Acknowledgements We are thankful to L. Nagy for the idea of performing embryo transplantation into host eggs; E. Zhurov for expert embryo transplantation; P. Lasko, L. Nagy, L. Dickinson, T. Reavell, T. Drysdale, P. Wigge and G. Velicer for critical reading of the manuscript; E. Zhurov and I. Craig for the artwork; J. Whistlecraft, L. Verdon and N. Terzin for *Copidosoma* rearing; and R. Kulperger for the statistical analysis. M.G. thanks V. Grbić and Ž. Srdić for introducing him to parasitic wasps. This work was supported by grants from the Canadian Foundation of Innovation, Premier's Research Excellence Award and National Science and Engineering Research Council to M.G.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.G. (mgrbic@uwo.ca).

Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA

Photini Kiepiela[1], Alasdair J. Leslie[2], Isobella Honeyborne[1,2], Danni Ramduth[1], Christina Thobakgale[1], Senica Chetty[1], Prinisha Rathnavalu[1], Corey Moore[3], Katja J. Pfafferott[2], Louise Hilton[2], Peter Zimbwa[2], Sarah Moore[4], Todd Allen[5], Christian Brander[5], Marylyn M. Addo[5], Marcus Altfeld[5], Ian James[3], Simon Mallal[3], Michael Bunce[6], Linda D. Barber[7], James Szinger[8], Cheryl Day[2], Paul Klenerman[2], James Mullins[4], Bette Korber[8], Hoosen M. Coovadia[1], Bruce D. Walker[1,5,9] & Philip J. R. Goulder[1,2,5]

[1]*HIV Pathogenesis Programme, The Doris Duke Medical Research Institute, University of KwaZuluNatal, 719 Umbilo Road, Durban 4015, South Africa*
[2]*Department of Paediatrics, Nuffield Department of Medicine, Peter Medawar Building for Pathogen Research, South Parks Road, Oxford OX1 3SY, UK*
[3]*Centre for Clinical Immunology and Biomedical Statistics, Royal Perth Hospital and Murdoch University, Level 2 North Block, Royal Perth Hospital, Wellington Street, Western Australia 6000, Australia*
[4]*Department of Microbiology, University of Washington, Seattle, Washington 98195-8070, USA*
[5]*Partners AIDS Research Center, Massachusetts General Hospital, 149 13th Street, Charlestown, Massachusetts 02129, USA*
[6]*Dynal Biotech Ltd, 11 Bassendale Road, Wirral CH62 3QL, UK*
[7]*Anthony Nolan Research Institute, Royal Free Hospital, Pond Street, London NW3 2QG, UK*
[8]*Santa Fe Institute, Santa Fe, New Mexico 87501, USA, and Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[9]*Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA*

The extreme polymorphism in the human leukocyte antigen (HLA) class I region of the human genome is suggested to provide an advantage in pathogen defence mediated by $CD8^+$ T cells[1–3]. HLA class I molecules present pathogen-derived peptides on the surface of infected cells for recognition by $CD8^+$ T cells. However, the relative contributions of HLA-A and -B alleles have not been evaluated. We performed a comprehensive analysis of the class I restricted $CD8^+$ T-cell responses against human immunodeficiency virus (HIV-1), immune control of which is dependent upon virus-specific $CD8^+$ T-cell activity[4,5]. In 375 HIV-1-infected study subjects from southern Africa, a significantly greater number of $CD8^+$ T-cell responses are HLA-B-restricted, compared to HLA-A (2.5-fold; $P = 0.0033$). Here we show that variation in viral set-point, in absolute CD4 count and, by inference, in rate of disease progression in the cohort, is strongly associated with particular HLA-B but not HLA-A allele expression ($P < 0.0001$ and $P = 0.91$, respectively). Moreover, substantially greater selection pressure is imposed on HIV-1 by HLA-B alleles than by HLA-A (4.4-fold, $P = 0.0003$). These data indicate that the principal focus of HIV-specific activity is at the HLA-B locus. Furthermore, HLA-B gene frequencies in the population are those likely to be most influenced by HIV disease, consistent with the observation that B alleles evolve more rapidly than A alleles[6–8]. The dominant involvement of HLA-B in influencing HIV disease outcome is of specific relevance to the direction of HIV research and to vaccine design.

Despite apparently similar roles in pathogen defence, there are well-established differences between the class I loci. HLA-C alleles are expressed at lower levels on the cell surface than HLA-A and HLA-B (ref. 9) and thus it is not unexpected that HLA-C shows the least diversity of the three classical HLA class I loci. However, an unexplained observation is that greater diversity exists at the HLA-B locus than at the HLA-A locus. Overall, there are 563 HLA-B alleles described, compared with 309 HLA-A alleles and 167 HLA-C alleles (http://www.ebi.ac.uk/imgt/hla/docs/release.html). It appears that the HLA-B locus is diversifying more rapidly than the HLA-A locus; the significance of which remains unknown.

In order to test the hypothesis that the observed differential diversity at the class I loci reflects functional differences between the HLA-A, -B and -C-restricted $CD8^+$ cytotoxic T lymphocyte (CTL) responses, we investigated the contribution of different HLA class I molecules within the $CD8^+$ T-cell activity directed against HIV—a pathogen whose control is strongly influenced by $CD8^+$ T cells.

We used an empirical approach to determine the contributions of individual HLA class I molecules in anti-HIV host defence in 375 infected, treatment-naive persons in southern Africa, a region burdened with more HIV infections than any other (http://www.unaids.org). This avoids the heavy biases inherent in the use of previously defined optimal epitopes that largely favour particular, well-studied, HLA class I molecules. We employed a panel of 410 overlapping synthetic peptides, spanning the entire expressed HIV genome, and characterized the T-cell responses to these peptides in interferon-γ enzyme-linked immunospot (elispot) assays. Previous studies have validated this approach, and have shown that the responses detected by this method are, with few exceptions, the result of $CD8^+$ T-cell activity[10].

This analysis revealed marked differences in the frequency of targeting of individual peptides; 132 peptides being targeted by no subjects, and five peptides targeted by >25% of persons. For many targeted peptides, we noted a strong association with specific class I allele expression: 180 peptides exhibited a strong allele-specific association ($P < 0.05$), of which 111 remained significant ($P < 0.001$) following correction for multiple comparisons (see Supplementary Tables 1 and 2). These included 58 previously defined epitopes (http://www.hiv.lanl.gov/content/immunology/tables/ctl_summary.html). To demonstrate further that the HLA–peptide associations reflect true HLA class I restricted $CD8^+$ T-cell responses, we experimentally defined the restriction elements for 12 additional, randomly selected epitopes (Fig. 1, Supplementary Fig. 1 and Supplementary Tables 1 and 2). In each case the restricting allele was precisely that predicted by the statistical association. This analysis therefore affords a stringent approach to define the HLA class I restriction for the large majority of $CD8^+$ T-cell responses detected in this population.

We next used this approach to determine the relative contributions of the different class I alleles to immune recognition of HIV-1. The association of HLA-B alleles with peptide-specific responses far exceeded that of HLA-A alleles (67 versus 25/111, $P = 0.0033$, two-tailed unpaired *t*-test; mean peptide associations per allele, 3.94 versus 1.56) and of HLA-C alleles (19/111, $P = 0.004$; Fig. 2a and b). Of the 30 most highly targeted peptides, 67% were HLA-B-restricted responses (Fig. 2c). Thus, HLA-B alleles contribute significantly more towards the total HIV-specific

$CD8^+$ T-cell response than either HLA-A or HLA-C in this population.

We next examined the influence of individual HLA alleles on the level of plasma viraemia, a strong predictor of time to AIDS in B-clade infection[11]. HLA allele expression and viral load were documented in an extended cohort ($n = 706$) of antiretroviral therapy-naive, chronically infected Zulu/Xhosa study subjects. Viral load did not vary in the cohort in relation to particular HLA-A allele expression ($P = 0.914$, one-way analysis of variance; Fig. 3a), but varied significantly according to the particular HLA-B allele expressed ($P < 0.0001$), suggesting that an individual's HLA-B type has a profound influence on viral load. In order to identify the particular alleles responsible for differential outcome, as indicated by viral load, we next sought individual alleles for which there were significant differences between the viral loads in subjects who either expressed or did not express the allele. Of 10 significant associations identified (uncorrected, $P < 0.05$), seven were through HLA-B alleles (Fig. 3b). Four of five associations that remained significant after correction for multiple comparisons ($P < 0.001$) were through HLA-B alleles; none was through HLA-A. Thus, the majority of alleles associated with either higher or lower than expected viral loads were HLA-B.

Re-analysis of these HLA associations with high/low viral load indicated that the apparent associations through HLA-A*0205, HLA-Cw*0401 and HLA-Cw*0602 were explained by linkage disequilibrium with HLA-B*5801, HLA-B*8101 and HLA-B*5802, respectively (Fig. 3c, Supplementary Table 3). In contrast, with one exception, the seven HLA-B associations were not altered by linkage disequilibrium effects. Thus, all six alleles (four, following correction for multiple comparisons) independently associated with either higher or lower than expected viral loads were HLA-B. Together, these data show that HLA-B alleles bear the major burden of HIV-specific CTL activity in chronic HIV infection, and that these are the alleles principally associated with diverse outcomes from HIV infection in this population.

In order to test further the inference that it is particular HLA-B allele expression that principally influences disease outcome, we analysed the absolute CD4 counts in the same cohort ($n = 706$) of chronically infected Zulu/Xhosa study subjects in relation to HLA type. Although there is no recognised steady-state for CD4 count as there is for viral load in chronic infection, CD4 T cell counts are an independent predictor of disease progression in B- and C-clade infection[11,12], and a close relationship was also observed between particular HLA-B alleles expressed and disease outcome inferred by CD4 count (Supplementary Fig. 2).

These studies suggest a dominant role for HLA-B alleles in successful or unsuccessful immune containment of HIV infection. To determine whether B alleles exert stronger immune selection pressure on HIV (ref. 13), we examined HIV amino-acid sequence variation in relation to HLA allele expression. Analysis of Nef and Gag, the most immunogenic proteins, demonstrated 25 viral polymorphisms in the Durban cohort associated with HLA-B, 12 with HLA-A and 9 with HLA-C. Eleven of the 46 polymorphisms identified lie within previously defined epitopes, of which ten are presented by the relevant HLA-B allele (data for Nef are shown in Fig. 4a). To investigate further the differential impact of class I loci

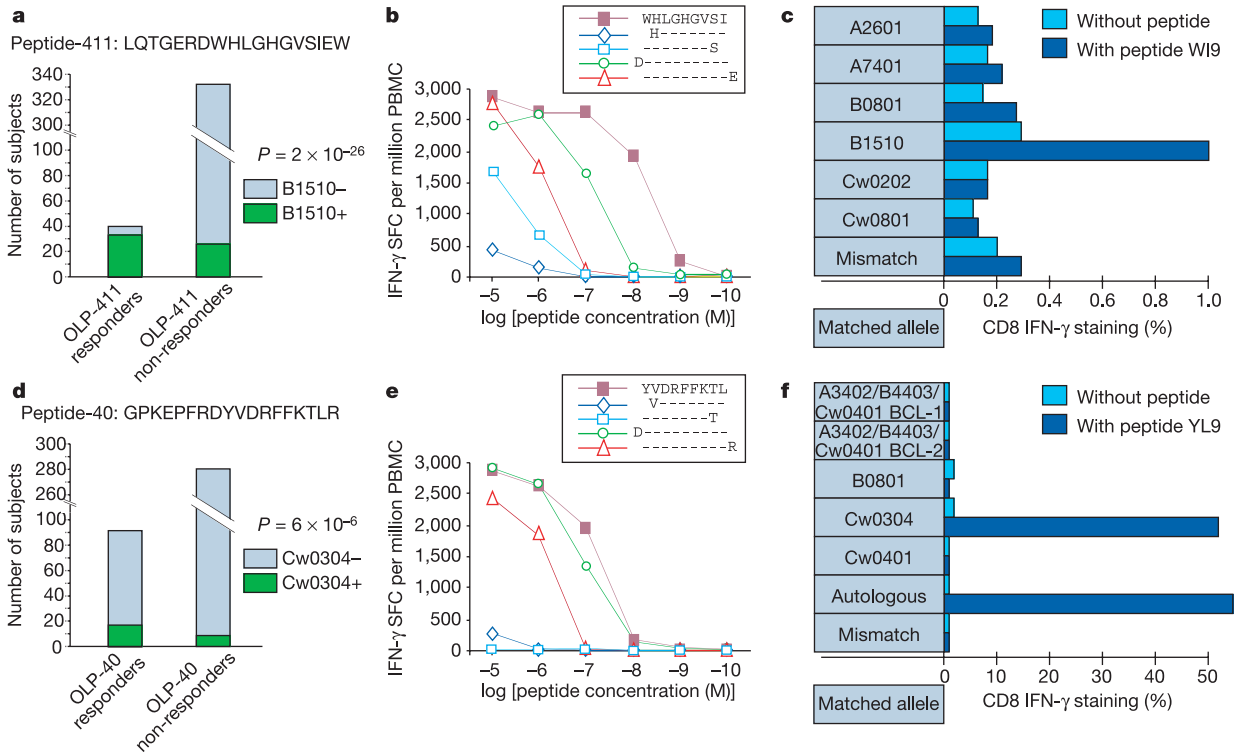

Figure 1 Epitope optimization and formal definition of HLA restriction from HLA–peptide associations. (Total number of subjects, $n = 375$.) **a**, Recognition of peptide-411 associated with HLA-B*1510 expression. **b**, WHLGHGVSI (WI9) is the optimal epitope. **c**, HLA-B*1510 is the restriction element. Intracellular interferon-γ (IFN-γ) staining assay using PBMC from donor SK-031 (HLA-A*2601/*7401 B*0801/*1510 Cw*0202/*0801) and BCL matched through individual HLA class I molecules as shown. **d**, Recognition of peptide-40 associated with HLA-Cw*0304 expression. (Total number of subjects, $n = 375$.) **e**, YVDRFFKTL (YL9) is the optimal epitope. SFC: spot forming cells; PBMC: peripheral blood mononuclear cells. **f**, HLA-Cw*0304 is the restriction element. Intracellular interferon-γ staining using YL9-specific CTL line from donor 1157-M (HLA-A*3402/−B*0801/*4403 Cw*0304/*0401) and BCL matched through individual HLA class I molecules as shown.

on HIV-1 sequence polymorphism, we analysed a previously reported cohort from Western Australia[14]. The HIV-1 Pol sequence polymorphisms in this B-clade infected cohort had shown 12 strong associations with particular HLA class I allele expression, of which 11 were with HLA-B alleles. HLA-C allele associations were not analysed[14]. We re-analysed these data to include HLA-C alleles and all the viral proteins. The rates of polymorphism associated with HLA-A, -B and -C were, respectively, 0.0022, 0.0096 and 0.0059 per amino acid (Fig. 4b). Polymorphism was significantly more commonly associated with HLA-B than with HLA-A ($P = 0.0003$). This analysis therefore provides further evidence that HLA-B-restricted $CD8^+$ T-cell responses are the most important in shaping the evolution of HIV.

This study of C-clade infected individuals from the region worst affected by the HIV epidemic shows that most HIV-specific CTL responses are HLA-B-restricted, and that outcome from HIV infection, inferred from viral set-point and CD4 count, is principally associated with particular HLA-B allele expression. Moreover, immune selection pressure on HIV is disproportionately driven by HLA-B alleles. This provides biological evidence to support theoretical data regarding the differential evolution of class I loci, where HLA-B is the most rapidly evolving[6–8], and to indicate significant functional differences among individual loci in infectious disease.

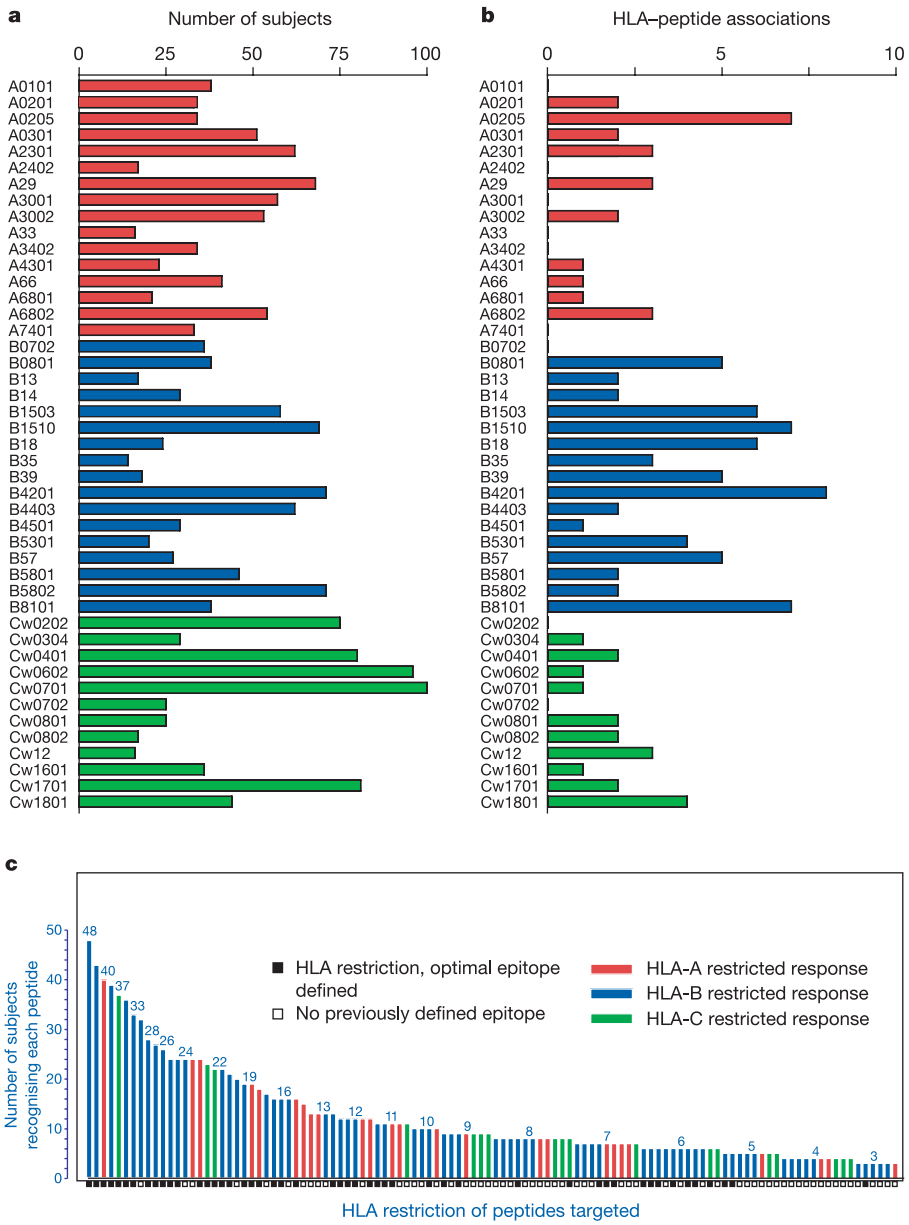

Figure 2 HLA frequencies in study cohort and peptide recognition. **a**, Phenotypic frequencies of HLA class I molecules in the study cohort ($n = 375$) showing alleles present at frequencies of >3%. **b**, Number of peptides recognised in association with individual HLA allele expression. More associations were observed for HLA-B alleles compared to HLA-A alleles ($P = 0.0033$) and for HLA-B compared to HLA-C ($P = 0.0004$). **c**, Number of responders for each peptide for which a peptide–HLA association (in panel **b**) was identified. Filled boxes: previously defined epitope presented by the relevant HLA molecule; open boxes: epitope within the peptide yet to be defined.

These data also support the hypothesis that pathogens controlled in part by the CTL response can adapt and impose selection pressure on the host, particularly on the HLA-B locus. In HIV, both survival and transmission risk depend substantially on viral load[15,16], and therefore on the HLA-B alleles expressed. Comparing one generation with the next in this Zulu/Xhosa population, the frequency of deleterious alleles, HLA-B*5802 and B*18, is greater in infected infants (52%, $n = 52$) than in infected antenatal mothers (31%, $n = 511$), while the frequency of protective alleles, B*57 and B*5801, is greater in the infected antenatal mothers (19% versus 6%; $P = 0.0013$, Fig. 5). These data suggest that the frequencies of B*18 and B*5802 will decrease, and those of B*57 and B*5801 will increase rapidly in the population as the epidemic progresses.

These conclusions are supported by previous studies in B-clade infected Caucasians which show that HLA-B alleles are most closely associated with non-progression/low viral load[17–19] or progression/high viral load[4,18,20]. Although our study was cross-sectional, and therefore lacks the statistical power afforded to longitudinal studies, the associations between B*57 and B*5801 with low viraemia, and between B*18 and B*5802 with high viraemia, were sufficiently strong to stand out. The dominant effect of HLA-B-restricted CTL responses on HIV outcome is also supported by work that shows HLA class I homozygous disadvantage in HIV infection[4]. The relative hazards for progression to AIDS or death were 2- to 3-fold higher for HLA-B homozygosity compared with HLA-A homozygosity. Moreover, studies of rare supertype advantage demonstrated an HLA-B effect only[21].

Although these data indicate a dominant role for HLA-B alleles in HIV infection, the mechanism underlying these observations is unknown. The effect is unlikely to be due to differences in NK cell activation through the HLA-B Bw4 motif[22], because the related alleles HLA-B*57, B*5801 and B*5802, associated with diverse outcomes, express the identical Bw4 motif[3]. Moreover, data describing the epistatic interaction between KIR3DS1 and HLA-B indicated that HLA-B*57- and B*27-mediated protection was entirely independent of this interaction[23].

One possible mechanism for the dominant HLA-B allele effect is the greater diversity of peptides bound by HLA-B alleles[3]. Amino acids binding the B pocket of HLA-A alleles are, without exception, broadly hydrophobic residues (excluding proline, however). In contrast, the B pocket of HLA-B alleles can variously accommodate these residues; but also proline, positively and negatively charged residues, and histidine and glutamine[3]. The principal source of

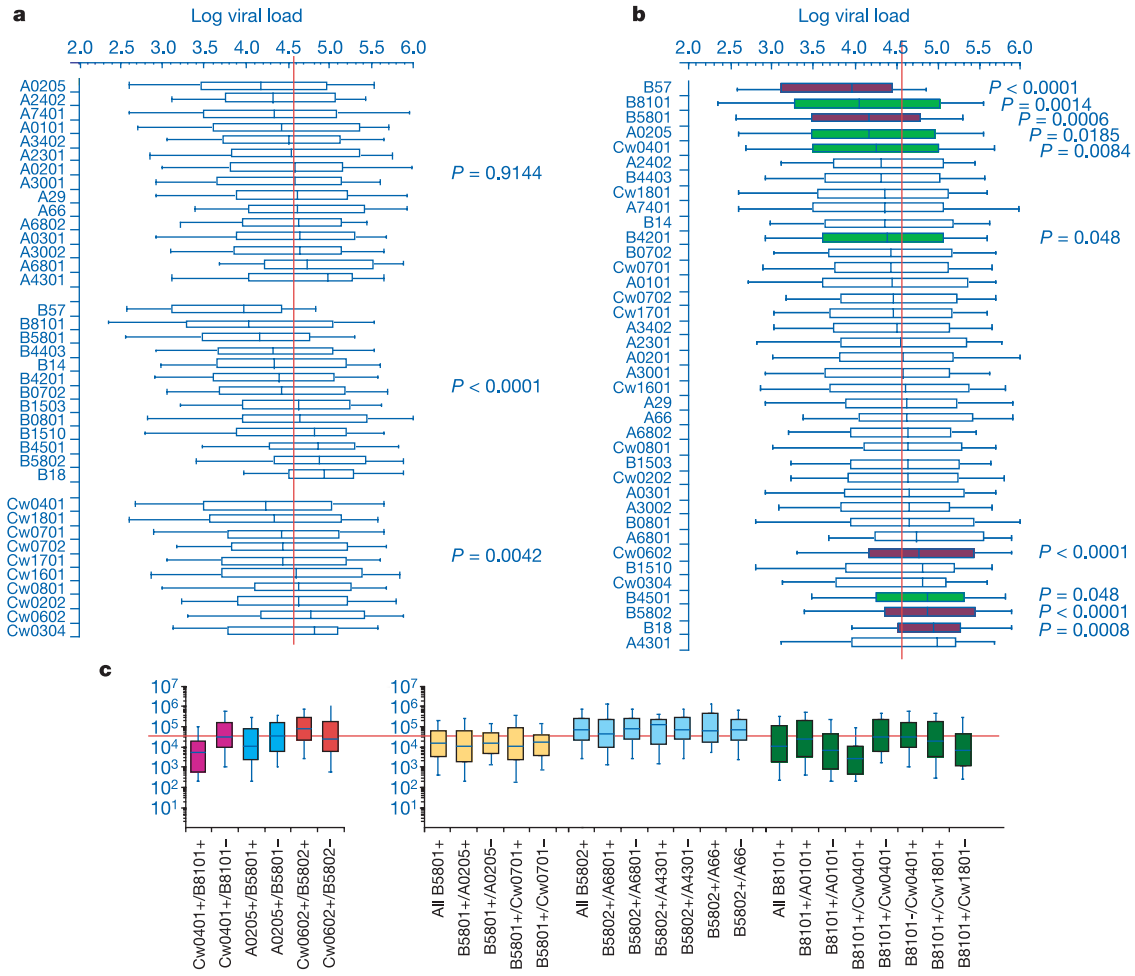

Figure 3 HLA class I molecule expression and viral load in chronically infected Zulu/Xhosa ($n = 706$). **a**, Contribution of class I loci to viral load variation. Viral loads shown as box plots (95th, 75th, 50th, 25th and 5th centiles); vertical red line indicates cohort median viral load. **b**, Association of individual HLA alleles with differential viral loads. Purple shading: $P < 0.001$. Green shading: $0.05 > P > 0.001$. **c**, Effect of linkage disequilibrium on the associations observed in **b**. Left panel: viral load associations of HLA-Cw*0401, A*0205 and Cw*0602 are lost. Right panel: viral load association of HLA-B*8101 is lost; however, HLA-B*5801, B*5802 (shown), and B*57, B*4201 and B*4501 (not shown) associations are maintained irrespective of linkage disequilibrium effects.

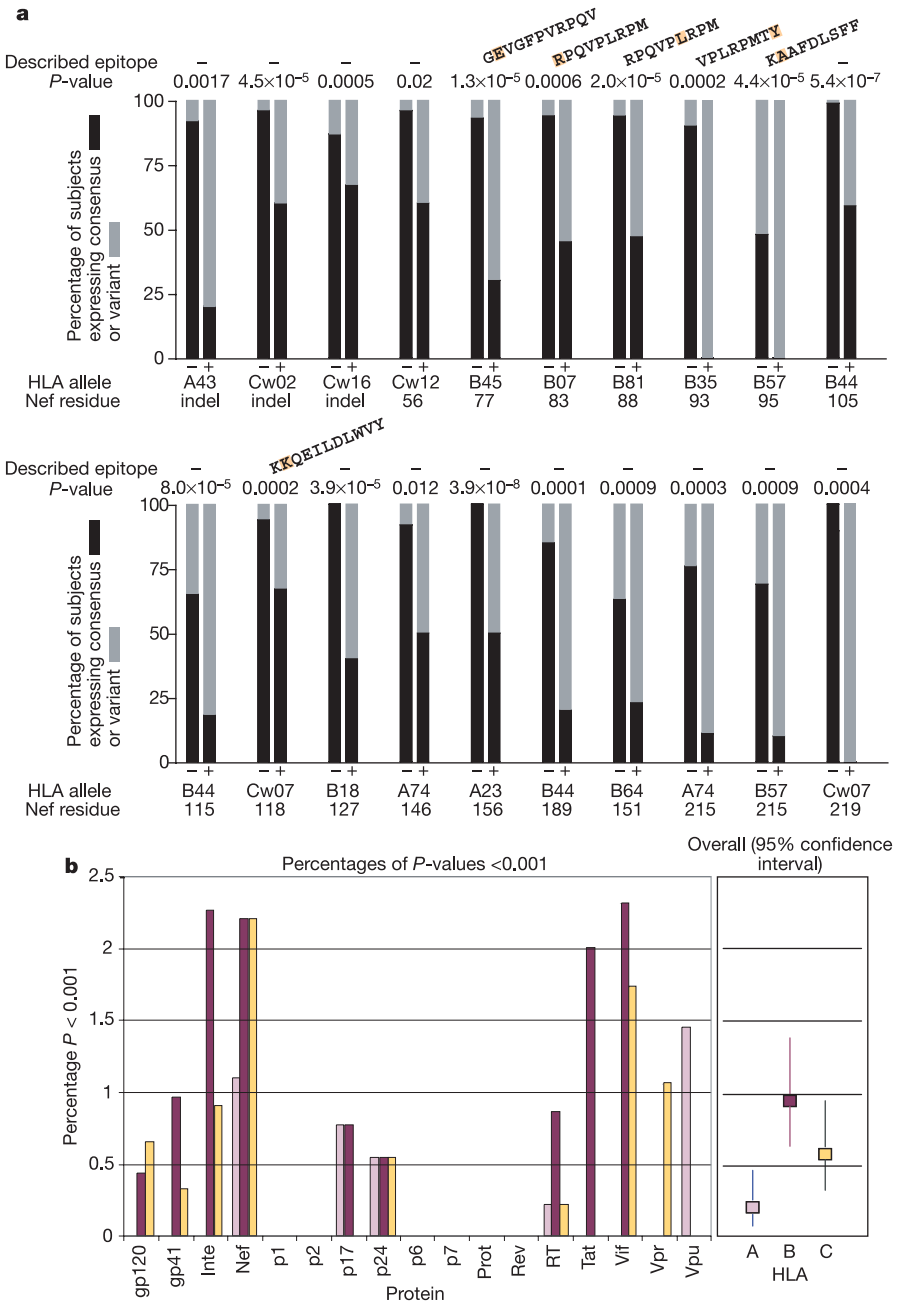

Figure 4 Expression of HLA class I molecules and HIV sequence polymorphism. **a**, Analysis of Nef sequences from 123 subjects from Durban, South Africa. Pairs of bars indicate the frequency of individuals whose viral sequence is consensus (black bars) or variant (grey bars) in subjects either expressing (+) or not expressing (−) the relevant allele. Twenty-one associations between expression of particular HLA class I alleles and sequence polymorphisms were identified, as shown. Associations identified within defined epitopes presented by the relevant HLA allele are as shown. Indel, insertion/deletion. **b**, Percentages of amino-acid positions at which polymorphism away from population consensus was significantly associated with the HLA-A, -B and -C allelic groups ($P < 0.001$). Numbers of positions considered: 14–263 per protein; a total of 2713 positions across all proteins. Comparisons (McNemar tests): $P = 0.0003$ and $P = 0.12$ for HLA-B versus HLA-A and HLA-C, respectively; $P = 0.04$ for HLA-C versus HLA-A.

HLA-B diversity, however, arises from intralocus recombination events within exon 3, primarily affecting the F, C and D pockets[3]. The three amino acid differences between HLA-B*5801 and B*5802, for example, would be predicted only to affect the C pocket. We speculate that a variable pathogen such as HIV may adapt more readily to the homogeneous terrain represented by HLA-A alleles, and less so to the functionally diverse HLA-B. This greater functional diversity of HLA-B suggests that, even as the epidemic evolves, the dominant focus of immune-mediated selection pressure will remain channelled through HLA-B alleles, though not necessarily through the particular alleles predominantly involved today.

These data indicate marked functional differences between HLA-A and -B alleles to the HIV-specific $CD8^+$ T-cell response. Differences in viral load and absolute CD4 count, and, by inference, in HIV disease outcome, are principally related to HLA-B allele expression, as is HLA-mediated immune selection pressure. This work does not prove a co-evolutionary process between HIV and

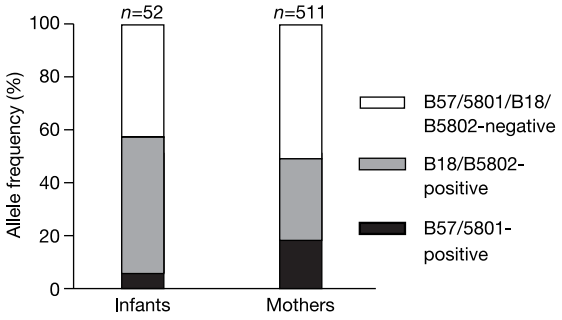

Figure 5 Altered frequencies of alleles associated with high or low viral set-point in infected infants and mothers. Frequencies of protective alleles (B*57 and B*5801) and susceptibility alleles (B*18 and B*5802) differed significantly in the two groups ($P = 0.0013$, Fisher's exact test).

HLA—which would require the availability of samples derived from the two co-evolving genetic entities obtained over great time periods. However, it provides a snapshot in time that is highly indicative of these dynamics in populations where HIV disease is showing high morbidity and mortality rates. The previous observation that HLA-B is evolving more rapidly than the HLA-A or -C loci suggests that, for pathogens in addition to HIV whose control depends upon the $CD8^+$ T-cell response, similar effects may be observed. The immunogenetic analyses of human diseases undertaken to date largely support this hypothesis[24–26], and suggest a dominant involvement of HLA-B alleles in the control of human pathogens. □

Methods

Study subjects

Characterization of HIV-specific T-cell responses was undertaken in study subjects from southern Africa. Of the 375 subjects studied, 314 were from Durban, South Africa, the remainder being from Zimbabwe, Angola, Malawi, Zambia and Botswana. In 219 cases, asymptomatic subjects were tested for HIV infection during pregnancy following voluntary counselling, and were studied subsequent to diagnosis; 156 subjects were recruited from outpatient follow-up clinics. This latter group were mostly diagnosed following clinical presentation suggestive of HIV infection. The subjects for whom analysis of HLA type and viral load was performed ($n = 706$) were all antiretroviral therapy naive, from Durban, South Africa. This group comprised 494 asymptomatic subjects recruited from antenatal clinics as described above and 212 subjects from out-patient follow-up clinics. Fifty-three of these 706 enrolled subjects (7.5%) were males. Overall, the median viral load in this cohort was 37,500 and the median absolute CD4 count was 387. Viral load measurement was undertaken using the Roche Amplicor version 1.5 assay.

Elispot, HLA restriction assays

Four hundred and ten 18-mer peptides, spanning the expressed entire HIV genome, were synthesized on the basis of the consensus of available C-clade sequences in 2001. These peptides were used in a matrix system of 11–12 peptides per pool to screen study subjects for HIV-specific T cell responses by interferon-γ elispot assay, as previously described[10]. Confirmation of recognised individual 18-mer peptides within a peptide pool was undertaken in a separate elispot assay. Formal definition of the HLA restriction of the responses was carried out either by intracellular cytokine staining or by elispot assay using PBMC or CTL lines as effectors.

HLA typing

For the subjects studied in the Durban cohort, genomic DNA samples were initially typed to an oligo-allelic level using Dynal RELITM reverse Sequence Specific Oligonucleotide (SSO) kits for the HLA-A, -B and -C loci (Dynal Biotech). Refining the genotype to the allele level was performed using the Dynal Biotech sequence-specific priming kits in conjunction with the previous SSO type. Where alleles were still not defined to the allele level, bespoke sequence-specific priming primer mixes were utilised. All HLA class I alleles in the IMGT allele release 2.4.0 were considered in the typing[27]. HLA class I typing for the Perth cohort was performed by direct sequencing methods[14].

Viral sequencing

Sequencing of HIV proviral DNA from patients in the Durban cohort was undertaken as previously described[28]. Sequencing of proviral DNA and viral RNA from patients in the Perth cohort was undertaken as described[14].

Statistics

See Supplementary Information.

Received 14 August; accepted 14 October 2004; doi:10.1038/nature03113.

1. Jeffery, K. J. & Bangham, C. R. Do infectious diseases drive MHC diversity? *Microbes Infect.* **2,** 1335–1341 (2000).
2. Parham, P. & Ohta, T. Population biology of antigen presentation by MHC class I molecules. *Science* **272,** 67–74 (1996).
3. Marsh, S. G. E., Parham, P. & Barber, L. D. *The HLA Facts Book* (Academic, London, 2000).
4. Carrington, M. *et al.* HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* **283,** 1748–1752 (1999).
5. McMichael, A. J. & Rowland-Jones, S. L. Cellular immune responses to HIV. *Nature* **410,** 980–987 (2001).
6. Belich, M. P. *et al.* Unusual HLA-B alleles in two tribes of Brazilian Indians. *Nature* **357,** 326–329 (1992).
7. Watkins, D. I. *et al.* New recombinant HLA-B alleles in a tribe of South American Amerindians indicate rapid evolution of MHC class I loci. *Nature* **357,** 329–333 (1992).
8. McAdam, S. N. *et al.* A uniquely high level of recombination at the HLA-B locus. *Proc. Natl Acad. Sci. USA* **91,** 5893–5897 (1994).
9. Snary, D., Barnstable, C. J., Bodmer, W. F. & Crumpton, M. J. Molecular structure of human histocompatibility antigens: the HLA-C series. *Eur. J. Immunol.* **7,** 580–585 (1977).
10. Addo, M. M. *et al.* 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,** 2081–2092 (2003).
11. Mellors, J. W. *et al.* Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science* **272,** 1167–1170 (1996).
12. Badri, M. *et al.* Initiating highly active antiretroviral therapy in sub-Saharan Africa: an assessment of the revised World Health Organization scaling-up guidelines. *AIDS* **18,** 1159–1168 (2004).
13. Goulder, P. J. & Watkins, D. I. HIV and SIV CTL escape: implications for vaccine design. *Nature Rev. Immunol.* **4,** 630–640 (2004).
14. Moore, C. B. *et al.* Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* **296,** 1439–1443 (2002).
15. Cooper, E. R. *et al.* Combination antiretroviral strategies for the treatment of pregnant HIV-1-infected women and prevention of perinatal HIV-1 transmission. *J. Acquir. Immune Defic. Syndr.* **29,** 484–494 (2002).
16. Garcia, P. M. *et al.* Maternal levels of plasma human immunodeficiency virus type 1 RNA and the risk of perinatal transmission. Women and Infants Transmission Study Group. *N. Engl. J. Med.* **341,** 394–402 (1999).
17. Kaslow, R. A. *et al.* Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nature Med.* **2,** 405–411 (1996).
18. O'Brien, S. J., Gao, X. & Carrington, M. HLA and AIDS: a cautionary tale. *Trends Mol. Med.* **7,** 379–381 (2001).
19. Migueles, S. A. *et al.* HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. *Proc. Natl Acad. Sci. USA* **97,** 2709–2714 (2000).
20. Gao, X. *et al.* Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *N. Engl. J. Med.* **344,** 1668–1675 (2001).
21. Trachtenberg, E. *et al.* Advantage of rare HLA supertype in HIV disease progression. *Nature Med.* **9,** 928–935 (2003).
22. Flores-Villanueva, P. O. *et al.* Control of HIV-1 viremia and protection from AIDS are associated with HLA-Bw4 homozygosity. *Proc. Natl Acad. Sci. USA* **98,** 5140–5145 (2001).
23. Martin, M. P. *et al.* Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nature Genet.* **31,** 429–434 (2002).
24. Cooke, G. S. & Hill, A. V. Genetics of susceptibility to human infectious disease. *Nature Rev. Genet.* **2,** 967–977 (2001).
25. Ramos, M. & Lopez de Castro, J. A. HLA-B27 and the pathogenesis of spondyloarthritis. *Tissue Antigens* **60,** 191–205 (2002).
26. Thio, C. L. *et al.* Comprehensive analysis of class I and class II HLA antigens and chronic hepatitis B virus infection. *J. Virol.* **77,** 12083–12087 (2003).
27. Bunce, M. PCR-sequence-specific primer typing of HLA class I and class II alleles. *Methods Mol. Biol.* **210,** 143–171 (2003).
28. Leslie, A. J. *et al.* HIV evolution: CTL escape mutation and reversion after transmission. *Nature Med.* **10,** 282–289 (2004).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank those who made possible the collection of blood samples for the study, in particular the patients and staff at Cato Manor Clinic and Sinikithemba Clinic, Durban, South Africa. We also thank S. McAdam, J. Webster and R. Phillips for comments and suggestions. This work was supported by a contract from the National Institutes of Health (B.D.W.), with additional support from the Wellcome Trust (P.J.R.G., A.J.L., P.Z., P.K.), the Elizabeth Glaser Pediatric AIDS Foundation (P.J.R.G., B.T.K.), the National Institutes of Health and the Doris Duke Charitable Foundation. P.J.R.G. and B.T.K. are Elizabeth Glaser Pediatric AIDS Foundation Scientists, B.D.W. is a Doris Duke Distinguished Science Professor.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.J.R.G. (philip.goulder@ndm.ox.ac.uk). GenBank accession numbers and Full Gag and Nef sequence alignments for the sequences from Durban, South Africa, included in this study are available at ⟨http://www.hiv.lanl.gov/content/immunology/hlatem/south_africa_sequences/⟩ and other accession numbers are available in the Supplementary Information.