, Faulhaber N, Frömmel C, Schneider-Mergener J, et al. (1999) Target structures of the CD8(+)-T-cell response to human cytomegalovirus: the 72-kilodalton major immediate-early protein revisited. J Virol 73: 8179–84.
11. Masemola A, Mashishi T, Khoury G, Mohube P, Mokgotho P, et al. (2004) Hierarchical targeting of subtype C human immunodeficiency virus type 1 proteins by CD8+ T cells: correlation with viral load. J Virol 78: 3233–43.
12. Brumme Z, Wang B, Nair K, Brumme C, de Pierres C, et al. (2009) Impact of select immunologic and virologic biomarkers on CD4 cell count decrease in patients with chronic HIV-1 subtype C infection: results from Sinikithemba Cohort, Durban, South Africa. Clin Infect Dis 49: 956–64.
13. Geldmacher C, Currier JR, Herrmann E, Haule A, Kuta E, et al. (2007) CD8 T-cell recognition of multiple epitopes within specific Gag regions is associated with maintenance of a low steady-state viremia in human immunodeficiency virus type 1-seropositive patients. J Virol 81: 2440–8.
14. Novitsky V, Gilbert P, Peter T, McLane MF, Gaolekwe S, et al. (2003) Association between virus-specific T-cell responses and plasma viral load in human immunodeficiency virus type 1 subtype C infection. J Virol 77: 882–90.
15. Zuñiga R, Lucchetti A, Galvan P, Sanchez S, Sanchez C, et al. (2006) Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. J Virol 80: 3122–5.
16. Gesprasert G, Wichukchinda N, Mori M, Shiino T, Auwanit W, 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.
17. Martinez-Picado J, Prado JG, Fry EE, Pfafferott K, Leslie A, et al. (2006) Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. J Virol 80: 3617–23.
18. Sacha JB, Chung C, Rakasz EG, Spencer SP, Jonas AK, et al. (2007) Gag-specific CD8+ T lymphocytes recognize infected cells before AIDS-virus integration and viral protein expression. J Immunol 178: 2746–54.
19. Goulder PJ, Watkins DI (2008) Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat Rev Immunol 8: 619–30.
20. Yusim K, Kesmir C, Gaschen B, Addo MM, Altfeld M, et al. (2002) Clustering patterns of cytotoxic T-lymphocyte epitopes in human immunodeficiency virus type 1 (HIV-1) proteins reveal imprints of immune evasion on HIV-1 global variation. J Virol 76: 8757–68.
21. Briggs JA, Simon MN, Gross I, Kräusslich HG, Fuller SD, et al. (2004) The stoichiometry of Gag protein in HIV-1. Nat Struct Mol Biol 11: 672–5.
22. Julg B, Williams KL, Reddy S, Bishop K, Qi Y, et al. (2010) Enhanced anti-HIV functional activity associated with Gag-specific CD8 T-cell responses. J Virol 84: 5540–9.
23. Kaslow RA, Dorak T, Tang JJ (2005) Influence of Host Genetic Variation on Susceptibility to HIV Type 1 Infection. J Infect Dis 191: S68–77.
24. Lieberman J, Fabry JA, Fong DM, Parkerson GR, 3rd (1997) Recognition of a small number of diverse epitopes dominates the cytotoxic T lymphocytes response to HIV type 1 in an infected individual. AIDS Res Hum Retroviruses 13: 383–92.
25. Miura T, Brockman MA, Schneidewind A, Lobritz M, Pereyra F, 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-lymphocyte [corrected] recognition. J Virol 83: 2743–55.
26. Seth A, Yasutomi Y, Jacoby H, Callery JC, Kaminsky SM, et al. (2000) Evaluation of a lipopeptide immunogen as a therapeutic in HIV type 1-seropositive individuals. AIDS Res Hum Retroviruses 16: 337–43.
27. Thakar MR, Bhonge LS, Lakhashe SK, Shankarkumar U, Sane SS, et al. (2005) Cytolytic T lymphocytes (CTLs) from HIV-1 subtype C-infected Indian patients recognize CTL epitopes from a conserved immunodominant region of HIV-1 Gag and Nef. J Infect Dis 192: 749–59.
28. Geels MJ, Dubey SA, Anderson K, Baan E, Bakker M, et al. (2005) Broad cross-clade T-cell responses to gag in individuals infected with human immunodeficiency virus type 1 non-B clades (A to G): importance of HLA anchor residue conservation. J Virol 79: 11247–58.
29. Gudmundsdotter L, Bernasconi D, Hejdeman B, Sandstrom E, Alaeus A, et al. (2008) Cross-clade immune responses to Gag p24 in patients infected with different HIV-1 subtypes and correlation with HLA class I and II alleles. Vaccine 26: 5182–7.

30. Malhotra U, Nolin J, Mullins JI, McElrath MJ (2007) Comprehensive epitope analysis of cross-clade Gag-specific T-cell responses in individuals with early HIV-1 infection in the US epidemic. Vaccine 25: 381–90.
31. Yu XG, Lichterfeld M, Perkins B, Kalife E, Mui S, et al. (2005) High degree of inter-clade cross-reactivity of HIV-1-specific T cell responses at the single peptide level. AIDS 19: 1449–56.
32. Yokomaku Y, Miura H, Tomiyama H, Kawana-Tachikawa A, Takiguchi M, 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: 1324–32.
33. Tenzer S, Wee E, Burgevin A, Stewart-Jones G, Friis L (2009) Antigen processing influences HIV-specific cytotoxic T lymphocyte immunodominance. Nat Immunol 10: 636–46.
34. Dong T, Stewart-Jones G, Chen N, Easterbrook P, Xu X, et al. (2004) HIV-specific cytotoxic T cells from long-term survivors select a unique T cell receptor. J Exp Med 200: 1547–57.
35. Meyer-Olson D, Brady KW, Bartman MT, O'Sullivan KM, Simons BC, et al. (2006) Fluctuations of functionally distinct CD8+ T-cell clonotypes demonstrate flexibility of the HIV-specific TCR repertoire. Blood 107: 2373–83.
36. Streeck H, Frahm N, Walker BD (2009) The role of IFN-gamma Elispot assay in HIV vaccine research. Nat Protoc 4: 461–9.
37. Betts MR, Nason MC, West SM, De Rosa SC, Migueles SA, et al. (2006) HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. Blood 107: 4781–9.

OPEN ACCESS Freely available online

PLoS one

Progression to AIDS in South Africa Is Associated with both Reverting and Compensatory Viral Mutations

Kuan-Hsiang Gary Huang[1,14☯], Dominique Goedhals[2,3☯], Jonathan M. Carlson[4], Mark A. Brockman[5,6,7], Swati Mishra[1], Zabrina L. Brumme[5,6,7], Stephen Hickling[1], Christopher S. W. Tang[1], Toshiyuki Miura[7,8], Chris Seebregts[9], David Heckerman[4], Thumbi Ndung'u[10], Bruce Walker[10,11,12], Paul Klenerman[1,15], Dewald Steyn[2], Philip Goulder[13], Rodney Phillips[1,14,15], Bloemfontein-Oxford Collaborative Group, Cloete van Vuuren,[2] John Frater[1,15]*

1 Nuffield Department of Clinical Medicine, Oxford University, Oxford, United Kingdom, **2** University of Free State, Bloemfontein, South Africa, **3** National Health Laboratory Services (NHLS), Johannesburg, South Africa, **4** eScience Group, Microsoft Research, Los Angeles, California, United States of America, **5** Simon Fraser University, Burnaby, British Columbia, Canada, **6** British Columbia Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada, **7** Ragon Institute of Massachusetts General Hospital, Harvard University and Massachusetts Institute of Technology (MIT), Boston, Massachusetts, United States of America, **8** Institute of Medical Science, University of Tokyo, Tokyo, Japan, **9** Biomedical Informatics Research Division, Medical Research Council, Cape Town, South Africa, **10** HIV Pathogenesis Programme, The Doris Duke Medical Research Institute, University of KwaZulu-Natal, Durban, South Africa, **11** Infectious Disease, Massachusetts General Hospital, Boston, Massachusetts, United States of America, **12** Division of AIDS, Harvard University, Boston, Massachusetts, United States of America, **13** Department of Paediatrics, University of Oxford, Oxford, United Kingdom, **14** The James Martin 21st Century School, Oxford, United Kingdom, **15** Oxford NIHR Biomedical Research Centre, Oxford, United Kingdom

Abstract

We lack the understanding of why HIV-infected individuals in South Africa progress to AIDS. We hypothesised that in end-stage disease there is a shifting dynamic between T cell imposed immunity and viral immune escape, which, through both compensatory and reverting viral mutations, results in increased viral fitness, elevated plasma viral loads and disease progression. We explored how T cell responses, viral adaptation and viral fitness inter-relate in South African cohorts recruited from Bloemfontein, the Free State (n = 278) and Durban, KwaZulu-Natal (n = 775). Immune responses were measured by γ-interferon ELISPOT assays. HLA-associated viral polymorphisms were determined using phylogenetically corrected techniques, and viral replication capacity (VRC) was measured by comparing the growth rate of gag-protease recombinant viruses against recombinant NL4-3 viruses. We report that in advanced disease (CD4 counts <100 cells/μl), T cell responses narrow, with a relative decline in Gag-directed responses ($p<0.0001$). This is associated with preserved selection pressure at specific viral amino acids (e.g., the T242N polymorphism within the HLA-B*57/5801 restricted TW10 epitope), but with reversion at other sites (e.g., the T186S polymorphism within the HLA-B*8101 restricted TL9 epitope), most notably in Gag and suggestive of "immune relaxation". The median VRC from patients with CD4 counts <100 cells/μl was higher than from patients with CD4 counts ≥500 cells/μl (91.15% versus 85.19%, $p = 0.0004$), potentially explaining the rise in viral load associated with disease progression. Mutations at HIV Gag T186S and T242N reduced VRC, however, in advanced disease only the T242N mutants demonstrated increasing VRC, and were associated with compensatory mutations ($p = 0.013$). These data provide novel insights into the mechanisms of HIV disease progression in South Africa. Restoration of fitness correlates with loss of viral control in late disease, with evidence for both preserved and relaxed selection pressure across the HIV genome. Interventions that maintain viral fitness costs could potentially slow progression.

Citation: Huang K-HG, Goedhals D, Carlson JM, Brockman MA, Mishra S, et al. (2011) Progression to AIDS in South Africa Is Associated with both Reverting and Compensatory Viral Mutations. PLoS ONE 6(4): e19018. doi:10.1371/journal.pone.0019018

Editor: Ashok Chauhan, University of South Carolina School of Medicine, United States of America

Received January 19, 2011; **Accepted** March 13, 2011; **Published** April 22, 2011

Funding: K-HGH is supported by the James Martin 21st Century School. JF is funded by the Medical Research Council. RP, PG and PK are supported by the Wellcome Trust. JF and PK are funded by NIHR Biomedical Research Centre Programme. ZLB is funded by the Canadian Institutes for Health Research (CIHR). The collection and analysis of 43 Free State sequences was funded by a grant from the ACACIA and Connectivity Africa programmes of the International Development Research Centre (IDRC), Canada. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: john.frater@ndm.ox.ac.uk

☯ These authors contributed equally to this work.

Introduction

With few exceptions, untreated individuals infected with Human Immunodeficiency Virus Type 1 (HIV-1) develop Acquired Immunodeficiency Syndrome (AIDS), associated with opportunistic infections, malignancies and, eventually, death. Some patients progress to AIDS quickly, whilst others maintain undetectable plasma viral loads without therapy and do not become unwell for many years. Deciphering the correlates of this heterogeneous protection is important, as there are implications for the design of vaccines and other interventions.

The pace of HIV disease progression is multifactorial - a mixture of host and pathogen genetics combined with factors such as the immune response and viral adaptation. In genome-wide

association studies a limited number of SNPs and alleles correlate with lower viral loads [1][2], and HLA Class I and the human MHC associate reproducibly [3]. The role of the cell-mediated immune system in HIV-associated disease has received much scrutiny, especially the effect of different HLA Class I alleles. Well-documented examples include the protection conferred by HLA B*57 and B*27 [4][5] in Caucasian individuals and HLA B*5801 and B*8101 in patients from South Africa [6].

What determines this differential HLA Class I effect is unclear. 'Beneficial' HLA Class I alleles may be associated with T cell clones with broader cross-reactivity to viral variants due to reduced thymic selection [7], and thus broader and more pervasive selection pressure. However, HIV is adept at adapting to selection pressures invoked by both antiretroviral drugs (ARVs) and the immune system in the forms of drug resistance [8][9] and immune escape mutations [10][11], respectively. The latter are widespread across the HIV-1 genome [12][13][14], and may influence outcome in individual patients [15] and across different populations [16].

Escape from an effective immune response is determined by the strength of the imposed selection pressure and may explain why the prevalence of HLA-associated polymorphisms is greater for HLA Class I alleles associated with protection [17]. Although the adapted virus maintains a fitness advantage in the presence of the selection pressure conferred by cytotoxic T cells (CTL), there may be a significant drop in replicative capacity compared to a wild-type virus in a selection-free environment [18][19]. We, and others, have previously hypothesised that immune escape may therefore result in the maintenance of relatively lower viral loads and clinical advantage [17] [20,21]. This is supported by high reversion rates of escape mutations selected by 'beneficial' HLA Class I alleles following transmission to HLA-mismatched recipients [22][23].

These interactions between HLA Class I imposed selection, viral escape and viral fitness are together likely to influence clinical progression, however the mechanisms that lead to progression to late disease and AIDS are poorly defined. We hypothesised that the nature of these interactions can be understood better by investigating patients with late-stage HIV infection to determine if, or how, CD8+ve T cells are maintaining selection pressure, and whether the virus shows different patterns of adaptation and fitness costs compared to patients with earlier infection. There are limited reports that, despite the loss of CD4+ve cells, CD8+ve T cells may still be functional in AIDS, although with varied avidity, less polyfunctionality and less differentiation [24], and targeting Env rather than Gag [25]. We proposed that if CD8 T cell pressure is 'relaxed' due to HIV-induced immunodeficiency this might facilitate reversion of costly escape mutations, leading to a restoration of viral fitness and the subsequent rise in viraemia seen in AIDS. Reversion of costly drug resistance mutations has been associated with a rise in viral load and clinical progression [26], and therefore a precedent exists to potentially explain the rise in viral load associated with the onset of AIDS. Alternatively, CTL pressure may be maintained, but the virus might develop secondary compensatory mutations, which restore the replicative cost of the initial escape mutation. Compensatory mutations have been reported in chronic infection [27,28], but whether they explain the AIDS-associated rise in viraemia is not known.

In this study we question whether progression to AIDS is associated with an increase in viral fitness, and whether this is related to changes in T cell imposed selection pressure. We start by comparing T cell ELISPOT responses in patients across different CD4 count strata, and then examine how the variation in these responses correlates with different patterns of selection across the HIV genome in patients with very low CD4 cell counts compared with less progressed infections. Finally, we measure the viral fitness of these variants to correlate the replicative cost of adaptation with outcome. We investigated 1053 untreated HIV-1 infected South African individuals from Bloemfontein and Durban, and found that viral fitness was greater in patients with advanced disease with examples of both viral compensatory mutations and of reduction in escape mutations, the latter possibly due to 'immune relaxation'. Despite the complex interplay between host and pathogen, these data support a key role for the restoration of viral fitness in association with AIDS progression and help to explain the rise in viraemia in terminal stages.

Methods

Ethics Statement

Full ethical approval was gained for the study of both cohorts. Written informed consent was provided by all study participants. The study ethics for the Bloemfontein cohort was approved by the University of the Free State (ETOVS 10/04 and ETOVS 206/05) and for the Durban cohort by the University of KwaZulu-Natal Review Board.

Study subjects

Participants from two South African cohorts were studied (total n = 1053) - the 'Bloemfontein' cohort from the Free State (n = 278) and the 'Durban' cohort from KwaZulu-Natal (n = 775). Both are established cohorts, described elsewhere ([29] [6]). In summary, each of the cohorts was comprised of antiretroviral naïve and chronically HIV-1 subtype C-infected adults from neighbouring provinces in the central east region of South Africa. Of these, 916 patients with data on CD4 cell count, HLA class I type and either HIV *gag*, *pol* or *nef* sequences were included in the analysis of HLA-linked polymorphisms, and comprised the 'total' cohort. Subsequent analyses involved stratification of the patients according to HLA type, plasma viral load and CD4 T cell count. In the CD4 T cell count stratification, patients were assigned to either "High CD4" (CD4 T cell counts >500 cells/μl, n = 299), or "Low CD4" (CD4 T cell counts <100 cells/μl, n = 196).

HLA typing

Participants' HLA Class I type was determined to the oligo-allelic level using Dynal RELITM Reverse Sequence-Specific Oligonucleotide kits for the HLA-A, -B and –C loci (Dynal Biotech). To obtain four-digit typing, Dynal Biotech Sequence-Specific priming kits were used, in conjunction with the Sequence-Specific Oligonucleotide type.

T Cell ELISPOT assays

Peptides, 18 amino acids in length (n = 410), overlapping by 10 amino acids, and spanning the entire expressed HIV genome, were synthesized based on the consensus of available C-clade sequences in 2001. These peptides were used in a 'mega-matrix' of 11–12 peptides per pool to test patient samples for HIV-specific T cell responses by interferon-gamma ELISPOT assay, as previously described [30]. Confirmation of recognised individual 18-mer peptides within a peptide pool was carried out in separate ELISPOT assays.

Sequencing of HIV *gag*, *pol* and *nef*

HIV *gag*, *pol* and *nef* were sequenced using previously described methods and primers [17]. In brief, viral RNA was extracted from plasma, reverse transcribed to cDNA and amplified using a nested polymerase chain reaction (PCR) protocol. Population sequencing

of the HIV pol gene was carried out using ABI Big Dye terminator sequencing kits (Applied Biosystems), according to manufacturer's instructions. Sequences were aligned manually using X11 and Se-Al software.

Identification of HLA class I linked polymorphisms

Amino acid polymorphisms that associated with host HLA class I alleles, were identified using previously described methods utilising 'phylogenetic dependency networks' [31]. Briefly, the analysis combines phylogenetic correction with a statistical model of evolution to evaluate associations between host HLA class I alleles and viral amino acid site-specific polymorphisms. The analysis adjusts for confounding factors including founder effects, linkage disequilibrium of host HLA types, co-variation in HIV codons and corrections for multiple statistical tests. The significance of an association is expressed as a 'q value', which estimates the false discovery rate for each p-value. In this study an association is considered significant for $q<0.2$. 'Escape' describes increased polymorphisms observed at a particular amino acid site in the presence of a specific HLA class I allele. 'Reversion' describes decreased polymorphisms observed at a particular amino acid site in the absence of a specific HLA class I allele. The strength of associations between HLA class I alleles and viral polymorphisms in the "total" cohort, "high CD4" group or "low CD4" group was derived using a logistic regression model that corrected for phylogeny, and is reported as a $\log_2$-adjusted odds ratio.

Identified polymorphisms were mapped to known cytotoxic T cell epitopes. Epitope maps were defined from experimental data from the Durban cohort [6], and from those listed in the A-list of Los Alamos Database (http://www.hiv.lanl.gov/content/immunology/pdf/2008/optimal_ctl_article.pdf). The epitope flanking region was also included in the analysis, defined according to the five amino acids neighbouring to the defined epitope in both C and N terminal directions.

Generation of chimeric NL4-3 viruses

Chimeric viruses were constructed by recombining a *gag-protease*-deleted pNL4-3 HIV plasmid (pNL4-3Δ*gag-protease*) with autologous *gag-protease* amplified from virus extracted from patient plasma, as described elsewhere [19,32,33,34].

In brief, to amplify the autologous HIV *gag-protease* viral RNA was extracted from plasma, reverse transcribed and amplified using specific primers in the first round of nested PCR using the Superscript III One-step RT-PCR kit with high fidelity Platinum *Taq* polymerase (Invitrogen) (forward primer: 5′ AAATCTCTAG-CAGTGGCGCCCGAACAG 3′, HXB2 nucleotides 623 to 649; reverse primer: 5′ TTTAACCCTGCTGGGTGTGGTATYCCT 3′, HXB2 nucleotides 2851 to 2825). The second round of the nested PCR was conducted using a 100-mer primer pair (forward primer: 5′ GACTCGGCTTGCTGAAGCGCGCACGGCAA-GAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAA-TTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGG 3′, HXB2 nucleotides 695 to 794; reverse primer: 5′ GGCCC-AATTTTTGAAATTTTTCCTTCCTTTTCCATTTCTGTAC-AAATTTCTACTAATGCTTTTATTTTTTCTTCTGTCAAT-GGCCATTGTTTAACTTTTG 3′, HXB2 nucleotides 2704 to 2605) that completely matched the pNL4-3 sequences using Platinum *Taq* polymerase (Invitrogen).

The pNL4-3 HIV plasmid had been mutated to contain unique restriction enzyme (*BstEII*) sites at the 5′ end of the *gag* gene and 45 nucleotides downstream from the 3′ end of the *protease* gene, as described elsewhere [19]. The *BstEII* enzyme digest results in deletion of HIV *gag-protease* and self ligation of the remaining plasmid (pNL4-3Δ*gag-protease*). For transfection, the pNL4-3Δ*gag-pro*tease was linearised by *BstEII* digestion for two hours at 60°C (10 U/μl of enzyme per 1000 μg/mL of plasmid). For each sample, 10 μg of the linearised pNL4-3Δ*gag-pro*tease was co-transfected with 5 μg of patient-derived *gag-protease* PCR product in 3.9×10^6 tat-driven GFP reporter T cells (GXR cells of CEM origin). After transfection by electroporation (250 V, 950 μF, for 30–40 msec), 10 μl polybrene (4 μg/μl) and 1×10^6 GXR cells were added and incubated at 37°C. Positive control samples were generated using co-transfection of pNL4-3Δ*gag-pro*tease with *gag-protease* amplified from pNL4-3, pHXB2 (subtype B), and pMJ4 (subtype C). After five days, GFP expression was measured every 1~2 days using a FACSCalibur flow cytometer (FACSCalibur; BD Biosciences, San Jose, CA). Supernatant was harvested when >15% of cells were positive for GFP, and stored at −80°C.

Measurement of viral replication capacity (VRC)

Viral titration was performed in 1×10^6 GXR cells to determine the multiplicity of infection (MOI). The VRC was assayed by adding recombinant virus supernatant to 1×10^6 GXR cells at an MOI of 0.003. The VRC was measured, at least in duplicate, using gating strategies as previously described [19]. The growth of virus was calculated from the slope derived from the natural log of the percentage of GFP positive cells between days 2 and day 7, as appropriate for an exponential growth curve. The final VRC of each variant is expressed as percentage growth rate of the variants compared to that of recombinant NL4-3 strains. In addition, for each VRC assay, viral supernatants were collected on day 7 for viral RNA extraction and sequencing to confirm the population sequences VRC measured.

Statistics

Student t-tests (two tailed) and Mann-Whitney tests (two tailed) were used to compare various clinical parameters and VRC measurements. The statistical analyses were carried out using Prism4 for Macintosh (GraphPad Software, version 4.0c). Associations between HLA Class I alleles and viral polymorphisms were determined using phylogenetically-corrected algorithms, as described above.

Results

Structure of cohorts

Patients were recruited from the Bloemfontein and Durban cohorts [29] [6]. The Bloemfontein cohort comprises 884 drug-naïve patients at all stages of HIV infection recruited through the South African government ARV program. Plasma samples were studied from 278 out of the 884 participants: 96 out of 110 patients in the high (CD4 count>500 cells/μl), 18 out of 308 in an intermediate (CD4 count = 201–400 cells/μl), and 164 of 183 in the low (CD4 count<100 cells/μl) strata. The samples were used for the HLA-association study and to make recombinant viral strains for the analysis of fitness. To increase the power of the HLA-association analysis we incorporated HIV viral sequences from the neighbouring Durban cohort, which comprised 775 adults, naive to antiretroviral therapy. Of these, 219 were asymptomatic women identified through antenatal clinics. The remaining 556 subjects were recruited from out-patient clinics. For all 775 patients host HLA class I alleles were genotyped. ELISPOT assays were carried out using overlapping peptides designed from the subtype C majority consensus. For 656 of these patients, sequences of either viral *gag*, *pol* and *nef* genes were available.

Breadth of ELISPOT responses narrows in advanced disease

In order to measure the T cell-imposed selection pressure acting on HIV at different stages of the infection, ELISPOT assays using overlapping peptides (OLPs) covering the entire HIV genome were analysed from 775 patients. Results were stratified according to CD4 cell count and are presented according to the number of OLPs seen per patient for each gene (Figure 1a), and the percentage contribution of each protein to all responses in each strata (Figure 1b). The figure shows that over the course of infection, there is a general narrowing of the absolute breadth of responses (Figure 1a). However, the relative contribution of Gag to the total responses decreased ($p<0.0001$, $r^2 = 0.055$) in contrast to the proportion of responses to other proteins such as Env ($p<0.0001$, $r^2 = 0.024$) and Vif ($p = 0.003$, $r^2 = 0.011$), which was greater in patients with lower CD4 cell counts. We hypothesized that the reduction in breadth of CD8+ve T cell responses might result in, and therefore correlate with, a decline in selection pressure, particularly in Gag in light of the relative shift away from Gag targeting. As a result, this 'immune relaxation' might allow the reversion of costly viral escape mutations back to a fitter wild type state. Accordingly, we next analysed HLA Class I-associated polymorphisms in the HIV-1 *gag*, *pol* and *nef* genes to identify any difference in the strength of selection pressure in patients with high and low CD4 cell counts.

Identification of HLA class I linked polymorphisms in HIV-1 *gag, pol and nef*

To identify evidence of differential HLA Class I imposed selection pressure in patients with high and low CD4 cell counts, we carried out a study to identify HLA Class I associations with particular amino acid polymorphisms in this cohort. We accounted for potential confounding effects such as phylogeny, founder effect and multiple comparisons by using phylogenetic dependency networks [31] - an approach which reports associations between specific viral polymorphisms and HLA Class I alleles according to p value, q value (a measure of false positive error) and a phylogenetically-adjusted odds ratio.

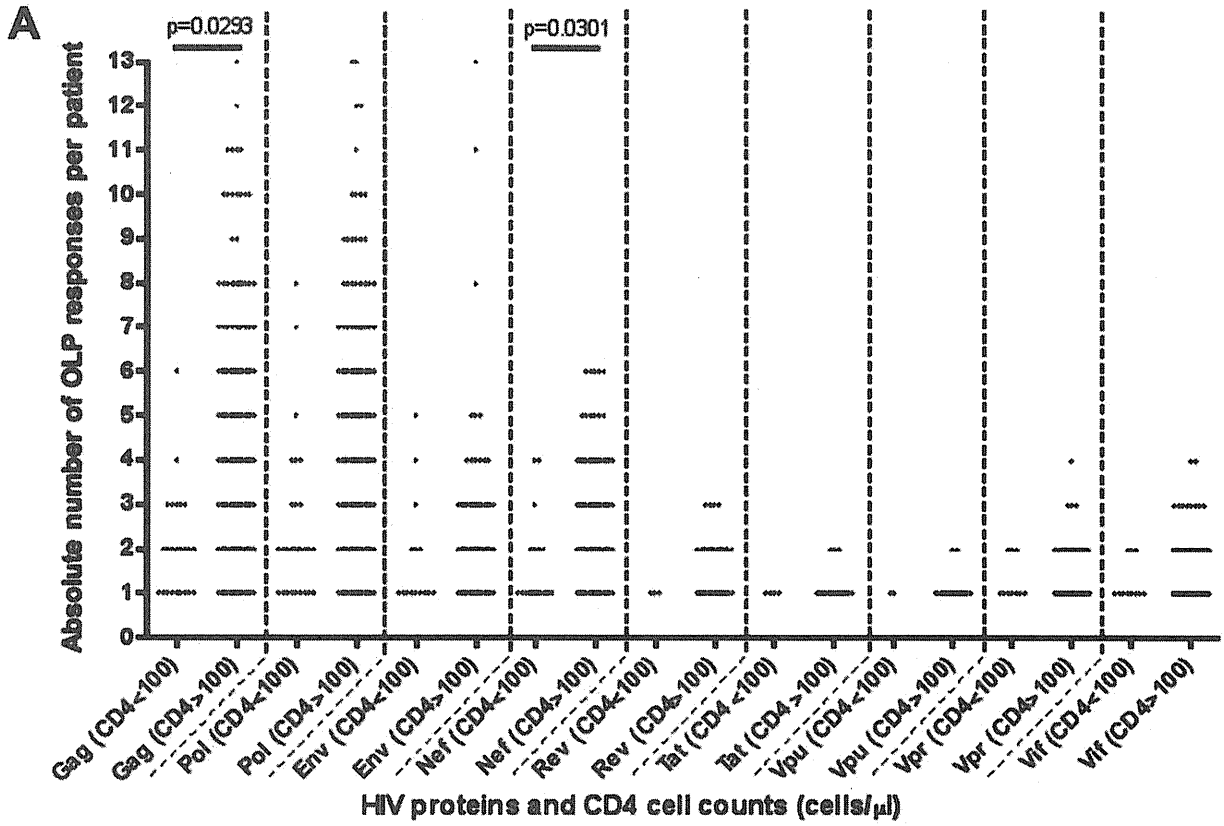
Initially, we analysed 916 HIV-1 *gag*, *pol* and *nef* sequences from the combined Durban and Bloemfontein cohorts and then we carried out sub-group analyses on patients with either 'high' (>500 cells/μl, n = 299) or 'low' (<100 cells/μl, n = 196) CD4 cell counts. We then selected those HIV amino acid polymorphisms significantly associated with an HLA Class I allele in either, or both of, the 'high' or 'low' CD4 count group of patients for further analysis. The results for *gag*, *pol* and *nef* are presented in Tables 1, 2 and 3, respectively. The tables detail each viral polymorphism and its associated HLA Class I allele, whether the association lies in, or within, the flanking region of a known restricted epitope and whether this association is expected to revert in HLA-mismatched hosts [35]. The q values for the statistically significant associations for the whole cohort, the 'high' CD4 count subgroup (>500 cells/μl) and 'low' CD4 count subgroup (<100 cells/μl) are shown, followed by the $\log_2$-adjusted odds ratios for the 'high' and 'low' groups. In the final column, the p value is reported, when there is a significant difference between the two odds ratios.

We identified 40, 91 and 20 associations, respectively, between an HLA Class I allele and a viral polymorphism in the HIV-1 *gag*, *pol* and *nef* genes, in either, or both of, the 'high' or 'low' CD4 count subgroups. Of these, 16, 28 and 7 associations in HIV-1 *gag*, *pol* and *nef*, respectively, had significantly different $\log_2$-adjusted odds ratios in the two subgroups. These 51 associations are shown in Figure 2, according to HIV-1 gene. In *gag*, *pol* and *nef*, respectively, there were 5/16 (31%), 24/28 (86%) and 7/7 (100%) associations that were more prevalent in the 'low CD4' group. These data show that, where we could detect a difference, associations were more prevalent at lower CD4 counts in *pol* and *nef*, suggestive of continued selection pressure, but that certain associations – particularly in HIV-1 *gag* – were weaker. Why might HLA associations become weaker as disease progresses? One possible explanation is that as CTL pressure weakens due to immune exhaustion and loss of breadth, escape mutations with replicative fitness costs are no longer required by the virus. These therefore revert back to wild-type, resulting in a reduction in the observed $\log_2$-adjusted odds ratio. Thus, low-cost mutations accumulate over the course of infection, leading to higher prevalence in Nef and Pol, whereas high-cost mutations begin reverting, leading to a lower prevalence in Gag.

If changes in selection pressure were resulting in an increase or decrease in the prevalence of wild type virus during late-stage disease, we would expect to find changes in viral replicative capacity in patients with low CD4 cell counts. We therefore undertook assays of viral fitness. As the HIV *gag* gene was associated with the most potential examples of immune relaxation we implemented a fitness assay targeting the *gag* gene and focussed on three beneficial HLA Class I alleles (HLA B*57, HLA B*5801, and HLA B*8101) and one allele associated with more rapid progression, HLA B*5802. In particular, we highlighted two associations: the T186S mutation in the TPQDLNTML (TL9) epitope, restricted by B*8101, with an observed decrease in $\log_2$-adjusted odds ratio from 18.38 to 7.25 in patients with high and low CD4 cell counts, respectively, (consistent with immune relaxation, $p = 0.0397$, likelihood ratio test), and the B*57/B*5801 restricted epitope TSTLQEQIAW (TW10), with $\log_2$-adjusted odds ratios of 25.4 and 23.2 for high and low CD4 cell count stratification analysis ($p = 0.12$, likelihood ratio test, showing no significant evidence of immune relaxation). We focussed on these two associations as published data suggest that they are both associated with significant viral fitness costs [32] [18,27] and yet in this dataset they behave very differently in terms of changes in $\log_2$-adjusted odds ratios.

Increased viral replication capacity (VRC) of chimeric HIV-1 NL4-3 at low CD4 counts and high plasma viral loads

To investigate further possible reasons for declining selection pressure within Gag epitopes with progression to AIDS, we considered the hypothesis that loss of CD8+ T-cell responses as CD4 count declines results in reversion of escape mutants, with resulting increase in viral replicative capacity. To investigate the impact of the *gag-protease* mutations selected by different HLA class I alleles at different clinical stages, we constructed 148 chimeric HIV-1 NL4-3 viruses containing autologous *gag-protease* from patients of the Bloemfontein cohort. Patients with specific HLA class I alleles previously associated with control of HIV in this population (HLA-B*57/5801/8101) or lack of control (B*5802) [6] were selected for chimeric virus construction and stratified by CD4 cell count. 88 viruses were constructed from patients with low CD4 cell counts (<100 cells/μl), comprising HLA-B*57 (n = 2), HLA-B*5801 (n = 15), HLA-B*5802 (n = 54) and HLA-B*8101 (n = 10) plus 14 'neutral/other' (i.e. neither protective or disadvantageous) HLA Class I alleles. From patients with high CD4 cell counts (>500 cells/μl), 42 viruses were made: HLA-B*57 (n = 7), HLA-B*5801 (n = 8), HLA-B*5802 (n = 15), HLA-B*8101 (n = 10) plus 10 'neutral/other' HLA Class I alleles. Sixteen patients co-expressed two of these HLA class I alleles. A further 18 viruses were made from patients randomly selected from intermediate CD4 T cell counts. The VRC of the chimeric virus


A
p=0.0293
p=0.0301
Absolute number of OLP responses per patient
0
1
2
3
4
5
6
7
8
9
10
11
12
13
Gag (CD4<100)
Gag (CD4>100)
Pol (CD4<100)
Pol (CD4>100)
Env (CD4<100)
Env (CD4>100)
Nef (CD4<100)
Nef (CD4>100)
Rev (CD4<100)
Rev (CD4>100)
Tat (CD4<100)
Tat (CD4>100)
Vpu (CD4<100)
Vpu (CD4>100)
Vpr (CD4<100)
Vpr (CD4>100)
Vif (CD4<100)
Vif (CD4>100)
HIV proteins and CD4 cell counts (cells/μl)

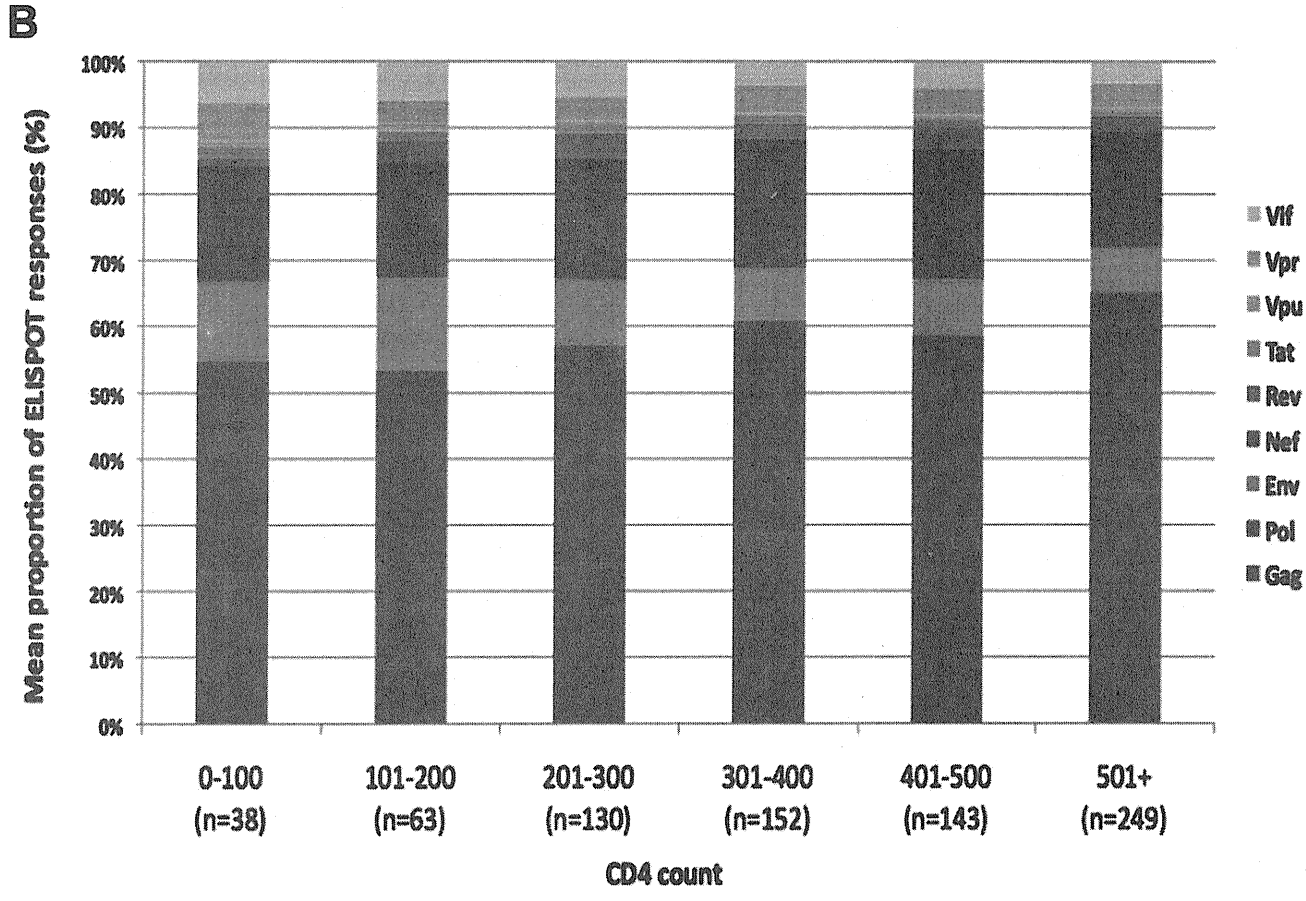

B
Mean proportion of ELISPOT responses (%)
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
0-100
(n=38)
101-200
(n=63)
201-300
(n=130)
301-400
(n=152)
401-500
(n=143)
501+
(n=249)
CD4 count
Vif
Vpr
Vpu
Tat
Rev
Nef
Env
Pol
Gag

Figure 1. Gamma interferon ELISPOT responses in patients with CD4 counts above and below 100 cells/μl. Recognition of overlapping peptides covering the full HIV genome is shown in patients with CD4 cell counts above and below 100 cells/μl. ELISPOT responses are displayed as (a.) the absolute breadth of responses, reflecting the mean number of overlapping peptides recognised by each patient for each HIV protein and, (b.) the relative breadth, reflecting the proportion of ELISPOT responses targeted against different HIV proteins in patients within different CD4 count strata.
doi:10.1371/journal.pone.0019018.g001

was expressed as percentage growth rate compared to the pNL4-3Δ*gag-protease* plasmid recombined with the wild type HIV-1 NL4-3 *gag-protease* insert.

Initially we determined whether CD4 cell count alone was associated with VRC. Figure 3a shows that the VRC of the chimeric viruses from patients from different CD4 T cell count strata was significantly different (Figure 3a). The VRC from viral isolates from patients with the lowest CD4 T cell counts (<100 cells/μl, median 91.15%; IQR: 86.1–98.0%) was significantly higher than from patients with high CD4 T cell counts (>500 cells/μl, median 85.19%; IQR: 79.8%–91.8%), $p = 0.0004$ (Mann-Whitney test), showing that progression to AIDS is associated with an increase in viral fitness.

There was also a weak but statistically significant positive correlation between VRC and $\log_{10}$ plasma viral load (Figure 3b, $p = 0.0003$, $r^2 = 0.088$, $n = 142$). The slope of the best-fit values (0.029 ± 0.0079) indicated that a 34.4% increase in VRC accounted for an increase of one $\log_{10}$ in plasma viral load.

VRC of chimeric *gag-protease* NL4-3 viruses varies according to HLA class I

As higher VRCs were associated with higher plasma VL and lower CD4 cell count, we examined how this effect varied according to HLA Class I (Figure 4a). The VRC of chimeric viruses derived from patients with the protective alleles HLA B*57 (median = 83%, IQR: 76–89%; n = 9) and HLA B*8101 (median = 84%, IQR: 78–92%; n = 20) was significantly lower than from patients with 'neutral/other' HLA class I alleles (median = 91%, IQR: 81–92%; n = 26); ($p = 0.007$ and $p = 0.005$, respectively), or from patients with HLA B*5802, associated with more rapid clinical progression (median = 91%, IQR: 85–98%, n = 68) ($p = 0.010$ and $p = 0.008$, respectively, Mann-Whitney test).

To determine whether the association between HLA Class I and viral fitness was additionally impacted by disease stage we stratified the VRCs according to high and low CD4 cell counts (Figure 4b). This sub-group analysis shows that the mean VRC values are higher for each HLA Class I allele in the low CD4 count group suggestive of an increase in viral fitness in advanced disease, although in most cases these values are not statistically significant. This is likely to be an effect of being underpowered, as patient numbers become smaller with each sub-categorisation. Interestingly, there is a significant difference in the VRCs of viruses in the non-protective 'neutral/other' HLA Class I alleles ($p = 0.0092$, Mann-Whitney test), showing that although there is an HLA-specific effect on viral fitness, there is an additional association with CD4 cell count, independent of the more beneficial HLA Class I alleles.

The impact of specific mutations on VRC in the context of HLA-B*8101 and HLA-B*B57/B*58

These analyses have shown that viral fitness is greater in viruses from patients with advanced disease, and that certain protective HLA Class I alleles (such as B*57, B*8101 and B*5801) are associated with greater fitness costs. The HLA-association analysis presented above (Tables 1,2 and 3) revealed that in patients with low CD4 counts HLA B*57/*5801 was associated with persistence of T242N in the TW10 epitope, whereas HLA B*8101 was associated with a potential loss of selection pressure reflected by the decreased prevalence of T186S mutants in the TL9 epitope. We therefore focused further on these well-characterised escape mutants to investigate how they impacted viral fitness at different stages of disease progression.

For HLA-B*8101, we excluded from the analysis any viruses with key mutations restricted by other HLA Class I alleles (that is, T242N and A163G) which were known to impact fitness. Viruses from patients with HLA B*8101 were significantly less fit than viruses from patients in the rest of the cohort ($p = 0.0059$, Mann-Whitney test; Figure 5a) and mutations in the B*8101-restricted Gag TL9 epitope were associated with lower VRCs ($p = 0.024$) both in the whole cohort and in patients with HLA B*8101 ($p = 0.025$). The main single mutation conferring loss of replicative fitness was the T186S immune escape mutation ($p<0.0001$). Neither the Q182S or T190X mutations had an impact on VRC in this cohort (data not shown). Although patient numbers were small, there was no significant difference between the VRCs of viruses with the T186S mutation according to high and low CD4 count stratification. These data show that although B*8101 imposes a fitness cost on the virus through selection of T186S, we found no evidence that, when this mutation persists, the fitness cost is restored at low CD4 cell counts.

We compared the fitness costs associated with HLA B*8101 with those associated with B*57 and B*5801, as the HLA-association analysis suggested that for these latter alleles escape mutations persisted in advanced disease. We pooled the data for patients with HLA B*57 and B*5801 as these alleles have highly related binding motifs, and both present the Gag TW10 epitope and select the T242N escape mutation [36]. As for HLA B*8101, viruses from patients with HLA-B*57 or HLA-B*5801 were significantly less fit than those from the rest of the cohort ($p = 0.0299$, Mann-Whitney test; Figure 5b), supporting the association between protective HLA Class I alleles and viral replicative cost (Figure 4a). In the whole cohort there was no significant effect of all mutations in the TW10 epitope on VRC. For patients with HLA B*57 or B*5801, there was a non-significant trend for all mutations within the TW10 epitope to impair fitness, although the power of this analysis was limited as only one of the 22 patients had not selected a mutant epitope. However, the fitness cost of the T242N mutation was confirmed when the whole cohort was analysed, when it was associated with a significant drop in VRC ($p = 0.0285$). This result was independent of the effect of mutations in other B*57/B*5801 associated epitopes such as Gag KF11 or Gag IW9 ($p = 0.024$; data not shown), showing that the fitness cost was associated with T242N, rather than reflecting being positive for HLA B*57/B*5801.

In the HLA-association analysis we had found that T242N was maintained in late disease, even though these data, as well as previous reports, suggest that this mutation is associated with a fitness cost. We explored this apparent discrepancy by stratifying the fitness analysis of T242N according to CD4 count (Figure 5b), and found that in patients with advanced disease (CD4 counts <100 cells/μl), viral fitness had been restored even though the T242N persisted ($p = 0.0028$).

Table 1. HLA-related associated polymorphisms in the HIV Gag protein in different CD4 count strata.

						q-value of HLA-related association in each strata			Phylogentically corrected Log2 Odds Ratio (pLOR) towards "escape"		
HIV-1 protein	HLA Class I	Population majority consensus amino acid at associated site	HXB2 site and direction of HLA associated polymorphism	Is the amino acid part of known epitope (E), or within flanking (F) region (within 5 amino acid)	Is there evidence for reversion in HLA mismatched host (R)?	Entire population	High CD4 cell count strata (>500 cells/μl)	Low CD4 cell count strata (<100 cells/μl)	High CD4 cell count strata (>500 cells/μl)	Low CD4 cell count strata (<100 cells/μl)	p-value (likelihood ratio test between pLOR of high and low CD4 strata)
Gag p17	A29	I	I7V				0.189				
Gag p17	A74	K	K12N/Q		R	0.000	0.000		18.83	- (infinity)	0.0070
Gag p17	B5801	K	K12Q		R		0.147				
Gag p17	A3303	K	K15S			0.026	0.062				
Gag p17	A6801	K	T15	F		0.131	0.086				
Gag p17	B45	H	H28R					0.058	- (infinity)	15.05	0.0005
Gag p17	C17	H	H28Q			0.000	0.183	0.017			
Gag p17	B4201	M	M30	E		0.004	0.122		14.14	9.44	0.0443
Gag p17	B0801	A	A45S		R			0.128	- (infinity)	5.29	0.0051
Gag p17	A6802	E	E55D					0.015	−6.35	18.13	0.0009
Gag p17	A3402	K	Q62				0.149		(infinity)	−10.24	0.0130
Gag p17	A3001	K	K62S				0.183		10.89	- (infinity)	0.0151
Gag p17	A30	Q	Q65H		R		0.086		9.27	−11.37	0.0086
Gag p17	B8101	Q	K69				0.163		(infinity)	−0.21	0.0027
Gag p17	A26	K	K76R				0.149				
Gag p17	B15	Y	79Y		R		0.188		14.22	−1.66	0.0193
Gag p17	C04	K	K91N		R			0.016	15.78	4.21	0.0004
Gag p17	B08	Q	Q112K				0.159				
Gag p17	B1402	K	K114				0.069		7.73	- (infinity)	0.0498
Gag p17	B1801	T	T115				0.146				
Gag p17	A2902	Y	Y132F				0.188				
Gag p24	B57	A	A146P	E		0.000	0.000	0.048			
Gag p24	C17	A	P146		R		0.102				
Gag p24	B1510	A	A146	E		0.008	0.153				
Gag p24	C0304	I	147I	E			0.102		(infinity)	1.41	0.0209
Gag p24	B57	I	I147L/M	E	R	0.000	0.000				
Gag p24	C17	Q	Q182					0.017	0.79	4.18	0.0019
Gag p24	B8101	Q	Q182S	E		0.000	0.000	0.017	13.97	16.79	0.0349
Gag p24	B8101	T	T186S	E	R	0.000		0.000	18.38	7.25	0.0397
Gag p24	C0210	V	V218		R			0.056			
Gag p24	B5801	T	T242N	E		0.000	0.000	0.000			

Table 1. Cont.

						q-value of HLA-related association in each strata			Phylogentically corrected Log2 Odds Ratio (pLOR) towards "escape"		
HIV-1 protein	HLA Class I	Population majority consensus amino acid at associated site	HXB2 site and direction of HLA associated polymorphism	Is the amino acid part of known epitope (E), or within flanking (F) region (within 5 amino acid)	Is there evidence for reversion in HLA mismatched host (R)?	Entire population	High CD4 cell count strata (>500 cells/μl)	Low CD4 cell count strata (<100 cells/μl)	High CD4 cell count strata (>500 cells/μl)	Low CD4 cell count strata (<100 cells/μl)	p-value (likelihood ratio test between pLOR of high and low CD4 strata)
Gag p24	B57	T	T242N	E	R	0.000	0.000	0.000			
Gag p24	A74	I	I247		R		0.087				
Gag p24	C04	S	N252				0.189				
Gag p24	B35	D	D260E	E		0.000	0.000				
Gag p24	B1401	K	K302R	E	R	0.000	0.015				
Gag p24	B44	D	D312E	E	R	0.000	0.183				
Gag p24	B0702	D	D319E				0.086				
Gag p24	C08	T	T332A			0.131		0.015			
Gag p24	B0702	G	S357G	E		0.000	0.001	0.000			

The table shows the results of the association analysis between HLA Class I alleles and HIV viral polymorphisms in HIV Gag. The columns of the table detail the HIV protein, the associated HLA Class I allele and viral polymorphism, whether the association lies in, or within, the flanking region of a known restricted epitope and whether this association is expected to revert in HLA-mismatched hosts. The q values for the statistically significant associations for the whole cohort, the 'high' CD4 count subgroup (>500 cells/μl) and 'low' CD4 count subgroup (<100 cells/μl) are shown, followed by the $\log_2$-adjusted odds ratios for the 'high' and 'low' groups. In the final column, the p value is reported, where there is a significant difference between the two odds ratios.

doi:10.1371/journal.pone.0019018.t001

Table 2. HLA-related associated polymorphisms in the HIV Pol protein in different CD4 count strata.

						q-value of HLA-related association in each strata			Phylogentically corrected Log2 Odds Ratio (pLOR) towards "escape"		
HIV-1 protein	HLA Class I	Population majority consensus amino acid at associated site	HXB2 site and direction of HLA associated polymorphism	Is the amino acid part of known epitope (E), or within flanking (F) region (within 5 amino acid)	Is there evidence for reversion in HLA mismatched host (R)?	Entire population	High CD4 cell count strata (>500 cells/μl)	Low CD4 cell count strata (<100 cells/μl)	High CD4 cell count strata (>500 cells/μl)	Low CD4 cell count strata (<100 cells/μl)	p-value (likelihood ratio test between pLOR of high and low CD4 strata)
Protease	C0401	L	L10		R		0.178				
Protease	C18	S	S12T					0.010	2.13	16.45	0.0005
Protease	B5802	S	T12S			0.000		0.007			
Protease	C06	S	12S					0.007			
Protease	B1801	I	I19T					0.162			
Protease	B1510	E	D35E				0.155		−5.82	20.55	0.0015
Protease	B44	E	E35D	E		0.000		0.000			
Protease	A6601	N	N37			0.167		0.121			
Protease	C0602	L	V63T	E		0.192		0.190	13.30	(infinity)	0.0044
Protease	C16	L	P63L		R			0.013			
Protease	C18	L	P63S		R	0.000		0.064			
Protease	C08	L	63L					0.100			
Protease	A43	I	I64L		R	0.167		0.097	- (infinity)	20.66	0.0477
Protease	C06	V	I77V			0.004		0.013	3.60	6.41	0.0129
Protease	A66	V	V77I					0.196			
RT	B81	P	P103S/T	E		0.000	0.155	0.000	10.43	22.85	0.0092
RT	C18	E	E105K			0.001	0.178		10.97	−1.39	0.0207
RT	B81	E	E105	E		0.000	0.178				
RT	B18	K	K119R		R			0.073			
RT	B15	E	E152D					0.035			
RT	B4101	E	E152D			0.141		0.080			
RT	A34	V	V159I					0.121	0.54	17.36	0.0155
RT	B35	E	E221K	E		0.003	0.178				
RT	C02	G	S222			0.002		0.015			
RT	B8101	I	I234R					0.131	−5.66	22.55	0.0379
RT	B4202	I	I234V					0.036			
RT	B0702	S	S261A/C	E		0.000		0.041	15.26	16.55	0.0348
RT	C0702	S	S261					0.029			
RT	B07	T	T264I	E		0.000		0.000			
RT	A03	K	K265R	E	R	0.005		0.121			
RT	C0702	E	E268					0.010			
RT	B1503	K	K273R	E		0.002		0.003	−14.09	18.11	0.0302
RT	A03	K	K273R	F		0.000		0.035	15.22	18.75	0.0444