Su EW, Lin JY, Kane LP (2008) TIM-1 and TIM-3 proteins in immune regulation. Cytokine 44:9–13

Sugimoto C, Watanabe S, Naruse T, Kajiwara E, Shiino T, Umano N, Ueda K, Sato H, Ohgimoto S, Hirsh V, Villinger F, Ansari AA, Kimura A, Miyazawa M, Suzuki Y, Yamamoto N, Nagai Y, Mori K (2010) Protection of macaques with diverse MHC genotypes against a heterologous SIV by vaccination with a deglycosylated live-attenuated SIV. PLoS One 5:e11678

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731–2739

Uccelli A, Giunti D, Capello E, Roccatagliata L, Mancardi GL (2003) EAE in the common marmoset *Callithrix jacchus*. Int MS J 10:6–12

Wichukchinda N, Nakajima T, Saipradit N, Nakayama EE, Ohtani H, Rojanawiwat A, Pathipvanich P, Ariyoshi K, Sawanpanyalert P, Shioda T, Kimura A (2010) TIM1 haplotype may control the disease progression to AIDS in a HIV-1-infected female cohort in Thailand. AIDS 24:1625–1631

Wong WS, Yang Z, Goldman N, Nielsen R (2004) Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. Genetics 168:1041–1051

Yang Z (2005) The power of phylogenetic comparison in revealing protein function. Proc Natl Acad Sci U S A 102:3179–3180

Yang Z, Nielsen R (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol Biol Evol 17:32–43

Yang Z, Wong WS, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol 22:1107–1118

Zhang J, Rosenberg HF, Nei M (1998) Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc Natl Acad Sci U S A 95:3708–3713

Development of a Novel *In Silico* Docking Simulation Model for the Fine HIV-1 Cytotoxic T Lymphocyte Epitope Mapping

Masahiko Mori[1,2], Kei Matsuki[1], Tomoyuki Maekawa[1], Mari Tanaka[1], Busarawan Sriwanthana[3], Masaru Yokoyama[4], Koya Ariyoshi[1]*

1 Institute of Tropical Medicine, Nagasaki University, Nagasaki City, Nagasaki, Japan, **2** Department of Paediatrics, The Peter Medawar Building for Pathogen Research, University of Oxford, Oxford, United Kingdom, **3** Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand, **4** Pathogen Genomics Center, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo, Japan

Abstract

Introduction: Class I HLA's polymorphism has hampered CTL epitope mapping with laborious experiments. Objectives are 1) to evaluate the novel *in silico* model in predicting previously reported epitopes in comparison with existing program, and 2) to apply the model to predict optimal epitopes with HLA using experimental results.

Materials and Methods: We have developed a novel *in silico* epitope prediction method, based on HLA crystal structure and a peptide docking simulation model, calculating the peptide-HLA binding affinity at four amino acid residues in each terminal. It was applied to predict 52 HIV best-defined CTL epitopes from 15-mer overlapping peptides, and its predictive ability was compared with the HLA binding motif-based program of HLArestrictor. It was then used to predict HIV-1 Gag optimal epitopes from previous ELISpot results.

Results: 43/52 (82.7%) epitopes were detected by the novel model, whereas 37 (71.2%) by HLArestrictor. We also found a significant reduction in epitope detection rates for longer epitopes in HLArestrictor ($p = 0.027$), but not in the novel model. Improved epitope prediction was also found by introducing both models, especially in specificity ($p<0.001$). Eight peptides were predicted as novel, immunodominant epitopes in both models.

Discussion: This novel model can predict optimal CTL epitopes, which were not detected by an existing program. This model is potentially useful not only for narrowing down optimal epitopes, but predicting rare HLA alleles with less information. By introducing different principal models, epitope prediction will be more precise.

Citation: Mori M, Matsuki K, Maekawa T, Tanaka M, Sriwanthana B, et al. (2012) Development of a Novel *In Silico* Docking Simulation Model for the Fine HIV-1 Cytotoxic T Lymphocyte Epitope Mapping. PLoS ONE 7(7): e41703. doi:10.1371/journal.pone.0041703

Editor: Jianming Tang, University of Alabama at Birmingham, United States of America

Received April 24, 2012; **Accepted** June 25, 2012; **Published** July 27, 2012

Funding: This study was financially supported by Japan foundation for AIDS prevention (JFAP), and the Ministry of Health, Labour and Welfare of Japan and the Grand-in-Aid for Scientific Research from the Japan Society for the Promotion of Science. 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) play a crucial role in HIV replication control by eliminating virus-infected cells by recognizing class I Human Leukocyte Antigen (HLA) molecule-viral peptides (= epitope) complex. This response is thought to be a major determinant of the viral set point, and consequent disease progression [1]. However the efficacy of the CTL response is affected by the extent of polymorphisms in HLA loci and viral sequences. The HLA region is found on chromosome 6 and is the most polymorphic loci in the human genome [2]; each individual expresses up to six different class I alleles out of a vast pool of allelic variants, the reported number of which reaches 5,399 for class I HLA molecules (1,757 of HLA-A, 2,338 of HLA-B, and 1,304 of HLA-C alleles) [3]. In addition, the extensive diversity of HIV-1 owing to its extreme capacity to mutate has led to a reported 13 prototype clades and 43 circulating recombinant forms (CRFs) [4]. Despite such HLA polymorphism and HIV viral diversity environment, recent genome wide association study (GWAS) reported the best contribution of class I HLA for viral control, suggesting the importance of CTL epitope mapping with responsible HLA information [5]. Several major HIV-1 epitopes and their restricting HLA alleles have been defined through fine epitope mapping; 1,344 epitopes and their restricting HLA alleles have been reported as of February 2012 (CTL Epitopes. Los Alamos National Lab. http://www.hiv.lanl.gov/). The limitation of the dataset currently available however, is that the majority of these epitope/HLA combinations are derived from subtype B-infected Caucasians or C-infected Africans, and epitope information from other subtypes or ethnicities is rare.

The traditional, *in vitro* method of epitope detection involves using a matrix of overlapping peptides (OLPs) encoding viral proteins in Enzyme-Linked Immunospot (ELISpot) assays to

identify a single candidate peptide, from which the 8-11mer epitope is mapped down. This is typically followed by the confirmation of the restricting HLA alleles using tetramers or in a ^{51}Cr release assay using peptide-specific lines [6,7]. It is a difficult and labor-intensive process, particularly time-consuming in the case of epitopes restricted by rare HLA alleles because of the limited number of samples available.

Recently, alternative, *in silico* models for epitope prediction have been developed [8]. These can broadly be divided into two models; the first is an algorithm based on the peptide-binding motif, and the second is a structural algorithm model based on the crystal structure of HLA molecules. The former is characterized by the use of motif matrices deduced from refined motifs based on the pool sequence, enlisting optimal amino acid sequences at anchor positions in specific HLA alleles. An example of such an algorithm is the SYFPEITHI [9] database, which predicts the HLA-binding affinities of peptides by ranking them according to the presence of primary and secondary anchor amino acids. However these models are based on reported epitopes and their restricting HLA alleles, so their predictions are powerful in the context of well-published HLA alleles but not suitable against rare or novel alleles with little previous information. Another model of epitope prediction is the binding affinity model, which calculates the peptides' binding affinity and scores it using quantitative matrices (QMs), a well-known example being the NetMHC [10,11] or the HLArestrictor [12]. This model scores binding strength as binding affinity with thresholds to differentiate strong binding peptides and weak ones in each calculation.

On the other hand, the structural algorithm model does not require binding motif information, which is advantageous for the definition of epitopes restricted by HLA alleles with less published epitope information. Recently, a docking simulation model (DSM) which takes into consideration binding energy such as electrostatic interactions and van der Waals (vdw) interactions, together with the crystal structure of HLA alleles, has been developed [13–17].

Our objectives here are 1) to evaluate the novel *in silico* DSM in predicting previously reported best-defined epitopes in comparison with existing binding motif-based program, and 2) to apply the model to predict optimal size of the epitopes and restricting HLA alleles using results obtained from our previous study in a HIV-1 CRF01_AE-infected Thai cohort.

Materials and Methods

Ethic Statement

This study was approved by Thai Ministry of Public Health Ethics Committee. Written informed consent was obtained from all patients after explaining the purpose and expected consequences of the study.

Computational program and calculation

We used the commercial softwares Molecular Operating Environment® (MOE) (CCG Inc., Montreal, Canada) and MOE-ASEDock® (Ryoka System Inc., Tokyo, Japan) for the molecular binding affinity calculation [18]. HLA's 3D models were obtained from the X-ray crystallography database in MOE's library (1OGA for HLA-A*02:01, IQ94 for HLA-A*11:01, 2BCK for HLA-A*24:02, 1XR9 for HLA-B*15:01, 1JGE for HLA-B*27:05, 2CIK for HLA-B*35:01, 1E27 for HLA-B*52:01, 2RFX for HLA-B*57:01, and 1EFX for HLA-C*03:04). In cases where the original X-ray crystallography information was unavailable, we generated a 3D structural model using highly homologous HLA alleles as template, using rotamer explorer or homology modeling to reconstruct their structures by changing sequential difference sites, a method originally used in the point mutation program attached in MOE AMBER99 [19] for force field, calculations. For solvent effect energy calculation, a generalized Born model [20], were introduced. As an indicator of the affinity between epitope candidate peptides and the class I HLA allele, we measured the U_dock score [U_ele (electric energy)+U_vdw (van der Waals energy)+U_solv (Solvation energy)+U_strain (Strain energy)] (kcal/mol) [18]. We calculated the U_dock score of four residues at each N- and C-terminal, spanning the anchor position at each of the terminals, and scored the sum of them as binding affinity. A lower score indicates a higher affinity between the HLA molecule and peptides.

Evaluation of the novel DSM through an analysis of best-defined HIV CTL epitopes and their restricting HLA alleles

For the quality evaluation of this novel program, we first calculated the U_dock score for 52 best-defined HIV epitopes restricted by the alleles HLA-A*02:01, HLA-A*11:01, HLA-A*24:02, HLA-B*15:01, HLA-B*27:05, HLA-B*35:01 and HLA-B*57:01 as enlisted in Los Alamos database (CTL Epitopes. Los Alamos National Lab. http://www.hiv.lanl.gov/). We calculated the U_dock score between the restricting HLA alleles and the 8 to 11-mer peptides within 15-mer peptides of the viral strain HXB2, in which best-defined epitopes were included. 26 variants of 8 to 11-mer peptides were calculated in one HLA and 15-mer peptide combination, then the lowest U_dock score was ranked as the 1st and the highest score as the 26th in each calculation (Figure 1). Combinations that ranked within the top five were regarded as positive. In parallel with our DSM, we also performed epitope prediction using the latest artificial neural network (ANN) model, the HLArestrictor [12], using the affinity thresholds of Strong Binder (SB), Weak Binder (WB), Combined Binder (CB) and Non-binder (NB), according to their definitions.

We evaluated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for each best-defined epitope prediction using the DSM, HLArestrictor, as well as those defined as dual positive by both models.

Analysis of *in vitro* HIV-1 CRF01_AE Gag epitope candidates by using both *in silico* epitope prediction models

We then applied both the DSM and the HLArestrictor to predict the optimal size of epitopes, based on results obtained from our previous study [21], in which 31 candidate epitopes were detected by ELISpot assays using Gag 15-mer OLPs and their HLA associations detected by Fisher's exact test in a cohort of 137 (107 female and 30 male) HIV-1 CRF01_AE-infected Thais. All were chronically infected and treatment naïve, with median 461/ul CD4+T cell count (range 204–1,191) and 4.2 log copies/ml viral load (2.6–5.9).

Epitope prediction for the immunogenic Gag OLP $p24_{276–285}$ MYSPVSILDI using a ^{51}Cr release assay and both *in silico* models

In our previous study [21], the 15-mer peptides Gag $p24_{271–285}$ NKIVRMYSPVSILDI (NI15) and $p24_{276–290}$ MYSPVSILDIRQGPK (MK15) induced the largest responses in terms of both breadth and magnitude, and were statistically associated with the alleles HLA-A*02:07, HLA-B*46:01, and HLA-C*01:02, which were under linkage disequilibrium (LD) association [21]. Presuming that the optimal epitope resides in the overlapping amino acid sequence between NI15 and MK15, that is, $p24_{276–285}$

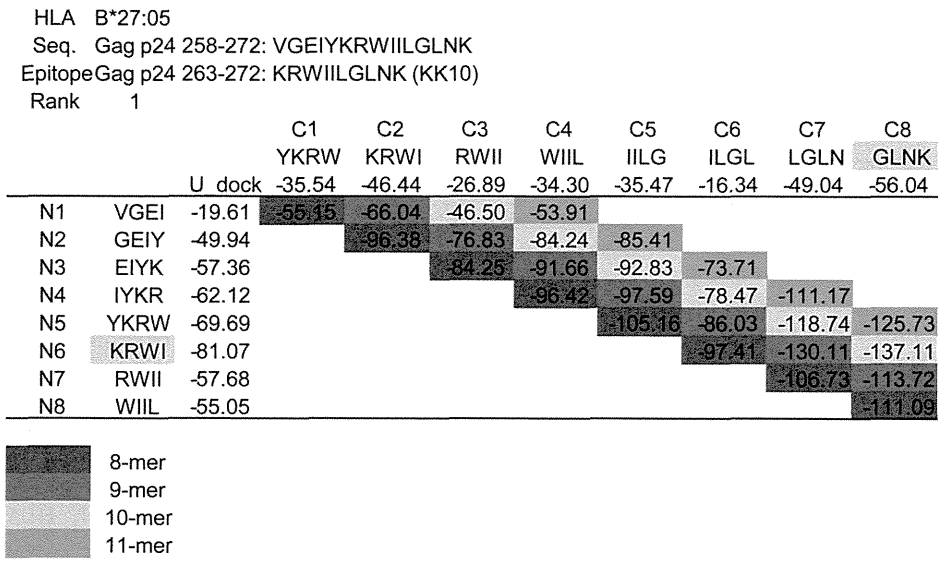


Figure 1. Example of epitope prediction using the novel *in silico* docking simulation model. U_dock scores of the N-terminal (Row N1–N8) and C-terminal (Column C1–C8) was calculated and their sum was scored as the U_dock score (kcal/mol) of each 8 to 11-mer peptide's. The lower score indicated stronger binding between the peptide and HLA. In this example, Gag $p24_{263-272}$ KRWIILGLNK (KK10), well-known as one of the best-defined epitopes, scored −137.11 kcal/mol against HLA-B*27:05 and was the lowest (ranked as the 1st) among 26 variants in 15-mer peptide of Gag $p24_{258-272}$ VGEIYKRWIILGLNK.
doi:10.1371/journal.pone.0041703.g001

MYSPVSILDI (MI10), we conducted a ^{51}Cr release assay as previously described [22].

Results

Prediction of best-defined epitopes by the DSM and the peptide binding motif model

We have evaluated the predictive power of our DSM by testing its ability to predict epitopes within 52 15-mer peptides spanning the epitopes for seven HLA alleles enlisted in the Los Alamos database as best-defined epitopes. Overall, DSM ranked 43/52 (82.7%) of the best-defined epitopes correctly within the top five candidates, within which 14 epitopes ranked as the 1st, 11 as the 2nd, 7 as the 3rd, 3 as the 4th, then 8 as the 5th (Table S1). This was comparable to the HLArestrictor, where 37/52 (71.2%, 43/52 vs 37/52, $p = 0.24$ by Fisher's exact test) best-defined epitopes scored within the threshold of binding affinity without having 4 or more other candidate epitopes: 20 as SB, 10 as WB and 7 as CB. Table 1 summarizes the performance on epitope prediction by each model and dual positives by both models, according to their sensitivity, specificity, PPV and NPV. The performance of the DSM is similar to that of HLArestrictor. Interestingly, by introducing both models, specificity increased with significance ($p<0.001$), and an additive effect was seen in the PPV. We believe this is the first study to report a structure-based epitope prediction model with comparable or greater predictive power than a peptide-binding motif based model.

32/52 (61.5%) epitopes were detected as a significant epitope candidate by both models. 11/52 (21.2%) epitopes were detected only by the DSM, while 5/52 (9.6%) were detected only by HLArestrictor. 4/52 (7.7%) epitopes were not detected by either methods. Within the 14 epitopes not correctly predicted by HLArestrictor, incorrect epitopes were predicted in 7 epitopes. It is noteworthy that two epitopes, Nef_{75-82} PLRPMTYK (PK8) restricted by HLA-A*11:01 and $Nef_{117-127}$ TQGYFPDWQNY (TY11) restricted by HLA-B*15:01 were detected as a NB by HLArestrictor, whereas they were ranked as the 2nd in PK8 and the 1st in TY11 in the DSM. $Integrase_{179-188}$ AVFIHNFKRK (AK10) restricted by HLA-A*11:01 was predicted as a SB, but because there were 5 other SB candidates, 3 WB candidates and 1 CB candidate, this prediction was regarded as failure.

A striking feature of the DSM was that it had a high detection rate of best-defined epitopes independent of the peptide's length. The prediction rate of shorter epitopes (8 and 9-mer) was 27/31 (87.1%) while the rate for longer epitopes (10 and 11-mer) was 16/21 (76.2%), between which we found no significant difference by Fisher's exact test ($p = 0.46$). In contrast, the ability of HLArestrictor to accurately predict best-defined epitopes was highly dependent on epitope length, as the prediction rate of longer

Table 1. Evaluation of best-defined epitope prediction among docking simulation model, HLArestrictor, and positives in dual models.

	DSM	HLArestrictor	Dual positives	p (mxn Fisher's exact test)
Sensitivity	0.83	0.71	0.62	0.056
Specificity	0.83	0.94	0.97	<0.001
PPV	0.17	0.31	0.43	0.095
NPV	0.99	0.99	0.98	0.46

Evaluation of best-defined epitope prediction among each model and positives in dual models were statistically evaluated, according to their sensitivity, specificity, positive prediction value (PPV) and negative prediction value (NPV) by maximum Fisher's exact test. DSM: Docking simulation model.
doi:10.1371/journal.pone.0041703.t001

epitopes (11/21, 52.3%) was significantly lower than that of shorter ones (26/31, 83.9%) ($p = 0.027$).

Successful prediction with the DSM was dependent on the HLA allele and its peptides: in HLA-B*15:01, HLA-B*27:05 and HLA-B*35:01, all of the best-defined epitopes were ranked within the top 5th. However, four best-defined epitopes restricted by HLA-B*57:01 and HLA-A*02:01 scored within the worst 5th candidates: $Nef_{120-128}$ YFPDWQNYT, $p15_{433-442}$ FLGKIWPSYK, RT_{33-41} ALVEICTEM, and $p24_{161-172}$ KAFSPEVIPMF.

Optimal epitope prediction to analyze HIV-1 CRF01_AE Gag ELISpot assay data using two *in silico* models

We next applied the model to predict optimal epitopes against HIV-1 CRF01_AE Gag based on our previously obtained results in a Thai HIV cohort study [21]. In total, 31 peptide-HLA associations were analyzed: 5 in HLA-A, 13 in HLA-B, and 13 in HLA-C (Table S2). Among these, 10 overlapping peptides spanned previously reported epitopes (6 were best-defined epitopes and 4 were published but not enlisted as best-defined epitopes). In the DSM, 9/10 (90%) reported epitopes were successfully ranked within the 5th as significant epitope candidates, and all of the six best-defined epitopes ranked either the 1st or 2nd. In HLArestrictor, 8/10 (80%) epitopes were predicted as significant binders; 3 as SB, 4 as WB, and 1 as CB, but 2 epitopes (best-defined epitopes HLA-A*02:07-restircted YL9, and HLA-B15-restricted KL9) were not predicted as significant binders. HLArestrictor also predicted another 16 sequences as potential epitope candidates: 1 as SB, 12 as WB, and 2 as CB. Intriguingly only one WB candidate was ranked within the top five by the DSM, reflecting a considerable degree of discrepancy between the two prediction methods.

8 previously unreported peptides were predicted by both models: HLA-B*38:02-restricted $p24_{198-205}$ MQMLKETI (rank 1st in DSM and WB in HLArestrictor), HLA-B*40:01-restricted $p24_{311-321}$ QEVKNWMTETL (2nd and SB), HLA-B*46:01-restricted $p24_{275-283}$ RMYSPVVSIL (5th and SB), HLA-B*58:01-restircted $p17_{79-86}$ YNTVVTLW (1st and WB), HLA-B*58:01-restricted $p17_{77-86}$ SLYNTVVTLW (4th and WB), HLA-C*01:02-restricted $p24_{277-285}$ YSPVSILDI (2nd and WB in $p24_{271-285}$ and 3rd and WB in $p24_{276-290}$), HLA-C*01:02-restricted $p24_{276-285}$ MYSPVSILDI (4th and WB both in $p24_{271-285}$ and $p24_{276-290}$), and HLA-C*01:02-restricted $p24_{296-304}$ YVDRFYKTL (1st and WB).

Application of the *in silico* DSM to define the restricting HLA molecule

We conducted a ^{51}Cr release assay with a truncated peptide titration spanning the overlapping region between Gag $p24_{271-285}$ NKIVRMYSPVSILDI (NI15) and $p24_{276-290}$ MYSPVSILDIRQGPK (MK15). These induced the largest responses both in breadth and magnitude in our previous study, and were statistically associated with HLA-A*02:07, HLA-B*46:01, and HLA-C*01:02, which we calculated to be under LD association [21]. We found strong killing against HLA-B*46:01 and HLA-C*01:02-matched $p24_{276-285}$ MYSPVSILDI (MI10)- and $p24_{277-285}$ YSPVSILDI (YI9)-pulsed target cells but not in any other condition (Figure S1). However, we could not further specify the restricting HLA molecule because a single HLA-matched target cell was not available due to the strong LD between them. Therefore, we conducted *in silico* analysis in order to identify the responsible HLA. Table 2 shows the results of the DSM between these two peptides (MI10 and YI9) and three candidate HLA alleles (HLA-A*02:07, HLA-B*46:01 and HLA-C*01:02). Firstly, with the DSM, none of these two peptides were predicted within the top five candidate epitopes when binding to HLA-A*02:07 or HLA-B*46:01, and neither scored significant binding using the HLArestrictor, eliminating these as the restricting HLA molecules. However in the model with HLA-C*01:02, both two peptides ranked within the 5th; MI10 ranked as the 3rd in NI15 and the 4th in MK15, while YI9 was ranked as the 2nd in NI15 and the 3rd in MK15. Significant binding affinity of MI10 and YI9 to HLA-C*01:02 was also predicted by HLArestrictor. Secondly, in the binding motif of HLA-C*01:02 (x[AL][P]xxxxx[L]), both MI10 and YI9 encoded compatible or similar hydrophobic amino acids with the binding motif x[Y]xxxxxxx[I] in MI10 and xx[P]xxxxx[I] in YI9. Together, these results indicate that the optimal epitopes MI10 and YI9 are equally likely candidates recognized by HLA-C*01:02, with YI9 ranking slightly higher in the DSM.

Discussion

In this study, we demonstrated that the structure-based DSM can predict the peptide binding affinity with various HLA molecules, independently of peptide binding motif information. To our knowledge, this novel DSM is the first model of its kind that succeeded in predicting HIV-1 CTL best-defined epitopes, with better or at least equivalent accuracy to the latest binding motif-based program. We also found a high detection rate of best-defined epitopes independent of peptide size in the DSM, while the detection rate significantly decreased with longer epitopes in the other model.

Historically, comparisons of epitope prediction methods has generally shown that peptide-binding motif based methods outperform structure-based methods [23]. However, the increased availability of crystal structures of MHC-peptide complexes is enabling the development of prediction methods using such structural models and the calculation of free energy of binding [23,24]. In the review by Liao *et al* [23], their comprehensive comparison of structure-based models and peptide-binding motif models in epitope prediction showed that the structure-based model was able to outperform all other methods except the ANN model, which performed equally well. In our novel program, we use a measure of the binding affinity between the HLA molecule and the peptides at four residues spanning the N- and C-terminal.

Table 2. Prediction of the HLA restriction of Gag $p24_{276-285}$ MYSPVSILDI (MI10) and $p24_{277-285}$ YSPVSILDI (YI9) using *in silico* methods.

			U_dock rank		
HLA	**Binding motif**	**Peptide**	**NI15**	**MK15**	**HLArestrictor**
A*02:07	x[L][D]xxxxx[L]	MI10	13	13	
		YI9	14	16	
B*46:01	x[M(I)]xxxxxx[YF]	MI10	15	20	
		YI9	19	21	
C*01:02	x[AL][P]xxxxxx[L]	MI10	3	4	WB
		YI9	2	3	SB

HLA restriction prediction against two reactive Gag peptides, Gag $p24_{276-285}$ MYSPVSILDI (MI10) and $p24_{277-285}$ YSPVSILDI (YI9) was performed by the docking simulation model, and the binding motif HLArestrictor 1.2. The U_dock rank by the docking simulation model against MI10 and YI9 was analyzed in the original 15-mer peptides of Gag $p24_{271-285}$ NKIVRMYSPVSILDI (NI15) and $p24_{276-290}$ MYSPVSILDIRQGPK (MK15). SB: Strong Binder, WB: Weak Binder.
doi:10.1371/journal.pone.0041703.t002

This covers not only the anchor position sites but also their flanking sites, which have a considerable effect on peptide-HLA binding; this may also have led to the high detection rate of best-defined epitopes independent of epitope size. Together with precise HLA crystal structure information, we have also incorporated a fine calculation model for binding affinity [18], giving the DSM a high detection rate of best-defined epitopes equivalent to that of the latest binding motif-based program.

Intriguingly there was a considerable degree of discrepancy between the two methods: 21.2% of the 52 best-defined epitopes were detected as significant epitope candidates only by the DSM, while 9.6% was detected only by the HLArestrictor. Furthermore, two epitopes which ranked within the bottom five by DSM were successfully predicted as a single candidate by HLArestrictor, whereas five epitopes which were not detected by HLArestrictor, were successfully predicted as the best candidates by the DSM. This result highlights the importance of combining programs with different approaches, for example those based on peptide binding motif information and those that do not require peptide binding motif information, consistent with previous report in class II HLA peptide binding prediction model [25].

We therefore applied both models to predict optimal epitopes in HIV-1 CRF01_AE Gag and found 8 previously unreported optimal epitopes supported by both models. These potential epitopes need to be further confirmed *ex vivo* that they are true epitopes capable of stimulating T cell responses with either a ^{51}Cr release assay or tetramer assay. However, since the DSM alone predicted 11 other candidates that were not predicted by the HLArestrictor, combining both models would be important to reduce the cost of such experiments. Furthermore a substantial number of OLPs were recognized using an ELISpot assay but within the peptides that induced a response, no epitope was predicted by the HLArestrictor. This DSM would save the cost of experiments by reducing 26 potential candidate peptides to five.

The ability of the DSM model to accurately predict peptides was dependent on the HLA molecule in question, and our results suggest that this is due to variations in the C-terminal binding groove. Four best-defined epitopes restricted by the alleles HLA-A*02:01 and HLA-B*57:01 ranked among the worst from the 22nd up to the 26th in our program. In HLA-A*02:01, both FK10 and AM9 coded Leucine (L) at the 2nd position of sequence, compatible with the HLA-A*02:01 binding motif at the B pocket and scored a low and therefore strongly binding U_dock score at the N-terminal site [−47.8 kcal/mol in FK10 (5th in N1-N8 terminal) and −54.4 kcal/mol in AM9 (2nd)]. However, the sequences did not match with the HLA-A*02:01 binding motif at the C-terminal which contains a Valine (V) at the F pocket, and they scored the worst U_dock scores [−14.1 kcal/mol in KF10 (8th) and −48.5 kcal/mol in AM9 (8th)]. A similarly low score at the C-terminal was also found in HLA-B*57:01-restricted KF11 [−24.5 kcal/mol (8th)] and YT9 [−23.8 kcal/mol (8th)]. The importance of the C-terminal for peptide-binding stability has been previously reported [26], and with respect to structural differences between the B and F pockets, it is generally known that the B pocket has a rather round shape while the F pocket has a deep cleft-like shape, suggesting stricter peptide binding restriction at the F pocket compared to the B pocket among HLA-A*02:01 and HLA-B*57:01. In contrast, HLA-B*27:05 and HLA-B*35:01 had none or only one variant of their binding motif at C-terminal: x[R(K)]xxxxxxx or x[R]xxxxxx[LFYRHK(MI)] in HLA-B*27:05 and x[P(AV)]xxxxxxx or x[P(AVYRD)]xxxxxx[YFMLI] in HLA-B*35:01. In these two alleles, all of the best-defined epitopes ranked within the 5th. These results strongly suggest that the diversity of peptide binding at the F pocket defines the accuracy or difficulty of epitope prediction by DSM.

Recent studies have highlighted the importance of HLA-C alleles for HIV viral control, for instance in the population-based study from Africa [27], existence of dominant HLA-C*04-restricted epitopes [28], stimulation of NK cells through HLA-C and Killer-cell Immunoglobulin-like receptors (KIRs) [29,30], and HLA-C expression control by 35 kb upstream genotype of HLA-C allele and HIV viral control [31]. However, epitope mapping of HLA-C antigens has been held back for several reasons. Firstly, in *in vitro* studies it has been difficult to find target and effector cell combinations with singly matched HLA alleles which are not under LD association, as we found in our ^{51}Cr release assay. *In silico*, in contrast to HLA-A or B alleles, epitope prediction programs against HLA-C alleles have been sparse [9–11]. This can be attributed to the lack of reported epitopes information from HLA-C alleles, since binding motif-based models were originally programmed based on such reported data. Furthermore, LD of HLA-C alleles, especially with HLA-B alleles, hinders the confirmation of HLA-C alleles as the restricting alleles in statistical analyses. In our previous study, among 13 HLA-C-associated epitope candidates, nine were reported with HLA-A or B alleles which were under LD association [21]. Novel DSM could contribute to epitope detection by bypassing such obstacles to epitope prediction against HLA-C alleles.

This study had several limitations. First, we could not define the threshold of the U_dock score degree itself in novel program as defined in HLArestrictor. Related with this limitation, considering the HLA polymorphism, reported epitope number, and comparison between alleles with/without original crystal structure information, further calculations will be warranted for the quality evaluation of DSM. Second, this is a computational epitope prediction model whose algorithm is solely based on the binding between the peptide and the HLA molecule. Although peptide-HLA binding is the most selective event for epitope determination [32], CTL activation is a multi-step process involving the processing of viral peptides by proteasome [22,33,34] and the recognition of the peptide-HLA complex by T cell receptors (TCRs) [35], both of which are not accounted for in the model.

In conclusion, we have shown here a novel *in silico* DSM which can be used for epitope mapping, and combined with a binding motif-based model, this will significantly reduce the required experimental burden for epitope identification in the development of a CTL-based vaccine for HIV.

Supporting Information

Figure S1 Identification of HLA-B*46:01/C*01:02-restricted Gag p24$_{276–285}$ MI10 and p24$_{277–285}$ YI9 by a ^{51}Cr release assay. ^{51}Cr release assays under HLA-B*46:01/C*01:02-matched conditions were performed for each peptide. Significant % lysis was found in target cells pulsed with Gag p24$_{276–285}$ MI10: MYSPVSILDI and p24$_{277–285}$ YI9: YSPVSILDI.
(PPTX)

Table S1 Predicted best-defined epitopes using the docking simulation model and a comparison with HLArestrictor. The docking simulation model was applied to predict epitopes within 15-mer peptides spanning best-defined epitopes and compared with those predicted with the HLArestrictor. The U_dock score and their rank were calculated for each peptide in the docking simulation model, while with HLArestrictor the affinity thresholds of SB: Strong Binder, WB: Weak Binder, and CB: Combined Binder, and Non-binder were given, according to their definitions.
(XLS)

Table S2 Epitope prediction using the docking simulation model and HLArestrictor against in vitro HLA-restricted HIV-1 CRF01_AE Gag epitope candidates. Using previously reported HIV-1 CRF01_AE Gag epitope candidates detected by ELISpot assays and statistical analysis, epitope prediction was performed by our novel docking simulation model and HLArestrictor. Among 31 15-mer peptide and HLA associations, six best-defined epitopes and four non-best defined epitopes were included. Bold, underlined sequences indicate positive candidates in dual models. SB: Strong Binder, WB: Weak Binder, and CB: Combined Binder.
(XLSX)

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 the Lampang hospital.

Author Contributions

Conceived and designed the experiments: TM MY KA. Performed the experiments: MM KM TM MT BS. Analyzed the data: MM KM TM. Contributed reagents/materials/analysis tools: TM MY KA. Wrote the paper: MM KA.

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. Mungall AJ, Palmer SA, Sims SK, Edwards CA, Ashurst JL, et al. (2003) The DNA sequence and analysis of human chromosome 6. Nature 425: 805–11.
3. Robinson J, Mistry K, McWilliam H, Lopez R, Parham P (2011) The IMGT/HLA Database Nucleic Acids Research 39 (Suppl 1): D1171–6.
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. Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, et al. (2010) The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 330: 1551–1557.
6. Draenert R, Altfeld M, Brander C, Basgoz N, Corcoran C, et al. (2003) Comparison of overlapping peptide sets for detection of antiviral CD8 and CD4 T cell responses. J Immunol Methods 275: 19–29.
7. Streeck H, Frahm N, Walker BD (2009) The role of IFN-gamma Elispot assay in HIV vaccine research. Nat Protoc 4: 461–9.
8. Lafuente EM, Reche PA (2009) Prediction of MHC-peptide binding: a systematic and comprehensive overview. Curr Pharm Des 15: 3209–20.
9. Rammensee HG, Bachmann J, Emmerich NPN, Bacho OA, Stevanovic S (1999) SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50: 213–9.
10. Buus S, Lauemoller SL, Worning P, Kesmir C, Frimurer T, et al. (2003) Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. Tissue Antigens 62: 378–84.
11. Lundegaard C, Lund O, Nielsen M (2008) Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. Bioinformatics 24: 1397–98.
12. Erup LM, Kloverpris H, Stryhn A, Koofhethile CK, Sims S, et al. (2011) HLArestrictor–a tool for patient-specific predictions of HLA restriction elements and optimal epitopes within peptides. Immunogenetics 63: 43–55.
13. Tong JC, Tan TW, Ranganathan S (2004) Modeling the structure of bound peptide ligands to major histocompatibility complex. Protein Sci 13: 2523–32.
14. Bui HH, Schiewe AJ, von Grafenstein H, Haworth IS (2006) Structural prediction of peptides binding to MHC class I molecules. Proteins 63: 43–52.
15. Fagerberg T, Cerottini JC, Michielin O (2006) Structural prediction of peptides bound to MHC class I. J Mol Biol 356: 521–46.
16. Knapp B, Omasits U, Frantal S, Schreiner W (2009) A critical crossvalidation of high throughput structural binding prediction methods for pMHC. J Comput Aided Mol Des 5: 301–7.
17. Bordner AJ, Abagyan R (2006) Ab initio prediction of peptide-MHC binding geometry for diverse class I MHC allotypes. Proteins 63: 512–26.
18. Goto J, Kataoka R, Muta H, Hirayama N (2008) ASEDock-docking based on alpha spheres and excluded volumes. J Chem Inf Model 48: 583–90.
19. Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comp Chem 21:1049–1074.
20. Kabsch W (1976) A solution for the best rotation to relate two sets of vectors. Acta Crystallogr Sect F Struct Biol Cryst Commun 32: 922–3.
21. 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: e22680.
22. 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.
23. Liao WW, Arthur JW (2011) Predicting peptide binding to Major Histocompatibility Complex molecules. Autoimmun Rev 10: 469–73.
24. Jojic N, Reyes-Gomez M, Heckerman D, Kadie C, Schueler-Furman O (2006) Learning MHC I-peptide binding. Bioinformatics 22: e227–35.
25. Wang P, Sidney J, Dow C, Mothe B, Sette A, et al. (2008) A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol 4: e1000048.
26. Bouvier M, Wiley DC (1994) Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. Science 265: 398–402.
27. Leslie A, Matthews PC, Listgarten J, Carlson JM, Kadie C, et al. (2010) Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. J Virol 84: 9879–88.
28. Makadzange AT, Gillespie G, Dong T, Kiama P, Bwayo J, et al. (2010) Characterization of an HLA_C-restricted CTL response in chronic HIV infection. Eur J Immunol 40: 1036–41.
29. Jennes W, Verheyden S, Demanet C, Adjé-Touré CA, Vuylsteke B, et al. (2006) Cutting edge: resistance to HIV-1 infection among African female sex workers is associated with inhibitory KIR in the absence of their HLA ligands. J Immunol 177: 6588–92.
30. Ravet S, Scott-Algara D, Bonnet E, Tran HK, Tran T, et al. (2007) Distinctive NK-cell receptor repertoires sustain high-level constitutive NK-cell activation in HIV-exposed uninfected individuals. Blood 109: 4296–305.
31. Thomas R, Apps R, Qi Y, Gao X, Male V, et al. (2009) HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nat Genet 41: 1290–4.
32. Jensen PE (2007) Recent advances in antigen processing and presentation. Nat Immunol 8: 1041–8.
33. Tenzer S, Wee E, Burgevin A, Stewart-Jones G, Friis L, et al. (2009) Antigen processing influences HIV-specific cytotoxic T lymphocyte immunodominance. Nat Immunol 10: 636–46.
34. Ranasinghe SR, Kramer HB, Wright C, Kessler BM, di Gleria K, et al. (2011) The antiviral efficacy of HIV-specific CD8(+) T-cells to a conserved epitope is heavily dependent on the infecting HIV-1 isolate. PLoS Pathog 7: e1001341.
35. 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.

The Effect of HLA Polymorphisms on the Recognition of Gag Epitopes in HIV-1 CRF01_AE Infection

Busarawan Sriwanthana[1], Masahiko Mori[2,3], Mari Tanaka[2], Sei Nishimura[2], Toshiyuki Miura[4], Panita Pathipvanich[5], Pathom Sawanpanyalert[1], Koya Ariyoshi[2]*

1 Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand, **2** Institute of Tropical Medicine, Nagasaki University, Nagasaki city, Nagasaki, Japan, **3** Department of Paediatrics, The Peter Medawar Building for Pathogen Research, University of Oxford, Oxford, United Kingdom, **4** Advanced Clinical Research Centre, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo, Japan, **5** Day Care Centre, Lampang Hospital, Lampang, Thailand

Abstract

Introduction: The design of a globally effective vaccine rests on the identification of epitopes capable of eliciting effective cytotoxic T lymphocyte (CTL) responses across multiple HIV clades in different populations. This study aims to discern the effect of HLA polymorphisms and the cross-clade reactivity or clade-specificity of epitopes in Thailand where HIV-1 CRF01_AE is circulating.

Materials and Methods: 14 peptides based on consensus HIV-1 CRF01_AE amino acid sequences were designed for use in IFN-γ ELISpot assays and ^{51}Cr release assays among 66 HIV-1 CRF01_AE-infected Thai patients. For ELISpot responders carrying HLA alleles currently unknown to restrict CRF01_AE epitopes, *in silico* epitope-HLA prediction was performed.

Results: 29/66 (43.9%) patients recognized at least one peptide. In total 79 responses were seen against all 14 peptides. 28/79 (35.4%) of the responses were in patients with HLA alleles previously reported to restrict CRF01_AE epitopes, 24/79 (30.4%) responses were in individuals with HLA alleles previously reported to restrict epitopes of HIV clades other than CRF01_AE, and the remaining 27/79 (34.2%) responses were not associated with HLA alleles previously known to restrict HIV epitopes. *In silico* epitope prediction detected 19 novel, epitope-HLA combinations, and 11/19 (57.9%) were associated with HLA-C alleles. We further confirmed a novel HLA restriction of a previously identified HIV-1 Gag epitope [p24$_{122-130}$: PPIPVGDIY (PY9)] by HLA-B*40:01 with a standard ^{51}Cr release assay.

Discussion: CTL recognition sites in HIV-1 Gag were similar among different clades but the HLA restriction differed in Thai patients. This disparity in HLA restriction along different populations illustrated the importance of clade- and population-specific HLA analysis prior to CTL vaccine design.

Citation: Sriwanthana B, Mori M, Tanaka M, Nishimura S, Miura T, et al. (2012) The Effect of HLA Polymorphisms on the Recognition of Gag Epitopes in HIV-1 CRF01_AE Infection. PLoS ONE 7(7): e41696. doi:10.1371/journal.pone.0041696

Editor: Jianming Tang, University of Alabama at Birmingham, United States of America

Received March 29, 2012; **Accepted** June 25, 2012; **Published** July 27, 2012

Funding: This study was financially supported by Japan foundation for AIDS prevention, and the Ministry of Health, Labour and Welfare of Japan and the Grand-in-Aid for Scientific Research from the Japan Society for the Promotion of Science. 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

The control of the Human Immunodeficiency Virus type 1 (HIV-1) epidemic requires the design of a globally effective HIV vaccine. However, the sequence diversity of HIV across clades and the host's human leukocyte antigen (HLA) polymorphism poses a major challenge in the development of a globally effective HIV-1 vaccine capable of inducing cross-clade reactivity [1–3]. The design of polyvalent vaccines aimed at inducing HIV-specific cytotoxic T lymphocyte (CTL) responses has been one of the main focuses in the field of HIV vaccinology for several reasons. Firstly, HIV-specific CTLs play a key role in the control of HIV-1 replication during acute infection and in determining the consequent viral set point [4]. Secondly, studies in macaques have shown that vaccine induced recruitment of Simian Immunodeficiency Virus (SIV)-specific CTLs can effectively control viral replication and slow disease progression [5,6]. Thirdly, unlike neutralizing antibodies, CTLs target proteins such as Gag and Pol, which are relatively conserved across various clades [1,2,7]. Currently, there are 13 prototype HIV clades and 43 circulating recombinant forms (CRF) of HIV-1 group M in the world which are of global importance [2]. However, most immunogenicity studies of the CTL epitopes are conducted in the setting of clade B infection in Caucasian cohorts or clade C infection in African or Indian populations (Epitope Maps, Los Alamos database. http://www.hiv.lanl.gov/), and limited information is available on the immunogenicity of CTL epitopes in the CRF01_AE subtype dominating the epidemic in south-east Asian countries such as Thailand. Here there is a unique class I HLA allele distribution and the prevalence of the highly protective HLA allele B*57 is lower than in other ethnicities; 7%–9% among Africans, 5%–7% among Caucasians, and less than 5% among Asians [1].

HLA polymorphisms can also present a challenge in the design of a vaccine. The HLA loci are the most polymorphic genes in the

human genome [8]. As of February 2012, 1,757 of class I HLA-A, 2,338 of HLA-B, and 1,304 of HLA-C alleles have been reported in the IMGT/HLA database [9]. The pattern of HLA distribution and their influence on clinical progression differs among ethnic groups [1,10–12]. How this divergence across populations plays an effect on the CTL recognition of HIV-1 peptides is not yet fully elucidated, but understanding this is critical for the development of a universal CTL vaccine which delivers protection across various populations.

In the present study, the extent of T cell cross-reactivity to published HIV-1 CRF01_AE sequences in 66 HIV-1 infected Thai patients was evaluated in *ex vivo* ELISpot assays using 14 peptides encoding the Gag protein of the CRF01_AE sequence. The cross-clade specific T cell responses were further elucidated in a standard chromium release assay. We report here that 43% of CRF01_AE infected individuals reacted to at least one peptide of the CRF01_AE sequences that were tested. In this study we aimed to discern the effect of HLA polymorphisms and the cross-clade reactivity or clade-specificity of epitopes among HIV-1 CRF01_AE infected Thai patients, in order to fill in the missing information on epitopes and HLA alleles in Asia.

Materials and Methods

Subjects

This study was approved by Thai Ministry of Public Health Ethics Committee as described elsewhere [13]. Written informed consent was obtained from all patients after explaining the purpose and expected consequences of the study. In case of patients who were school-age, we obtained the written informed consent from their parents as well. 66 HIV-1 CRF01_AE chronically infected patients were recruited at the Lampang hospital, a government referral hospital in northern Thailand. Patients were eligible for inclusion if they were antiretroviral drug naïve at the time of the study. When we attempted to confirm the transmission route, we found one study patient transmitted as an intra-venous drug user (IDU). Thus we excluded this patient from the analysis and corrected the candidate number from 67 to 66. The heterosexual transmission is the predominant mode of HIV transmission in Thailand. It is known that CRF01_AE spread mainly in heterosexually transmitted population. Although rare, subtype B is detected among IDUs [14–16]. In our 66 study patients, the transmission route of all infection was confirmed to be heterosexual by direct interview. Furthermore a part of study patients were confirmed to carry a CRF01_AE virus by direct sequencing [13].

Class I HLA typing

Genomic DNA was extracted from buffy coats using the QIAamp DNA blood Mini Kit (Qiagen, Hilden, Germany) and 4-digit class I HLA typing for A, B and C loci was performed by bead-based array hybridization (WAKFlow HLA typing kit, Wakunaga Pharmaceutical, Hiroshima, Japan) at the Kyoto HLA Laboratory, Kyoto, Japan.

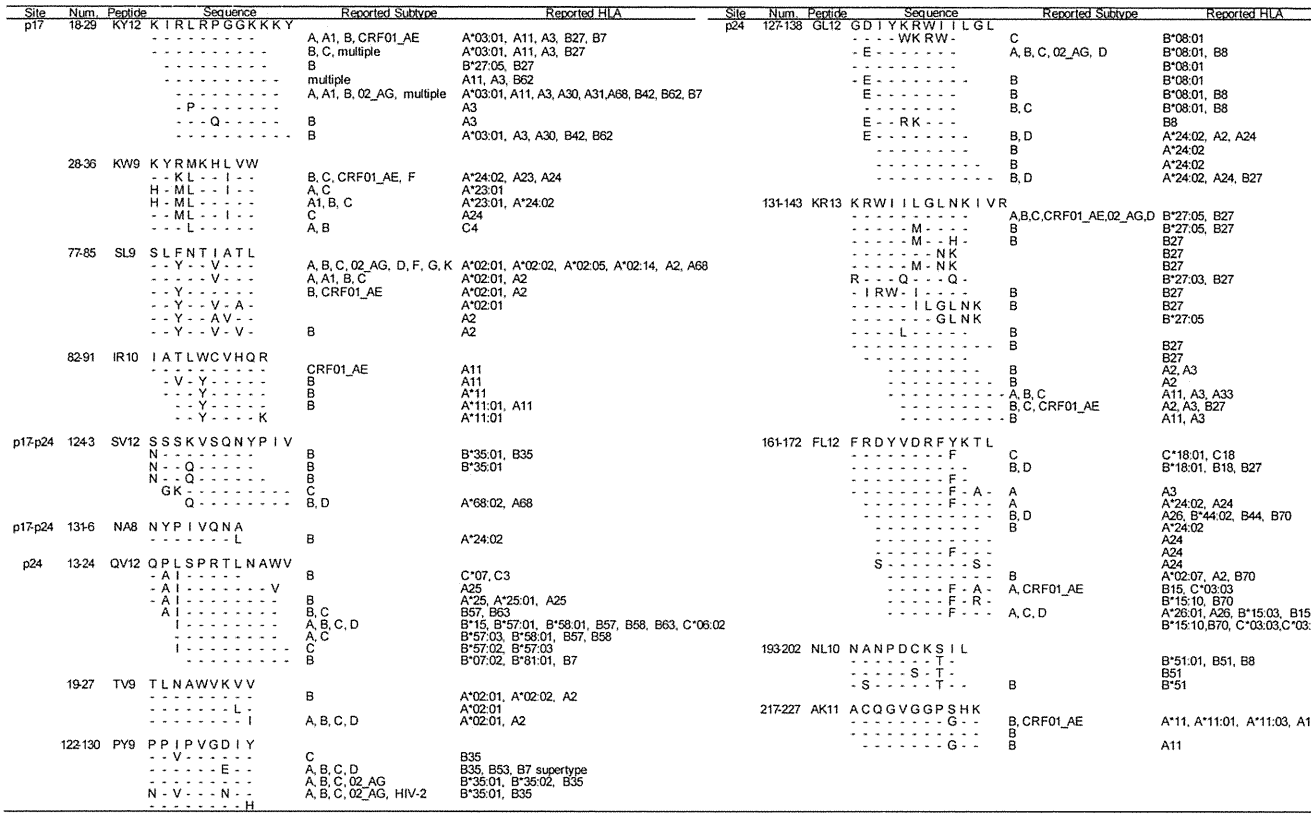
Synthetic peptides containing previously reported CTL epitopes

A set of 14 HIV-1 CRF01_AE Gag peptides were designed based on CTL epitope regions published in the Los Alamos database at the time this study was planned in the year 2000. Of these, 7 encoded previously reported CRF01_AE epitopes. For the remaining epitopes unreported in CRF01_AE infections, the peptide sequences were altered to fit the dominant CRF01_AE sequence as published in the Los Alamos database. The most predominant sequences among available single isolate sequences were selected to design the peptides. We often extended the peptide length up to 12-mer to maximize the frequency of responses if the extension spanned other epitopes restricted by different allele. $p24_{131-143}$ KRWIILGLNKIVR (KR13) was also included despite 13-mer, as it spans both HLA-B27-restricted KRWIILGLNK (KK10) and HLA-A11 and A3-restricted ILGLNKIVR (IR9). Peptides were synthesized by Sigma Genosys (Hokkaido, Japan) with a high purity of >90% as determined by high-pressure liquid chromatography.

We further summarised the optimal epitope sequences of the tested 14 peptides and their variants reported in the Los Alamos database for consequent cross-clade reactivity or clade-specificity analysis (Figure 1). In total, 97 variants of the 14 epitopes have been previously reported from various clades; all of the 14 peptides included reported epitopes from clade B, and 7 of these included epitope reports from CRF01_AE [$p17_{18-29}$: KIRLRPGGKKKY (KY12), $p17_{28-36}$ KYRMKHLVW (KW9), $p17_{77-85}$ SLFNTIATL (SL9), $p17_{82-91}$ IATLWCVHQR (IR10), $p24_{131-143}$ KRWIILGLNKIVR (KR13), $p24_{161-172}$ FRDYVDRFYKTL (FL12) and $p24_{217-227}$ ACQGVGGPSHK (AK11)]. The greatest diversity was found in KR13 with 18 epitope variants, followed by 14 in FL12 and $p24_{127-138}$ GDIYKRWIILGL (GL12). In contrast, $p17_{131}$- $p24_6$: NYPIVQNA (NA8) had only one epitope variant, reported from clade B, while $p24_{19-27}$ TLNAWVKVV (TV9) had 3 epitope variants. Almost all of the restricting HLA alleles were derived from HLA-A or HLA-B alleles, and only 6 epitope variants in 3 peptides included HLA-C alleles as their restricting HLA allele. Out of 66 tested patients, 63 (95.5%) patients had at least one of previously reported HLA alleles which were responsible for CRF01_AE epitope recognition: 63 (95.5%) patients with relevant A alleles, 21 (31.8%) with B alleles, and 7 (10.6%) with C alleles.

Peptide-based IFN-γ ELISpot assay

In ELISpot assay, 14 peptides were tested against all of 66 patients. Peripheral blood mononuclear cells (PBMCs) were isolated by density gradient separation using a Vacutainer CPT Cell Preparation Tube (BD, Franklin Lakes, NJ, USA) and washed twice with RPMI-1640 medium (Sigma-Aldrich, St. Louis, MO, USA). 1×10^5 fresh PBMCs/well were then plated onto multi-Screen plates (MAHA54510; Millipore, Billerica, MA, USA) that had been coated overnight at 4°C with 50 µl of anti-IFN-γ capture Ab 1-D1-K (2 µg/ml; Mabtech, Nacka Strand, Sweden). 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. PBMCs were stimulated with either media alone in negative control wells, 10 µg/ml phytohaemagglutinin (PHA; Sigma-Aldrich) in positive control wells or peptides (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 IFN-γ 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 and incubated for 1 hr at room temperature. Spots were visualized by adding BCIP/NBT substrate (Millipore) and counted by an independent scientist in a blinded fashion using an automated Enzyme-Linked Immunospot (ELISpot) Plate Reader System with the KS 4.3 software. Positive spot forming units (SFU) were counted and results were expressed as SFU/1×10^6 PBMCs. A response was considered positive if it was four times higher than the negative control and greater than 55 SFU/1×10^6 PBMC [17].



Site	Num.	Peptide	Sequence	Reported Subtype	Reported HLA
p17	18-29	KY12	K I R L R P G G K K K Y		
			- - - - - - - - - -	A, A1, B, CRF01_AE	A*03:01, A11, A3, B27, B7
			- - - - - - - - - -	B, C, multiple	A*03:01, A11, A3, B27
			- - - - - - - - - -	B	B*27:05, B27
			- - - - - - - - - -	multiple	A11, A3, B62
			- - - - - - - - - -	A, A1, B, 02_AG, multiple	A*03:01, A11, A3, A30, A31,A68, B42, B62, B7
			- P - - - - - - -		A3
			- - - Q - - - - -	B	A3
			- - - - - - - - - -	B	A*03:01, A3, A30, B42, B62
	28-36	KW9	K Y R M K H L V W		
			- - K L - - I - -	B, C, CRF01_AE, F	A*24:02, A23, A24
			H - M L - - I - -	A, C	A*23:01
			H - M L - - - - -	A1, B, C	A*23:01, A*24:02
			- - M L - - I - -	C	A24
			- - - L - - - - -	A, B	C4
	77-85	SL9	S L F N T I A T L		
			- - Y - - V - - -	A, B, C, 02_AG, D, F, G, K	A*02:01, A*02:02, A*02:05, A*02:14, A2, A68
			- - - - V - - -	A, A1, B, C	A*02:01, A2
			- - Y - - - - -	B, CRF01_AE	A*02:01, A2
			- - Y - - V - A -		A*02:01
			- - Y - - A V - -		A2
			- - Y - - V - V -	B	A2
	82-91	IR10	I A T L W C V H Q R		
			- - - - - - - - -	CRF01_AE	A11
			- V - Y - - - - -	B	A11
			- - - Y - - - - -	B	A*11
			- - Y - - - - - -	B	A*11:01, A11
			- - Y - - - - K		A*11:01
p17-p24	124-3	SV12	S S S K V S Q N Y P I V		
			N - - - - - - - -	B	B*35:01, B35
			N - - Q - - - - -	B	B*35:01
			N - - Q - - - - -	B	
			G K - - - - - - - -	C	
			Q - - - - - - - -	B, D	A*68:02, A68
p17-p24	131-6	NA8	N Y P I V Q N A		
			- - - - - - - L	B	A*24:02
p24	13-24	QV12	Q P L S P R T L N A W V		
			- A I - - - - - -	B	C*07, C3
			- A I - - - - - - - V		A25
			- A I - - - - - - -	B	A*25, A*25:01, A25
			A I - - - - - - - -	B, C	B57, B63
			I - - - - - - - -	A, B, C, D	B*15, B*57:01, B*58:01, B57, B58, B63, C*06:02
			- - - - - - - - -	A, C	B*57:03, B*58:01, B57, B58
			I - - - - - - - - -	C	B*57:02, B*57:03
			- - - - - - - - -	B	B*07:02, B*81:01, B7
	19-27	TV9	T L N A W V K V V		
			- - - - - - - - -	B	A*02:01, A*02:02, A2
			- - - - - - - L -		A*02:01
			- - - - - - - - I	A, B, C, D	A*02:01, A2
	122-130	PY9	P P I P V G D I Y		
			- - V - - - - - -	C	B35
			- - - - - - E - -	A, B, C, D	B35, B53, B7 supertype
			- - - - - - - - -	A, B, C, 02_AG	B*35:01, B*35:02, B35
			N - V - - - N - -	A, B, C, 02_AG, HIV-2	B*35:01, B35
			- - - - - - - - H		
p24	127-138	GL12	G D I Y K R W I I L G L		
			- - - - W K R W -	C	B*08:01
			- E - - - - - - -	A, B, C, 02_AG, D	B*08:01, B8
			- - - - - - - - -		B*08:01
			- E - - - - - - - -	B	B*08:01
			E - - - - - - -	B	B*08:01, B8
			- - - - - - - -	B, C	B*08:01, B8
			E - - R K - - -		B8
			E - - - - - - - -	B, D	A*24:02, A2, A24
			- - - - - - - -	B	A*24:02
			- - - - - - - - -	B	A*24:02
			- - - - - - - - -	B, D	A*24:02, A24, B27
	131-143	KR13	K R W I I L G L N K I V R		
			- - - - - - - - - -	A,B,C,CRF01_AE,02_AG,D	B*27:05, B27
			- - - - - M - - - -	B	B*27:05, B27
			- - - - - M - - - H -	B	B27
			- - - - - - - N K		B27
			- - - - - M - N K		B27
			R - - - Q - - - Q -		B*27:03, B27
			- - I R W - I - - - -	B	B27
			- - - - - - I L G L N K	B	B27
			- - - - - - - G L N K		B*27:05
			- - - - L - - - - -	B	
			- - - - - - - - - -	B	B27
			- - - - - - - - -		B27
			- - - - - - - - -	B	A2, A3
			- - - - - - - - -	B	A2
			- - - - - - - - - -	A, B, C	A11, A3, A33
			- - - - - - - - -	B, C, CRF01_AE	A2, A3, B27
			- - - - - - - - - -	B	A11, A3
	161-172	FL12	F R D Y V D R F Y K T L		
			- - - - - - - - F	C	C*18:01, C18
			- - - - - - - - - -	B, D	B*18:01, B18, B27
			- - - - - - - - F -		
			- - - - - - - - F - A -	A	A3
			- - - - - - - F - - -	A	A*24:02, A24
			- - - - - - - - - -	B, D	A26, B*44:02, B44, B70
			- - - - - - - - -	B	A*24:02
			- - - - - - - - -		A24
			- - - - - - F - - -		A24
			S - - - - - - - S -		A24
			- - - - - - - - -	B	A*02:07, A2, B70
			- - - - - F - A -	A, CRF01_AE	B15, C*03:03
			- - - - - F - R -		B*15:10, B70
			- - - - - F - - -	A, C, D	A*26:01, A26, B*15:03, B15
					B*15:10,B70, C*03:03,C*03:04
	193-202	NL10	N A N P D C K S I L		
			- - - - - - - T -		B*51:01, B51, B8
			- - - - - S - T -		B51
			- S - - - - - T - -	B	B*51
	217-227	AK11	A C Q G V G G P S H K		
			- - - - - - - - G - -	B, CRF01_AE	A*11, A*11:01, A*11:03, A11
			- - - - - - - - - -	B	
			- - - - - - - G - -	B	A11

Figure 1. Sequence variants of 14 peptides and their HLA restrictions. Epitope variations of the 14 peptides used in our ELISpot assays reported in the Los Alamos database and their restricting HLA alleles were listed.
doi:10.1371/journal.pone.0041696.g001

In silico epitope prediction

For ELISpot responses induced by patients with HLA alleles previously unknown to restrict the tested CRF01_AE epitopes, we performed *in silico* prediction of HLA restriction elements and epitopes within peptides using the online prediction tool HLArestrictor [18]. For each HLA-epitope combination we defined the binding affinity using the thresholds Strong Binder (SB) with binding affinity $\leq$50 nM or or % Rank $\leq$0.5, Weak Binder (WB) with binding affinity $\leq$500 nM or % Rank $\leq$2.0, and Combined Binder (CB) with binding affinity $\leq$500 nM, and % Rank $>$2.0, according to the original definitions by the program designers [18].

^{51}Cr release assay

CTL lysis assays were performed using standard ^{51}Chromium release assays as previously described [19]. In brief, peptide-specific CTL lines were generated from freshly isolated PBMCs of HIV-1 infected donors. One seventh of the PBMCs were stimulated with PHA (2 μg/ml) and incubated for 24 hrs, then pulsed with corresponding peptides at 100 μM and incubated for a further hour before being added to the remaining PBMCs. 3×10^5 cells per well were cultured in R10 supplemented with recombinant human IL-7 (25 ng/ml) in 96 well U-bottom plates. The CTL lines were maintained by adding fresh R10 media every 3–4 days. Cells were fed on day 7 with recombinant IL-2 at a concentration of 25 U/ml. Class I HLA-matched or unmatched B cell lines pulsed with peptides were used as target cells, aliquoted at 5×10^4 cells per well in 96-well U bottom plates. Effector cells were added to target cells at an effector/target (E:T) ratio of 20:1 unless otherwise specified. The amount of ^{51}Cr released into the culture supernatants was quantified after 6 hrs of incubation, and the percent specific lysis was determined by using the following formula: $[(E\text{-}S)/(M\text{-}S)]\times100$, where E is the experimental ^{51}Cr release, S for spontaneous ^{51}Cr release in the presence of culture medium and M for maximum release of incorporated ^{51}Cr from target cells treated with 4% Triton X-100. The result was regarded as positive when recognition of the HIV target was $>$10% above the control.

Results

Subjects' background

Of 66 individuals recruited, 55 were female and 11 were male. At the time of enrolment, the median age was 29 years old (range 15–50), the CD4+ T cell count was 473 cells/μl (range 149–1,191) and the viral load was 4.24 log copies/ml (range 2.60–5.97). Table S1 shows HLA distribution.

Frequency of Gag peptide responses and its HLA association

In this study we examined the CTL responses to 14 published Gag sequences to determine the extent of cross-clade reactivity in 66 Thai HIV-1 CRF01_AE infected individuals. Among 66 individuals, 29 individuals (43.9%) showed an IFN-γ response to at least one peptide in the ELISpot assay. In total, 79 responses were identified across all 14 peptides (Table S2). The most frequently recognized peptides were TV9 and FL12 recognised by 10 individuals, followed by $p24_{122-130}$ PPIPVGDIY (PY9) and KY12 recognised by 9 and 8 patients, respectively. In contrast, $p17_{124-}$

$p24_3$: SSSKVSQNYPIV (SV12) had the least number of responders, recognized by one individual, followed by NA8 and KR13 recognised by two and three patients, respectively.

Amongst the 79 responses detected in the ELISpot assay, 42 (53.2%) were responses against the 7 peptides that contained previously reported CRF01_AE epitopes. 52 of the 79 ELISpot responses (65.8%) were induced in patients carrying at least one HLA allele previously reported to restrict the tested epitopes. Of these, 28 out of 52 responses (53.8%) were induced in patients with HLA alleles known to restrict the tested CRF01_AE epitopes, and the remaining 24 responses (46.2%) were in patients carrying HLA alleles reported to restrict the epitopes in other subtypes. Furthermore, 18 out of these 24 responses (75%) were against peptides containing epitopes previously unreported in CRF01_AE infections, indicating that these are cross-clade reactive epitopes, while 6 out of these 24 responses (25%) were against peptides containing known CRF01_AE epitopes.

The remaining 27 out of 79 (34.2%) responses were in individuals carrying HLA alleles previously not known to restrict the tested epitopes, suggesting that at least one third of the peptide responses were restricted by unknown HLA alleles (Table S2). 19 out of these 27 responses were against the peptides containing epitopes previously unknown to be CRF01_AE epitopes, and the remaining 8 were against peptides containing reported CRF01_AE epitopes.

Prediction of epitopes and their HLA restriction using an in silico model

For the 27 responses induced in patients carrying HLA alleles previously not known to restrict the tested epitopes, we performed a prediction of the epitope and its restricting HLA allele using the latest peptide-binding motif based *in silico* program, HLArestrictor (Figure 2). In total, 19 dominant epitope candidates and their associated HLA alleles were detected amongst 6 peptides. Within these 6 peptides, the 2 peptides KW9 and SL9, which contained previously reported CRF01_AE epitopes, 7 dominant epitope candidates were detected, while we identified 12 epitope candidates and their associated HLA alleles in the remaining 4 peptides containing epitopes unreported in CRF01_AE (SV12, QV12, GL12, and NL10). 11 out of 19 (57.9%) epitope candidates were associated with HLA-C alleles, while 7 were associated with HLA-B alleles, and 1 was associated with an HLA-A allele. According to the binding affinity thresholds set by HLArestrictor, out of the 19 epitope-HLA complex candidates identified in our study, 1 was detected as SB, 13 as WB, and 5 as CB. Surprisingly, more than half (11/19) were associated with HLA-C alleles including SB, suggesting the possibility that there may be many HLA-C-associated epitopes that remain unreported.

Identification of a novel epitope-HLA association in CRF01_AE infection with a ^{51}Cr release assay

Using a ^{51}Cr release assay, we further succeeded in demonstrating a novel HLA association of the CRF01_AE Gag epitope $p24_{122-130}$ PPIPVGDIY (PY9), which had been previously reported to be restricted by HLA-B35 and B53 in clade A, B, C, D, CRF02_AG and HIV-2, but not in CRF01_AE. Significant lysis against the PY9 CRF01_AE variant was only detected in the presence of HLA-B*40:01-matched target cells, confirming HLA-B*40:01 as the restricting HLA allele (Figure 3). Out of 9 patients who made an ELISpot response to the peptide PY9, one patient carried the previously reported B7 supertype HLA-B*55:02, and two responders carried the newly detected HLA-B*40:01 allele (Table S2).

Discussion

HIV-1 vaccines in clinical trials today are based on sequences derived from clades B, C, or A, but the identification of conserved HIV-1 CTL epitopes and an understanding of cross-clade CTL responses will be essential to broaden the vaccine responses to include other subtypes of HIV. In the present study, we performed the IFN-γ secreting CTL responses in a Thai cohort with CRF01_AE infection using a set of 14 well-established epitopes and designing CRF01_AE peptide analogues to these epitopes. Our Thai patients infected with CRF01_AE responded to all CRF01_AE analogues, which were not previously reported as CRF01_AE epitopes. Interestingly, however, these analogues were recognized by patients carrying HLA alleles that differed from those reported to restrict the published epitopes, exemplified by our confirmation of a novel HLA restriction of the p24 epitope PY9 by HLA-B*40:01 with a ^{51}Cr release assay. Our findings indicate that the CTL recognition sites in HIV-1 Gag may be shared among different clades but these can be restricted by different HLA molecules, depending on the HLA polymorphism within the cohort.

The process of CTL activation is a highly sensitive and specific process, and a single mutation in the epitope can result in the lack of recognition by the CTLs, the impairment of peptide processing [19,20], or the inhibition of the formation of peptide-HLA complexes [21,22]. Therefore, the inter-clade and intra-clade sequence diversity of HIV-1 has been considered to be the primary barrier to the development of a globally effective vaccine. However, in this study, we have identified the cross-clade epitope candidates which had previously not been reported in CRF01_AE. 37 responses were found across 7 non-CRF01_AE epitopes, suggesting that these are novel cross-clade epitope candidates. It is noteworthy that although CRF01_AE is a recombinant HIV-1 with its Gag sequence derived from clade A (13.4% of Gag sequence discrepancy between clade A and CRF01_AE) [2], the three peptides SV12, NA8 and NL10 have not been reported in clade A.

Over the years, there has been much effort to identify HIV-1 epitopes that mediate potent cross-clade T cell responses [23,24]. However, previous methods utilised peptide pools [23–26] or CTL clones with predetermined HLA alleles to observe cross-clade reactivity [27,28]. The results of our study has shown at the level of single peptide, the lack of peptide recognition seen in the previous studies may have been due to the difference in HLA restriction allele among different population, rather than lack of CTL recognition.

Compared to clade B or C, there is far less epitope information available for clade A. This is the dominant clade circulating in eastern Europe, central Asia and eastern to central Africa, and given that HLA frequencies differ greatly between each region, we anticipate that a detailed epitope mapping study would further reveal the effect of HLA polymorphisms on a particular epitope's immunodominance and its association with viral control, as has been observed among other clades and ethnic groups [29,30].

There are many reports of cross-clade reactivity in HIV-specific CTL responses [31–33]. The mechanism of epitope cross-clade reactivity is poorly understood, however it has been proposed that the more conserved the epitope, the more likely it is to instigate cross-clade reactivity [7], and the sequence variability at anchor positions of the HLA binding motif is thought to be the determining factor [2,32–35]. However, our previous studies of ELISpot assays using overlapping peptides have shown that some peptides containing previously reported epitopes did not induce T cell responses in patients carrying HLA alleles known to restrict

Site	Peptide	HXB2	Sequence	HLA	Binding motif	Affinity (nM)	% Rank	BL
p17	KW9	28-36	K Y R M K H L V W					
	YW8	29-36	- - - - - - - -	C*06:02	xxxxxxxx[LIVY]	197	0.25	SB
	KW9	28-36	- - - - - - - - -	C*06:02	xxxxxxxx[LIVY]	764	1.5	WB
	KV8	28-35	- - - - - - - -	C*06:02	xxxxxxxx[LIVY]	789	1.5	WB
	YW8	29-36	- - - - - - - -	C*07:01	x[RHK]xxxxxx[Y]	NA	1.5	WB
	YW8	29-36	- - - - - - - -	C*07:02	xxxxxxxx[YFL]	NA	1.5	WB
	SL9	77-85	S L F N T I A T L					
	SL9	77-85	- - - - - - - - -	C*03:02	x[A]xxxxxx[FWY]	145	4.0	CB
	SL9	77-85	- - - - - - - - -	C*03:04	x[A]xxxxxx[LM]	399	5.0	CB
p17-p24	SV12	124-3	S S S K V S Q N Y P I V					
	KI8	127-2	- - - - - - - - -	C*07:04	x[RQ]xxxxxx[LM]	NA	2.0	WB
p24	QV12	13-24	Q P L S P R T L N A W V					
	QA10	13-22	- - - - - - - - - -	B*35:60	(B*35:01) x[P]xxxxxx[YFMLI]	NA	1.5	WB
	SV9	16-24	- - - - - - - - - - -	B*56:21	(B*56:01) x[P]xxxxxx[A(L)]	1565	1.5	WB
	SW8	16-23	- - - - - - - - -	B*56:21	(B*56:01) x[P]xxxxxx[A(L)]	1950	1.5	WB
	QL8	13-20	- - - - - - - -	B*56:21	(B*56:01) x[P]xxxxxx[A(L)]	2188	1.5	WB
	QW11	13-23	- - - - - - - - - - -	B*56:21	(B*56:01) x[P]xxxxxx[A(L)]	2243	2.0	WB
	QA10	13-22	- - - - - - - - - -	B*56:21	(B*56:01) x[P]xxxxxx[A(L)]	2298	2.0	WB
	LV10	15-24	- - - - - - - - - -	C*12:03	x[A]xxxxxx[FWY]	214	5.0	CB
	LW9	15-23	- - - - - - - - -	C*12:03	x[A]xxxxxx[FWY]	282	6.0	CB
	LV10	15-24	- - - - - - - - - -	C*14:02	x[YP]xxxxxx[FWY]	488	6.0	CB
	GL12	127-138	G D I Y K R W I I L G L					
	DL11	128-138	- - - - - - - - - - -	A*26:01	x[VTIFL]xxxxxx[YF]	647	1.5	WB
	NL10	193-202	N A N P D C K S I L					
		195-202	- - - - - - - -	B*07:02	x[P(V)]xxxxxx[L]	302	1.5	WB

Figure 2. *In silico* epitope prediction for ELISpot responders carrying HLA alleles currently unknown to restrict CRF01_AE epitopes. We found 27 CRF01_AE specific CTL responses induced in patients carrying HLA alleles previously unknown to restrict CRF01_AE epitopes. Prediction of the optimal epitope within the peptide and its restricting HLA allele was performed using the *in silico* epitope prediction model HLArestrictor. In total, 19 epitope-HLA combinations were detected with binder levels defined as SB (Strong Binder), WB (Weak Binder), or CB (Combined Binder). NA: Not available, and BL: Binder level.
doi:10.1371/journal.pone.0041696.g002

these epitopes even if their epitope sequences at anchor positions were compatible with the binding motif of the restricting HLA alleles [36]. The lack of a T cell response despite the binding motif matching with the epitope sequences may be accounted for by the amino acid sequence of the flanking regions, especially when the epitopes are shorter than the peptide tested, [19,20] or sequence variation in positions other than the anchor regions, both of which may influence the recognition of the MHC-peptide complex by T cell receptors (TCRs) [21,22].

Recently, *in silico* algorithms have been used to biometrically recombine and design vaccine epitopes that elicit CTL responses of higher breadth (number of peptide recognition) and depth (response of variants within an epitope) for experimental studies in rhesus macaques [37,38]. Taken together with our data suggesting the potential of multiple epitopes with cross-clade reactivity, these studies support the possibility of the development of a cross-clade reactive vaccine.

In our study we were also able to use epitope prediction models to identify epitopes within peptides that induced responses in patients carrying HLA alleles previously unknown to restrict the tested CRF01_AE epitopes, highlighting the potential of such *in silico* models to identify epitopes restricted by rare HLA alleles like HLA-C. CTL-epitope information for HLA-C alleles has been sparse compared to A and B alleles, as can be observed in the Los Alamos database. The lower level of cell surface expression of HLA-C alleles compared to the other alleles [39–41] and the lack of protective HLA-C allele information in population studies [10–12] has hampered epitope mapping against HLA-C alleles. However, a recent single nucleotide polymorphism (SNP) study in a Caucasian cohort identified a region 35 kb upstream of the gene encoding the HLA-C molecule (−35(C/T)) as the second strongest determining factor for viral control [42], capable of inducing either a higher (C) or lower (T) expression level of HLA-C on the cell surface [43]. As seen from the recent identification of dominant HLA-C-restricted epitopes [44], we can expect an increasing number of studies extending our understanding of how the HLA-C alleles contribute to viral control. Epitope information from HLA-C alleles will also contribute to our knowledge of a given epitope's cross-clade reactivity or clade-specificity. It is warranted to study further especially to demonstrate the potentially novel HLA alleles restriction listed in Figure 2.

As a post hoc analysis, this study also has several limitations. Firstly, due to limited PBMCs available from the patients we focused on Gag peptides and did not investigate responses to whole viral proteins. However, Gag, especially the p24 protein, is one of the most important target antigens for viral control [36,45,46], due to their role in the selection of escape mutations that lead to viral fitness costs [47], its sequence stability compared

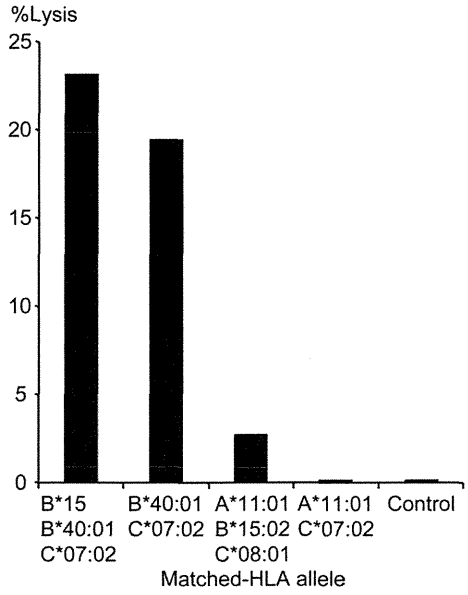


Figure 3. Cytotoxicity assay with T cells Demonstration of a novel epitope-HLA association by ^{51}Cr release assay. Specific lysis of Gag peptide $p24_{122-130}$ PPIPVGDIY (PY9) pulsed allogeneic target cells by effector CTLs from a HLA-B*40:01+ donor was assessed in a chromium release assay. The *Y* axis shows percentage specific lysis at an E:T ratio of 20:1 with the lysis (%) of unpulsed target cells subtracted. Effector cells were derived from patient 1509. HLA-B*40:01 matched cells pulsed with PY9 were also recognized by patient 326 (data not shown). HLA alleles shared by target cells and effector cells are shown; control indicates HLA-unmatched cells.
doi:10.1371/journal.pone.0041696.g003

to other viral particles [1,2,7], the abundance of the protein on incoming virions [48], as well as its rapid antigen presentation following viral infection [49]. For the development of a globally effective CTL-induced vaccine, detailed mapping of Gag epitopes and their restricting HLA alleles will be essential. Secondly, in this study we analysed 66 subjects, but further studies with larger population sample sizes may help identify cross-clade CTL epitopes restricted by minor HLA alleles and allow us to differentiate linkage disequilibrium effects from true associations. Thirdly, since the CTL epitope information was limited when these 14 peptides were designed in the year 2000, this peptide selection may not be optimal in the current setting. However, even after 2000, surprisingly only three CRF01_AE-associated epitopes with responsible four digits HLA allele were reported according to the latest Los Alamos database 2012; $p24_{35-43}$ EVIPMFSAL restricted by HLA-A*26:01 and A*26:03 [50], $p24_{145-153}$ YSPVSILDI by HLA-C*01:02 [51], and $p24_{209-217}$ ATLEEMMTA by HLA-A*02:06 [52]. This is why we think that our data is still worth reporting. Fourthly, we used five 12-mer peptides and one 13-mer peptide in this study to maximize the frequency of responses by spanning more epitopes. However, this extension of the peptides may have lowered the peptide responses but we did not find an obvious tendency that the longer peptides had less response. For optimal epitope and responsible HLA confirmation, further experiment with peptide narrow down will be required.

Fifthly, in computational epitope prediction, although more than half of candidates were detected from HLA-C alleles, prediction of HLA-C allele-associated epitopes was not accurate compared to that of HLA-A or B-associated as commented by programmers [18]. It is warranted to demonstrate the potentially novel HLA alleles restriction listed in Figure 2.

From our findings, we conclude that the HLA restriction of an epitope in a given population is dictated by two factors: HLA polymorphism within the population and viral sequence diversity. Previous studies on CTL cross-recognition have focused on the sequence divergence between the clades, and have promoted the inclusion of highly conserved epitopes in CTL-epitope vaccines [7,32–35]. However, we have shown that in different cohort populations, CTL recognition of the same epitopes may occur through unique HLA restrictions. We believe that the HLA restriction of epitopes should be determined for a given population prior to the selection of vaccine candidate immunogens, as certain epitopes will be able to induce broad, cross-clade responses which will increase the potential efficacy of the vaccine in the given population. We hope that the novel epitopes and HLA restrictions identified in this study will contribute to the development of a cross-clade reactive HIV vaccine.

Supporting Information

Table S1 HLA distribution among 66 HIV-1 CRF01_AE-infected Thais. HLA distribution by population frequency is shown.
(XLSX)

Table S2 HLA allele information of ELISpot assay responders and its compatibility with previous report. In total, 79 responses among 14 epitopes were identified. HLA allele information of ELISpot assay responders and its compatibility with previous report of responsible HLA alleles listed in Los Alamos database are shown.
(XLSX)

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 the Lampang hospital.

Author Contributions

Conceived and designed the experiments: BS PS KA. Performed the experiments: BS MT SN. Analyzed the data: MM TM KA. Contributed reagents/materials/analysis tools: BS MT SN PP KA. Wrote the paper: MM SN TM KA.

References

1. Goulder PJ, Watkins DI (2008) Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat Rev Immunol 8: 619–30.
2. 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.
3. Stephens HA (2005) HIV-1 diversity versus HLA class I polymorphism. Trends Immunol 26: 41–7.
4. 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 10: 11–23.
5. Amara RR, Villinger F, Altman JD, Lydy SL, O'Neil SP, et al. (2001) Control of a mucosal challenge and prevention of AIDS by a multiprotein DNA/MVA vaccine. Science 292: 69–74.

6. Barouch DH, Santra S, Schmitz JE, Kuroda MJ, Fu TM, et al. (2000) Control of viremia and prevention of clinical AIDS in rhesus monkeys by cytokine-augmented DNA vaccination. Science 290: 486–92.
7. 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.
8. Mungall AJ, Palmer SA, Sims SK, Edwards CA, Ashurst JL, et al. (2003) The DNA sequence and analysis of human chromosome 6. Nature 425: 805–11.
9. Robinson J, Mistry K, McWilliam H, Lopez R, Parham P, et al. (2011) The IMGT/HLA database. Nucleic Acids Res 39: D1171–6.
10. 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.
11. Leslie A, Matthews PC, Listgarten J, Carlson JM, Kadie C, et al. (2010) Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. J Virol 84: 9879–88.
12. O'Brien SJ, Gao X, Carrington M (2001) HLA and AIDS: a cautionary tale. Trends Mol Med 7: 379–81.
13. 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.
14. Ou CY, Takebe Y, Weniger BG, Luo CC, Kalish ML, et al. (1993) Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. Lancet 341: 1171–1174.
15. Saengwonloey O, Jiraphongsa C, Foy H (2003) Thailand report: HIV/AIDS surveillance 1998. J Acquir Immune Defic Syndr 32 Suppl 1: S63–67.
16. Lau KA, Wang B, Saksena NK. (2007) Emerging trends of HIV epidemiology in Asia. AIDS Rev 9: 218–29.
17. Streeck H, Frahm N, Walker BD (2009) The role of IFN-gamma Elispot assay in HIV vaccine research. Nat Protoc 4: 461–9.
18. Erup Larsen M, Kloverpris H, Stryhn A, Koofhethile CK, Sims S, et al. (2011) HLArestrictor–a tool for patient-specific predictions of HLA restriction elements and optimal epitopes within peptides. Immunogenetics 63: 43–55.
19. 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.
20. Tenzer S, Wee E, Burgevin A, Stewart-Jones G, Friis L, et al. (2009) Antigen processing influences HIV-specific cytotoxic T lymphocyte immunodominance. Nat Immunol 10: 636–46.
21. 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.
22. 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.
23. Buseyne F, Chaix ML, Fleury B, Manigart O, Burgard M, et al. (1998) Cross-clade-specific cytotoxic T lymphocytes in HIV-1-infected children. Virology 250: 316–24.
24. Cao H, Mani I, Vincent R, Mugerwa R, Mugyenyi P, et al. (2000) Cellular immunity to human immunodeficiency virus type 1 (HIV-1) clades: relevance to HIV-1 vaccine trials in Uganda. J Infect Dis 182: 1350–6.
25. 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 U S A 94: 1396–401.
26. Bansal A, Gough E, Ritter D, Wilson C, Mulenga J, et al. (2006) Group M-based HIV-1 Gag peptides are frequently targeted by T cells in chronically infected US and Zambian patients. AIDS 20: 353–60.
27. Cao H, Kanki P, Sankalé JL, Dieng-Sarr A, Mazzara GP, et al. (1997) Cytotoxic T-lymphocyte cross-reactivity among different human immunodeficiency virus type 1 clades: implications for vaccine development. J Virol 71: 8615–23.
28. Fukada K, Tomiyama H, Wasi C, Matsuda T, Kusagawa S, et al. (2002) Cytotoxic T-cell recognition of HIV-1 cross-clade and clade-specific epitopes in HIV-1-infected Thai and Japanese patients. AIDS 16: 701–11.
29. Masemola AM, Mashishi TN, Khoury G, Bredell H, Paximadis M, et al. (2004) Novel and promiscuous CTL epitopes in conserved regions of Gag targeted by individuals with early subtype C HIV type 1 infection from southern Africa. J Immunol 173: 4607–17.
30. Frahm N, Kiepiela P, Adams S, Linde CH, Hewitt HS, et al. (2006) Control of human immunodeficiency virus replication by cytotoxic T lymphocytes targeting subdominant epitopes. Nat Immunol 7: 173–8.
31. Coplan PM, Gupta SB, Dubey SA, Pitisuttithum P, Nikas A, et al. (2005) Cross-reactivity of anti-HIV-1 T cell immune responses among the major HIV-1 clades in HIV-1-positive individuals from 4 continents. J Infect Dis 191:1427–34.
32. 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.
33. 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.
34. 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.
35. 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.
36. 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: e22680.
37. Barouch DH, O'Brien KL, Simmons NL, King SL, Abbink P, et al. (2010) Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. Nat Med 16: 319–23.
38. Santra S, Liao HX, Zhang R, Muldoon M, Watson S, et al. (2010) Mosaic vaccines elicit CD8+ T lymphocyte responses that confer enhanced immune coverage of diverse HIV strains in monkeys. Nat Med 16: 324–8.
39. McCutcheon JA, Gumperz J, Smith KD, Lutz CT, Parham P (1995) Low HLA-C expression at cell surfaces correlates with increased turnover of heavy chain mRNA. J Exp Med 181: 2085–95.
40. Neisig A, Melief CJ, Neefjes J (1998) Reduced cell surface expression of HLA-C molecules correlates with restricted peptide binding and stable TAP interaction. J Immunol 160: 171–9.
41. Snary D, Barnstable CJ, Bodmer WF, Crumpton MJ (1977) Molecular structure of human histocompatibility antigens: the HLA-C series. Eur J Immunol 7: 580–5.
42. Fellay J, Frahm N, Shianna KV, Cirulli ET, Casimiro DR, et al. (2007) A whole-genome association study of major determinants for host control of HIV-1. Science 317: 944–7.
43. Thomas R, Apps R, Qi Y, Gao X, Male V, et al. (2009) HLA_C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA_C. Nat Genet 41: 1290–4.
44. Makadzange AT, Gillespie G, Dong T, Kiama P, Bwayo J, et al. (2010) Characterization of an HLA-C-restricted CTL response in chronic HIV infection. Eur J Immunol 40: 1036–41.
45. 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.
46. 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.
47. 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.
48. 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.
49. 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.
50. Kawashima Y, Satoh M, Oka S, Takiguchi M. (2005) Identification and characterization of HIV-1 epitopes presented by HLA-A*2603: comparison between HIV-1 epitopes presented by A*2601 and A*2603. Hum Immunol 66: 1155–66.
51. Buranapraditkun S, Hempel U, Pitakpolrat P, Allgaier RL, Thantivorasit P, et al. (2011) A novel immunodominant CD8+ T cell response restricted by a common HLA-C allele targets a conserved region of Gag HIV-1 clade CRF01_AE infected Thais. PLoS One 6: e23603.
52. Kantakamalakul W, de Souza M, Bejrachandra S, Ampol S, Cox J, et al. (2006) Identification of a novel HIV type 1 CRF01_AE cytotoxic T lymphocyte (CTL) epitope restricted by an HLA-Cw0602 allele and a novel HLA-A0206/peptide restriction. AIDS Res Hum Retroviruses 22: 1271–82.

Epidemiol. Infect., Page 1 of 9. © Cambridge University Press 2012
doi:10.1017/S0950268812002397

Chronic hepatitis B and C co-infection increased all-cause mortality in HAART-naive HIV patients in northern Thailand

N. TSUCHIYA[1,2], P. PATHIPVANICH[3], A. ROJANAWIWAT[4], N. WICHUKCHINDA[4], I. KOGA[5], M. KOGA[6], W. AUWANIT[4], P. E. KILGORE[7], K. ARIYOSHI[1,2]* AND P. SAWANPANYALERT[4]

[1] *Department of Clinical Medicine, Institute of Tropical Medicine (NEKKEN), Nagasaki University, Japan*
[2] *Global COE program, Nagasaki University, Japan*
[3] *Day Care Center, Lampang Hospital, Thailand*
[4] *National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Thailand*
[5] *Department of Internal Medicine, Teikyo University, Japan*
[6] *The Institute of Medical Science, The University of Tokyo, Japan*
[7] *Wayne State University, Detroit, MI, USA*

Received 22 May 2012; Final revision 5 October 2012; Accepted 8 October 2012

SUMMARY

A total of 755 highly active antiretroviral therapy (HAART)-naive HIV-infected patients were enrolled at a government hospital in Thailand from 1 June 2000 to 15 October 2002. Census date of survival was on 31 October 2004 or the date of HAART initiation. Of 700 (92·6%) patients with complete data, the prevalence of hepatitis B virus (HBV) surface antigen and anti-hepatitis C virus (HCV) antibody positivity was 11·9% and 3·3%, respectively. Eight (9·6%) HBV co-infected patients did not have anti-HBV core antibody (anti-HBcAb). During 1166·7 person-years of observation (pyo), 258 (36·9%) patients died [22·1/100 pyo, 95% confidence interval (CI) 16·7–27·8]. HBV and probably HCV co-infection was associated with a higher mortality with adjusted hazard ratios (aHRs) of 1·81 (95% CI 1·30–2·53) and 1·90 (95% CI 0·98–3·69), respectively. Interestingly, HBV co-infection without anti-HBc Ab was strongly associated with death (aHR 6·34, 95% CI 3·99–10·3). The influence of hepatitis co-infection on the natural history of HAART-naive HIV patients requires greater attention.

Key words: Co-infection, hepatitis B, hepatitis C, mortality, resource-limited settings.

INTRODUCTION

In resource-limited countries, the prevalence of chronic hepatitis B virus (HBV) and hepatitis C virus (HCV) infection is often high [1], and populations with a high prevalence of HBV and HCV usually overlap with those seriously affected by HIV. In a study from northern India, the reported prevalence of HBV and HCV co-infection in HIV-infected patients was 5·3% and 2·4%, respectively [2]. A study in Tanzania reported 17·3% and 18·1% of HIV-infected patients were co-infected with HBV and HCV, respectively [3], and an earlier report from Thailand showed that the prevalence of HBV infection was 8·7% and HCV infection 7·8% in HIV-infected patients [4].

* Author for correspondence: Dr K. Ariyoshi, Department of Clinical Medicine, Institute of Tropical Medicine (NEKKEN), Nagasaki University, Nagasaki, Japan.
(Email: kari@nagasaki-u.ac.jp)

Accumulating evidence suggests that HIV co-infection adversely affects the clinical course of hepatitis. Increased HBV carriage rates, greater levels of HBV viraemia, more rapid decline in HBV surface antibody, increased reactivation episodes, and faster progression to liver cirrhosis are all characteristic of HIV/HBV co-infected patients [5, 6]. In HIV/HCV co-infected patients, faster progression of fibrosis resulting in decompensated cirrhosis have been shown in previous studies [7, 8]. As patients on highly active antiretroviral therapy (HAART) survive much longer, liver failure is becoming the major cause of death in patients with hepatitis co-infection [9].

Current International AIDS Society guidelines for the management of HIV recommend that HIV/viral hepatitis co-infected patients should start HAART, the same as HIV mono-infected patients. Moreover, initiation of HAART is recommended regardless of CD4 cell count when treatment for HBV is considered [10]. However, there is a big gap between the recommendation of the guidelines and the real-life situation in resource-limited countries. While access to HAART has markedly increased, even in resource-limited countries [11], overall HAART coverage remains as low as 36% (95% CI 33–39%) [11] based on 2010 WHO guidelines (treatment initiation at CD4 cell count <350 cells/μl. Moreover, monitoring (viral load testing and genotyping) and treatment for hepatitis are not available due to the cost in most resource-limited countries. Thus, we assume that there are still large groups of patients with HIV/chronic hepatitis co-infection who are not receiving HAART in resource-limited countries. For better management of these patients, it is important to know the association between hepatitis co-infection and the natural history of HIV infection.

The effect of viral hepatitis co-infection on HIV progression and all-cause mortality before initiating HAART remains uncertain. Most studies conducted in the late 1990s and early 2000s did not include HIV viral load in their analysis. Some studies presented a more rapid progression to AIDS and reduced survival in patients who have chronic HBV or HCV infection [12–15] while others have shown conflicting results [16–19]. In the majority of previous studies examining HBV and HCV co-infection, the main transmission mode of HIV in participants was homosexual intercourse or injecting drug use (IDU) and none were conducted with a substantial sample size in Asian or African countries where the majority of HIV-infected individuals with HBV vertical transmission reside. The present study aims to evaluate the impact of hepatitis co-infection on all-cause mortality in HAART-naive HIV-infected individuals in northern Thailand.

METHODS

Study site and population

To address the current research question, we re-analysed our previously conducted natural history cohort of HIV-infected patients in northern Thailand [20, 21]. This patient cohort was assembled from volunteers at the HIV centre of a government referral hospital with about 800 beds situated in the centre of Lampang province in upper northern Thailand. The centre was established in October 1995 as an out-patient clinic providing treatment, care and support for HIV-infected patients. The recruitment of this cohort was from 1 July 2000 to 15 October 2002 before the national antiretroviral treatment programme was launched. All adult (aged >18 years) HIV-infected individuals attending the HIV clinic who were HAART-naive at the first visit were approached by the research team and enrolled if written consent were obtained. All participants were requested to visit the clinic at least once every 3 months regardless of the presence of clinical symptoms and were followed up from the date of study enrolment until 15 October, 2004. This study was approved by the Thai Government Ethics Committee in December 1999 and December 2005.

Data collection

Demographic (gender and age at enrolment) data and medical history [HIV-related symptoms, history of antiretroviral therapy (ART) and mode of transmission] of patients were obtained at the study enrolment by well-trained research staff through face-to-face interviews based upon a structured questionnaire. In addition to physical examination by two research physicians, complete blood count (CBC), platelet count, CD4 cell count and HIV viral load (copies/ml) were measured. CD4 cell count was determined by flow cytometric technique FACScan (BD Biosciences, USA) and HIV viral load was measured using a commercial kit (Amplicor HIV-1 Monitor Test, Roche Molecular Systems Inc., USA). To address the present research question, the remaining freeze-stored plasma samples from our previous study were retrospectively tested for hepatitis B surface

antigen (HBsAg), anti-hepatitis B core antibody (anti-HBcAb) and antibody to hepatitis C virus (anti-HCV) were retrospectively tested using commercially kits: Cobas Core HBsAg II EIA, Anti-HBc EIA, and ETI-AB-HCVK-4 (DiaSorin S.p.A., Italy). HBsAg positivity and anti-HCV positivity were determined to define HBV and HCV co-infection, respectively. Cohort patient survival was assessed on 15 October 2004. Survival status for each patient was ascertained by hospital records, death certificates, mailing letters, and contacting families or relatives. Causes of death in HIV/hepatitis co-infected patients were investigated by reviewing hospital records.

Analysis

In survival analysis, patients who started HAART before 15 October 2004 were regarded as censored on the date of starting HAART. Kaplan–Meier survival analysis was performed to estimate survival in relation to the existence of HIV/HBV or HCV co-infection. HBV co-infected patients were divided into two subgroups according to the existence of anti-HBcAb. We used the log-rank test to compare Kaplan–Meier curves. Additionally, Cox's proportional hazard model was conducted to evaluate the influence of HBV or HCV co-infection on survival adjusted by several factors. In multivariate models, other than hepatitis co-infection status and existence of anti-HBcAb, we included all variables with $P<0{\cdot}1$ in univariate analysis. Results were presented as hazard ratios (HRs) with 95% confidence intervals (CIs). The proportional hazard assumption was explored using Nelson–Aalen plots and the likelihood ratio test. Statistical analyses were conducted using SPSS version 17.0 (SPSS Inc., USA) and Stata version 11.0 (StataCorp., USA).

Ethical approval

This study was conducted as part of the Lampang HIV Cohort Phase I and Lampang & Phayao HIV Cohort Phase II, which were approved by Research Ethics Committee of the Thai Ministry of Public Health.

RESULTS

Characteristics of participants

Seven hundred and fifty-five patients (over 95% of patients who attended the clinic during the targeted period) were enrolled in the study. Of 755 HIV-infected persons, 55 patients were excluded (32 patients had received HAART in private clinic or other clinical trial before recruitment, 22 patients had incomplete testing for hepatitis co-infection, and four patients had HBV/HCV dual co-infection). Baseline characteristics of patients are summarized in Table 1. Male:female ratio was 0·75. Median age at enrolment was 33 years (95% CI 29–37), and 21·8% had received mono or dual ART. The major transmission route was heterosexual (96·7%) and half of participants were asymptomatic. The CD4 cell count was >200 cells/μl in 299 (41·5%) patients and <50 cells/μl in 261 (36·3%) patients. The prevalence of HBV and HCV co-infection was 11·9% and 3·3%, respectively. Of 700 patients, 681 (97·3%) had both HBsAg and anti-HBcAb status. HCV co-infection was strongly associated with IDU ($P<0{\cdot}001$). All IDU cases were co-infected with HCV. The baseline CD4 cell count was significantly lower in HBV *vs.* HCV co-infected patients whereas there was no difference between these patient groups in their baseline viral load. Patients with HBV co-infection were more likely to have AIDS symptoms compared to HIV mono-infected patients ($P=0{\cdot}02$). Platelet counts were lowest in HCV co-infected individuals followed by HBV co-infected patients ($P=0{\cdot}03$). Table 2 shows the status of HBsAg and anti-HBcAb. Of 86 HBV co-infected patients, 78 (90·7%) had anti-HBcAb. Interestingly, eight (9·3%) patients did not have anti-HBcAb despite being HBsAg positive.

Impact of chronic hepatitis infection on all-cause mortality

Complete follow-up data were available for 694 (99·1%) patients in the evaluated cohort. Total follow-up time was 1166·7 person-years of observation (pyo) with a median patient follow-up of 588 days [interquartile range (IQR), 317–967]. During the observation period, 258 (36·9%) patients died, resulting in a mortality rate of 22·1/100 pyo (95% CI 16·7–27·8). When stratified according to baseline CD4 cell count, mortality was 64·6/100 pyo (95% CI 59·1–70·9) in patients with CD4 <50 cells/μl compared to 22·9/100 pyo (95% CI 16·3–29·7) in the group with CD4 <50–200 cells/μl and 5·17/100 pyo (95% CI 2·57–7·63) in patients with CD4 >200 cells/μl. When stratified with baseline clinical status, mortality was 8·71/100 pyo (95% CI 5·76–11·6) in asymptomatic patients, 27·5/100 pyo (95% CI 19·6–34·4) in

Table 1. *Baseline characteristics of HIV mono-infected patients and hepatitis co-infected patients*

Characteristics ($N=700$)	HIV mono-infection ($n=594$)	HBV co-infection ($n=83$)	HCV co-infection ($n=23$)
Age (median, IQR)*	33 (29–37)	32 (29–35)	30 (26–34)
Male gender (%)	244 (41·1)	42 (50·6)	14 (50·9)
Mode of transmission (%)*†‡			
Heterosexual	583 (98·1)	79 (95·2)	15 (65·2)
Homosexual	6 (1·0)	1 (1·2)	1 (4·3)
IDU	0 (0·0)	0 (0·0)	5 (21·7)
Others or multiple	5 (0·8)	2 (2·4)	2 (8·6)
Unknown	0 (0·0)	1 (1·2)	0 (0·0)
Previous ART (%)§	129 (21·7)	17 (20·5)	4 (17·4)
Clinical status (%)†‖			
Asymptomatic	312 (52·5)	32 (38·6)	8 (34·8)
Non-AIDS symptomatic	115 (19·4)	15 (18·1)	7 (30·4)
AIDS symptomatic	167 (28·1)	36 (43·4)	8 (34·8)
Baseline CD4 cell count (cells/μl) (median, IQR)*†‡	150 (22–351)	64 (19–219)	334 (31–459)
Baseline vial load (copies/ml) (median, IQR)	157431 (34891–446396)	169773 (69806–437116)	112334 (21415–515515)
Baseline platelet count (median, IQR)†	261000 (215000–319000)	237000 (191000–300000)	213000 (170000–328000)

IQR, Interquartile range; IDU, injecting drug user; ART, antiretroviral therapy; AIDS, acquired immunodeficiency syndrome.

* $P<0·05$, HIV mono-infection *vs.* HCV co-infection.

† $P<0·05$, HIV mono-infection *vs.* HBV co-infection.

‡ $P<0·05$, HBV co-infection *vs.* HCV co-infection.

§ Experience with antiretroviral therapy is limited to monotherapy or dual therapy.

‖ Definition of AIDS according to the National guidelines for the clinical management of HIV infection in children and adults, 6th edn. Thailand: Ministry of Public Health, 2000.

Other than *, †, ‡ there were not any significant differences between the groups.

Table 2. *Status of HBsAg and anti-HBcAb*

	anti-HBcAb(+)	anti-HBcAb(–)	Total
HBsAg(+)	78	8	86
HBsAg(–)	318	277	595
Total	396	285	681

HBsAg, Hepatitis B surface antigen; anti-HBcAb, anti-hepatitis B core antibody.

Both HBsAg and anti-HBcAb status data were available for 681 patients.

symptomatic but non-AIDS patients and 60·6/100 pyo (95% CI 54·4–66·7) in symptomatic AIDS patients. Kaplan–Meier survival analysis (Fig. 1) revealed that HBV co-infection significantly increased mortality ($P<0·001$, log-rank test). HCV co-infection also tended to increase mortality, but the statistical significance was marginal. The curves for HIV/HBV and HCV/HIV co-infection converge at about 500 days. The likelihood ratio test for interaction by time band with a cut-point at 500 days revealed a P value of 0·2, confirming lack of evidence for a relevant violation of the proportional hazard assumption, allowing the use of Cox regression analysis. The influence of hepatitis co-infection on survival analysed by the Cox proportional hazard model is presented in Table 3*a*. In univariate analysis, factors associated with death were male gender, previous ART treatment, clinical symptom at enrolment, baseline CD4 cell count, baseline viral load, HBV co-infection and existence of anti-HBcAb. Patients with a low platelet count ($<150\,000$) were more likely to die than those with a higher platelet count. In multivariate analysis, there was no significant association between platelet count and death. HCV co-infection showed a tendency towards association with decreased survival.

HBV serology and mortality

Survival estimates focused on HBV serology interestingly showed that HBV co-infected individuals without anti-HBcAb had the poorest survival compared

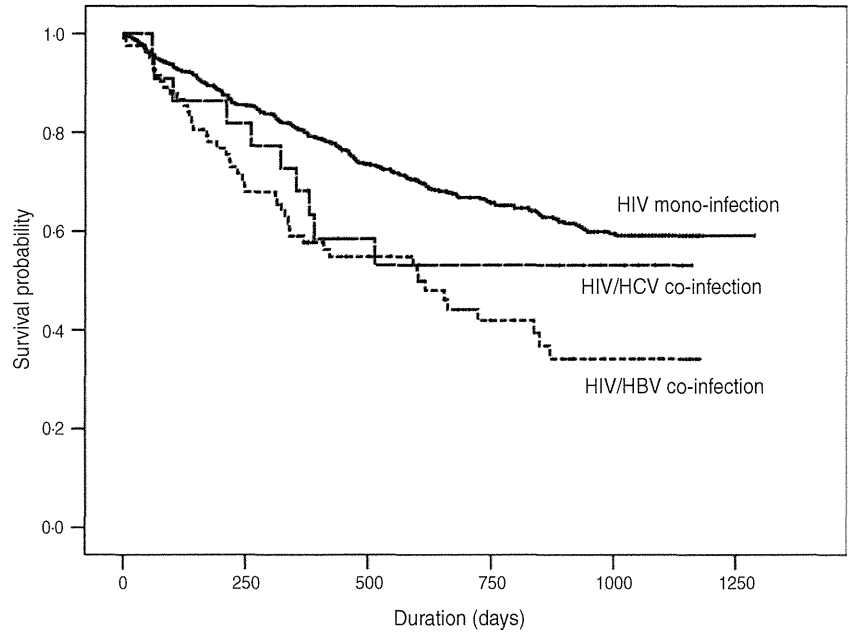


Fig. 1. Kaplan–Meier survival probability estimate of co-infected and mono-infected individuals showing that HBV co-infection significantly increased mortality.

to HIV mono-infected or HBV co-infected patients with anti-HBcAb ($P < 0{\cdot}0001$ by log-rank test, Fig. 2). In multivariate analysis (Table 3*b*), in addition to symptomatic AIDS and low CD4 cell count, HBV co-infection without anti-HbcAb was a strong risk factor for death (adjusted HR 6·34, 95% CI 3·99–10·3).

Cause of death in hepatitis co-infected patients

Although primary causes of death were unknown for nine out of 56 HIV/hepatitis co-infected patients, the majority of deaths (46/56, 82·1%) were attributed to AIDS-related diseases. No patients were diagnosed with liver failure before death except for one patient who was hospitalized for 6 days due to newly diagnosed cirrhosis with portal hypertension. This patient was lost to follow-up after discharge with a CD4 cell count of 22 cells/μl and died 3 months later.

DISCUSSION

In the present study with a substantial sample size, we have clearly demonstrated that HBV co-infection significantly increased all-cause mortality of HAART-naive HIV patients. It was only after the advent of HAART substantially increased life expectancy of HIV-infected individuals that the importance of liver-related death due to hepatitis co-infection was highlighted. Our data indicate that the influence of hepatitis co-infection on the natural history of HAART-naive HIV patients should not be ignored. We also discovered an increased mortality of HCV co-infected patients compared to HIV mono-infected patients. The results of multivariate analysis showed only a trend, but this is probably due to the small number of HCV co-infection in this study.

There are some limitations of the present study. First, the lack of data on HBV DNA and HCV RNA viral load is a limitation. Patients spontaneously clear HCV infection after acquisition. The absence of HCV viraemia in anti-HCV-positive patients might explain why HCV co-infection did not show the strong association with death. Second, cause of death was determined only by reviewing medical charts while the information of survival status was ascertained by contacting families and relatives in addition to reviewing hospital records. It is possible that cause of death might have been misclassified in some cases. However, we focused on all-cause mortality and these limitations do not alter the main results of the study.

To our knowledge, this is the first study from a resource-limited country to address the adverse influence of hepatitis co-infection on survival in HAART-naive HIV patients. We also found a higher mortality rate in this Thai population compared to white patients in New York in the 1980s [22]. Our finding

Table 3 (*a*). *Influence of HBV or HCV co-infection on survival*

Variables	HR (95% CI)	*P* value	aHR (95% CI)	*P* value
Male sex	2·59 (2·02–3·33)	<0·001	1·23 (0·94–1·60)	0·14
Age <30 years*	0·79 (0·61–1·01)	0·06	0·81 (0·63–1·05)	0·11
Previous ART†	0·67 (0·49–0·92)	0·01	1·16 (0·84–1·61)	0·37
Transmission mode				
IDU	Ref.	—	—	—
Homosexual	0·36 (0·05–2·58)	0·31	—	—
Heterosexual	0·68 (0·17–2·72)	0·58	—	—
Others	1·12 (0·21–6·13)	0·89	—	—
Clinical symptom*‡				
Asymptomatic	Ref.	—	Ref.	—
Symptomatic, non-AIDS	3·06 (2·16–4·33)	<0·001	1·42 (0·96–2·08)	0·08
AIDS, symptomatic	6·65 (4·93–8·98)	<0·001	2·05 (1·44–2·93)	<0·001
Baseline CD4 cell count† (cells/μl)				
⩾200	Ref.	—	Ref.	—
50–199	4·46 (2·91–6·84)	<0·001	2·98 (1·85–4·79)	<0·001
<50	12·7 (8·72–18·6)	<0·001	6·44 (4·04–10·3)	<0·001
Baseline viral load (copies/ml)				
<10000	Ref.	—	Ref.	—
10000–49999	2·44 (1·04–5·75)	0·04	1·20 (0·48–3·01)	0·70
50000–99999	3·30 (1·39–7·85)	0·007	1·67 (0·70–4·00)	0·25
⩾100000	8·79 (4·13–18·7)	<0·001	2·19 (0·96–5·00)	0·06
Baseline platelet count				
<150000	1·52 (1·00–2·29)	0·05	1·16 (0·76–1·78)	0·50
HBV co-infection	2·05 (1·49–2·82)	<0·001	1·81 (1·30–2·53)	<0·001
HCV co-infection	1·26 (0·67–2·38)	0·47	1·90 (0·98–3·69)	0·06

HR, Hazard ratio; CI, confidence interval; aHR, adjusted hazard ratio; ART, antiretroviral therapy; IDU, injecting drug user; AIDS, acquired immunodeficiency syndrome.

* At the time of enrolment.

† Experience with antiretroviral therapy is limited to monotherapy or dual therapy.

‡ Definition of AIDS according to the National guidelines for the clinical management of HIV infection in children and adults, 6th edn. Thailand: Ministry of Public Health, 2000.

implies that the high prevalence of hepatitis co-infection is at least partially responsible for the high mortality observed in HIV-infected individuals in resource-limited countries. Several studies have investigated the influence of hepatitis co-infection on the natural course of HIV infection but these studies often include patients receiving HAART and show conflicting findings [12–14, 16, 17, 23]. One of the reasons for this discrepancy appears to be the sample sizes in previous studies [16, 23]. In addition, none of the previous studies has analysed the association between HIV and HBV or HCV co-infection after adjusting for clinical status, viral load, and CD4 cell count in a cohort not receiving HAART with substantial sample size.

Our investigation of hospital records did not identify any patients with liver failure, with one exception. We might have undiagnosed hepatocellular carcinoma in some patients. However, we believe that the majority of chronic HBV or HCV co-infected HIV patients died from AIDS-defined illness rather than liver failure, since they all had significant opportunistic infections with a very low CD4 cell count. These results suggest that hepatitis co-infection accelerates the natural course of HIV infection itself. An *in vitro* study has demonstrated that HBV-X protein super-induces ongoing HIV replication and HIV-1 long-terminal repeat transcription [24]. However, according to our results, HBV co-infection increased the mortality independent of HIV viral load. Thus, the increased HIV viral load *per se* does not fully explain the higher mortality in HBV co-infected patients.

We also considered the possibility that the higher mortality in HBV co-infected patients might have been due to confounding factors. However, our multivariate analysis demonstrated that the association