Vaccine 36 (2018) 4142–4151

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Computational comparison of availability in CTL/gag epitopes among patients with acute and chronic HIV-1 infection

Gabriel Dario Damilano [a,*], Omar Sued [b], Maria Julia Ruiz [a], Yanina Ghiglione [a], Flavia Canitano [c], Maria Pando [a], Gabriela Turk [a], Pedro Cahn [b], Horacio Salomón [a], Dario Dilernia [a,d]

[a] Instituto de Investigaciones Biomédicas en Retrovirus y SIDA Facultad de Medicina, Universidad de Buenos Aires, Argentina
[b] Fundación Huésped, Buenos Aires, Argentina
[c] Instituto de Investigaciones Médicas Alfredo Lanari, Buenos Aires, Argentina
[d] Emory University, Atlanta, USA

ARTICLE INFO

Article history:
Received 25 January 2018
Received in revised form 23 April 2018
Accepted 29 April 2018
Available online 22 May 2018

ABSTRACT

Background: Recent studies indicate that there is selection bias for transmission of viral polymorphisms associated with higher viral fitness. Furthermore, after transmission and before a specific immune response is mounted in the recipient, the virus undergoes a number of reversions which allow an increase in their replicative capacity. These aspects, and others, affect the viral population characteristic of early acute infection.

Methods: 160 single gag-gene amplifications were obtained by limiting-dilution RT-PCR from plasma samples of 8 ARV-naïve patients with early acute infection (<30 days, 22 days average) and 8 ARV-naive patients with approximately a year of infection (10 amplicons per patient). Sanger sequencing and NGS SMRT technology (Pacific Biosciences) were implemented to sequence the amplicons. Phylogenetic analysis was performed by using MEGA 6.06. HLA-I (A and B) typing was performed by SSOP-PCR method. The chromatograms were analyzed with Sequencher 4.10. Epitopes and immune-proteosomal cleavages prediction was performed with CBS prediction server for the 30 HLA-A and -B alleles most prevalent in our population with peptide lengths from 8 to 14 mer. Cytotoxic response prediction was performed by using IEDB Analysis Resource.

Results: After implementing epitope prediction analysis, we identified a total number of 325 possible viral epitopes present in two or more acute or chronic patients. 60.3% (n = 196) of them were present only in acute infection (prevalent acute epitopes) while 39.7% (n = 129) were present only in chronic infection (prevalent chronic epitopes). Within p24, the difference was equally dramatic with 59.4% (79/133) being acute epitopes ($p < 0.05$). This is consistent with progressive viral adaptation to immune response in time and further supported by the fact that cytotoxic responses prediction showed that acute epitopes are more likely to generate immune response than chronic epitopes. Interestingly, only 27.5% of acute epitopes match the population-level consensus sequence of the virus.

Conclusions: Our results indicate that certain non-consensus viral residues might be transmitted more frequently than consensus-residues when located in immunological relevant positions (epitopes). This observation might be relevant to the rationale behind development of an effective vaccine to reduce viral reservoir and induce functional cure of HIV infection based in prevalent acute epitopes.

1. Introduction

Following transmission, Cytotoxic CD8+T lymphocytes (CTLs) mount a powerful response to transmitted HIV in the acute phase of infection [1]. However, the vast majority of cases is an inefficient response and directed to a limited number of epitopes [2]. This response is manifested producing a viral set point generally after first month of infection [3–5] with great specificity on the gag protein, which is strongly associated with control of viral replication [6–9]. At this point, just after the peak of viremia, the first viral

* Corresponding author.
E-mail addresses: gabrieldario06@hotmail.com (G.D. Damilano), omar.sued@huesped.org.ar (O. Sued), mariajulia83@gmail.com (M.J. Ruiz), yghiglione@fmed.uba.ar (Y. Ghiglione), flacanitano@hotmail.com (F. Canitano), mpando@fmed.uba.ar (M. Pando), gturk@fmed.uba.ar (G. Turk), pcahn@huesped.org.ar (P. Cahn), hsalomon@fmed.uba.ar (H. Salomón), ddilern@emory.edu (D. Dilernia).

https://doi.org/10.1016/j.vaccine.2018.04.086
0264-410X/

populations with HLA/CTL escape mutations are generated [2], which will increase over time and will have an impact on viral diversity during the chronic phase [10–12]. Therefore, it is interesting to investigate if these HLA-dependent reversions that occur during the acute phase are associated with a greater availability of epitopes in this stage.

Reversions are another important aspect to be considered among viral factors shaping the viral diversity during the acute stage of infection [13,14]. While their impact on viral fitness on transmitted/founder virus has not yet been elucidated, it is believes reversal mutations make the virus more fit [13,15,16]. These post-transmission reversions may range from those associated with HLA alleles present on the donor [17], to the CD8 T-cell TCR receptor [13,18,19], to the proteasome [20,21], or reversions that arise at random and increase the viral fitness that end up being established in the major viral population [22,24]. Also, these reversions could occur prior to transmission, in a process of viral compartmentalization in mucosal associated tissues [25–27].

In this research, our objective was to determine, in patients under study, whether the viral sequences approach phylogenetically in acute patients compared with chronic patients. In turn, we evaluated whether possible reversions in acute viral sequences could significantly influence to decrease viral phylogenetic distance in these patients. Also, using computational prediction analysis, we evaluated the relationship of phylogenetically relevant amino acid positions (possibly reversions), with epitopes associated with cytotoxic immune response mediated by HLA I. Finally, we evaluated the distribution and characteristics of CTL epitopes, in conserved viral segments. We did an availability and frequency analysis of possible CTL epitopes found in the early acute phase compared to those found in the chronic phase.

2. Methods

2.1. Study subjects

A total of 16 subjects participated in this study, 8 were enrolled during acute/early primary HIV infection (<30 days, 22 days average), and 8 were chronically infected patients, Table 1. All subjects were enrolled by the Grupo Argentino de Seroconversión Study Group under the following inclusion criteria for acute infection; detection of HIV RNA or p24 antigen with a simultaneous negative or indeterminate screening (ELISA and particle agglutination test) and western blot assay. Clinical stage of primary infection subjects was determined according to the system of Fiebig et al. [28]. Patients with chronic infection were defined as subjects with established HIV infection with an estimated average of 221 days of infection, with detectable viral load (VL > 50 copies of HIV RNA/ml plasma). Both acute and chronic patients were HAART naive at the time of sampling.

2.2. Samples

Blood samples were collected from study participants at enrollment and centrifuged to separate plasma, which was stored at −80 °C until use. Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Hypaque density gradient centrifugation (Amersham, Sweden) and cryopreserved. In the case of chronic, and acute subjects, plasma VL (branched-DNA, Versant HIV-1 RNA 3.0 assay; Siemens Healthcare) and CD4+T-cell count (flow cytometry double platform, BD FACSCanto; BD Biosciences) were determined.

2.3. DNA/Amino acids sequence production

Viral RNA extraction was from plasma and DNA from PBMC using commercial QIAamp viral extraction kit and QIAamp DNA mini kit (QIAGEN, Hilden, Germany), respectively. 160 single Gag-gene amplifications were obtained by limiting-dilution RT-PCR from plasma samples of 8 ARV-naïve patients with early acute infection and 8 ARV-naive patients with chronic infection (10 full gag amplicons per patient, from viral RNA). For thefirst strand DNA synthesis was used1rdgag-Reverse primer, SuperScript™ III Reverse Transcriptase (Thermo Fisher Scientific) whit 10ul of RNA sample in a final volume of reaction of 20 µl. Then, the c DNA was diluted successively until obtaining the limit dilution. Two to five dilutions were required to obtain single-gag sequence amplifications. Amplification was conducted under following conditions: for reverse transcription 5 min at 65 °C (1°rdgagRev/RNA), 50 min at 42 °C, 15 min at 70 °C (1°rdgagRev/RNA/SSIII), one cycle. The conditions of the nested PCR were: Platinum Taq polymerase HiFi (Thermo Fisher Scientific) was used under appropriate reaction conditions; 3 min at 95 °C one cycle; 15 s at 95 °C, 15 s at 56 °C, 1 min 40 s at 72 °C for 5 cycles; 15 s at 90 °C, 15 s at 56 °C, 1 min 40 s at 72 °C for 30 cycles and one cycle at 72 °C for 10

Table 1
Clinical and virological characteristics of patients enrolled for this research. All were Argentine patients. ND: no data; WB: western blot; Screening: ELISA and particle agglutination test. (*) about a year of infection.

Patients	ID	Gender	Phase	Year of enrollment	Estimated time since infection (days)	WB/screening	Viral Load	HLA A/B	Viral subtype (*gag*)	Fiebig stage
1	A700	M	Acute	2014	21	p24+/±	>500000	02AZKD 24/13YFVU 35AMYE	B	IV
2	A327	M	Acute	2015	15	-/-	550000	11 33/44CXJC 51	CRF12 BF	II
3	A901	M	Acute	2015	15	-/-	>500000	ND	B	II
4	A166	M	Acute	2015	15	-/-	>500000	ND	CRF12 BF	II
5	A134	F	Acute	2013	26	p24+/±	>500000	2 24 / 15YCK 07	CRF12 BF	IV
6	A429	M	Acute	2013	30	p24+/±	>500000	24 68 / 40CXJJ(61) 35	B	IV
7	A585	M	Acute	2013	30	p24,gp160+/±	>500000	2 32 / 40CXJM 39	B	IV
8	A959	M	Acute	2014	25	p24+/±	260028	03DFPJ 24HBMP/07GZTE 35AMYE	CRF12 BF	IV
9	C123	M	Chronic	2012	320	+/+	19250	2 68 / 15CXHX(62) 14	B	-
10	C193	M	Chronic	2013	>100*	ND	264322	03 31/15YCK 08	CRF12 BF	-
11	C377	M	Chronic	2012	>100*	ND	30020	ND	CRF12 BF	-
12	C572	M	Chronic	2014	>100*	+/+	383841	2 68 / 7CXPJ 15CXHW	B	-
13	C571	M	Chronic	2013	410	+/+	27194	2 X / 57 40CXJJ(61)	CRF12 BF	-
14	C273	M	Chronic	2015	>100*	+/+	21697	3 1 / 7CXPJ 44CXHZ	B	-
15	C957	F	Chronic	2013	460	+/+	4005	2 3 / 44CXJC 35	CRF12 BF	-
16	C502	M	Chronic	2013	180	+/+	20469	ND	CRF12 BF	-

min. Primers: 1°rd gagFw 5′-CTAGCAGTGGCGCCCGAACAGG-3′, 1°rd gagRev 5′-CAGTCTTTCATTTGGTGTCCTTC-3′, 2°rd gagFw5′-TC TCTCGACGCAGGACTCG-3′, 2°rd gagRev 5′-TTTCCACATTTCCAA CAGCCC-3′.

Sanger sequencing and NGS SMRT technology (Pacific Biosciences) were implemented to sequence the amplicons, method as described previously [29]. The chromatograms obtained from the sequencing of the viral quasispecies were analyzed with Sequencher 4.10 (http://www.genecodes.com/). Gen Bank accession numbers: MF356702-MF356861.

2.4. HLA typing

Genotyping of the genetic polymorphisms of HLA class I A and B was performed using PCR-SSOP (polymerase chain reaction-sequence specific oligonucleotide probe, Fujirebio INNO LiPA) [30,31], and was reported in medium and high resolution. For this technique, an amplification step was performed by the polymerase chain reaction (PCR) from PBMC DNA with biotinylated primers designed to encompass the polymorphic zone. The biotinylated amplified product was chemically denatured (strand separation) and hybridized to nitrocellulose strips containing specific immobilized oligonucleotide probes, complementary to the polymorphisms known in the alleles of our population (between 30 and 40 probes for HLA IA and B). Each strip has two internal control bands (hybridization and conjugation controls). Once the process was finished, the interpretation of results to determine the possible combinations of alleles was done through tables (manual) and Liras® HLA software v6.00 (Fujirebio-Europe Company).

2.5. DNA/Amino acids sequence analysis

Epitope prediction was performed with NetMHC CBS prediction server (http://www.cbs.dtu.dk/services/NetMHC/) for the 19 HLA-A and -B alleles most prevalent in our population with peptide lengths from 8 to 14 mer. NetChop 3.1 CBS prediction server (http://www.cbs.dtu.dk/services/NetChop/) was used to estimate the immune-proteosomal cleavage sites in gag sequences. Cytotoxic response prediction was performed by using the IEDB Analysis Resource (http://tools.iedb.org). For phylogenetic analysis, MEGA 6.06 (http://www.megasoftware.net/) was used, with which the phylogenetic reconstruction was performed for the acute and chronic viral amino acid sequences separately. For the phylogenetic trees, Neighbor-joining was used as a statistical method based on p-distances. On the other hand, an estimation of acute and chronic intragroup phylogenetic distances was performed. Phylogenetic distances were calculated from the nucleotide and amino acid sequences by constructing distance matrix using the Poisson models with uniform rate between sites and the Equal input model with uniformity of velocity and homogeneity of patterns, without considering deletions and insertions in both cases [32–34]. The same result was obtained with both models.

For consensus sequence construction was used Consensus Maker from Los Alamos HIV Database tools.

No bias was found for HLA alleles and viral subtype between the acute and chronic groups.

2.6. Ethics statement

The study was reviewed and approved by two institutional review boards (IRB): Comité de Ética Humana, Facultad de

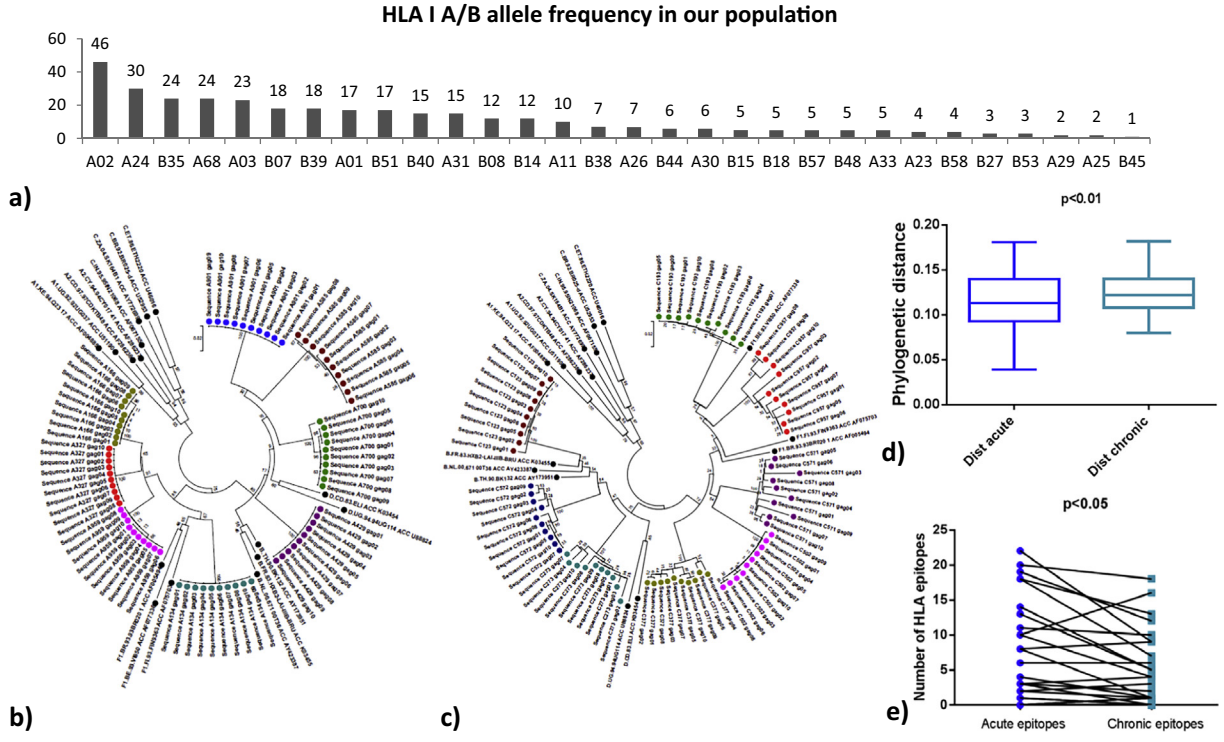

Fig. 1. (a) Prevalence of each HLA I allele in our population that were used as reference in this research. All amino acid sequences obtained were subjected to prediction of epitopes for the 30 described alleles. Phylogenetic trees of amino acid gag sequences from patients in acute phase (b) and from patients in chronic phase (c) using Neighbor joining test, bootstrap method (1000 rep), viral subtype references A, B, C, D and F were included. (d) Result of the comparison of the phylogenetic distances between both groups, (Distance matrix through equal model test, MEGA 6.06 software. Unpaired *t* test, $p < 0.05$). (e) Comparison between the number of possible epitopes found in the acute phase versus the chronic phase for each analyzed HLA I allele, (paired *t* test, $p < 0.05$).

Medicina, Universidad de Buenos Aires, and Comité de Bioética, Fundación Huésped (Buenos Aires, Argentina). Both HIV-infected participants provided written informed consent accepting to participate in this study and are under medical monitoring.

3. Results

In this study, a total of 16 HIV-positive ARV naïve patientswe studied: eight had very early acute infections and eight had chronic infection, Table 1. Viral RNA was extracted from plasma and amplified and sequenced, combining limiting dilution cloning (single genome amplification) and NGS SMRT technology (Pacific Biosciences) [29,35,36] a total of 10 full gag sequences per patient (160 full gag sequences in total). HLA I and B alleles were typed (see methods). The prevalence for each allele in our population were used as a reference in this research [37] Fig. 1a.

3.1. Phylogenetic analysis

Phylogenetic trees of both groups (acute and chronic) were constructed using the Neighbor joining test, which shows greater diversification of gag in virus from chronic patients compared to the virus from acute patients, Fig. 1b and c. First, and because of previous observations that certain polymorphisms were transmitted more frequently than others, we found it interesting to begin this analysis by comparing the phylogenetic distances between the gag amino acid sequences in the group of patients with acute infection versus the group of chronic patients. Phylogenetic distances were calculated from nucleotide and aminoacid sequences by constructing distance matrices using two models in parallel (see Methods). The results of the comparison of the phylogenetic distances between the two groups are shown in Fig. 1d (only aminoacid-Equal model results are shown), where the phylogenetic distances of the gag sequences are significantly lower compared to the distances found between viral sequences in chronic patients. This suggests, at least for the population under study, and in agreement with previous studies [38–41], that viral populations found in early acute stages tend to approach phylogenetically. In addition, we could verify by consensus amino acid sequence construction for each group that the acute phase consensus sequence presents a lower phylogenetic distance to the HXB2 consensus compared to the chronic stage consensus, which coincides with previous studies [38]. Fig. 4b shows a summary of consensus sequences highlighting the polymorphisms associated with HXB2 found.

3.2. Analysis of affinity of epitopes to MHC I A/B

Following the same line of analysis, we performed a computational analysis, for possible CTL epitopes in the amino acid sequences of the gag clones in both groups. To do this, we first evaluated the affinity of the peptides present in the amino acid gag sequences of all the clones of both groups by means of NetMHC CBS prediction server [42,43]. In this analysis we measured the affinities of 30 HLA I A and B alleles present in our population for peptides of 8 to 14 amino acids. Then, we counted only those peptides with moderate to high affinity for the HLA I molecule (500 nM to 1 nM), which are exclusive to the acute or chronic stage and were present in two or more patients (frequency $\geq$25% on the total of acute or chronic patients). A total of 325 possible epitopes that met these requirements were obtained. Of these, 60.3% (196 epitopes) belonged to the acute phase and 39.7% (129 epitopes) to the chronic phase, Fig. 1e. A similar relationship was found within the sequence of protein p24, 59.4% (79 epitopes) in the acute stage compared to 40.6% (54 epitopes) in the chronic phase ($p < 0.01$), Table 2. It is interesting that, despite the lower phylogenetic distance of the acute phase sequences with consensus HXB2, out of the 196 possible epitopes found, only 27.5% (21.5% in p24) coincided with this consensus.

Fig. 2 (a and b) shows the distribution of epitopes in the gag protein accounted for based on the inclusion criteria described above. It was observed that most of the epitopes for the different HLA I analyzed are in defined segments and do not present a random distribution along the gag protein. Fig. 2c shows the distribution of these conserved epitopes on both stages of the infection. Only possible epitopes with a CTL prediction score greater than zero (see below) is shown. Of the entropy analysis as a measure of variability in the analyzed gag sequences, we observed a high association of acute-chronic epitopes with more variable segments (91.9% for acute epitopes and 88.3% for chronic epitopes) in comparison with conserved epitopes in both phases of infection, which showed a lower association with variable segments (57.4%) ($p < 0.01$), Fig. 2 (a, b and c Graphics).

As for the frequencies of each epitope on both phases of infection, we found that the possible acute phase epitopes presented a higher frequency than the possible chronic phase epitopes (unpaired t test, $p < 0.05$), Fig. 3a. Regarding the frequency at the individual level, the difference was more significant, with the finding that 93% of the possible acute phase epitopes were found in the majority viral population, compared to 67% of the possible epitopes in the chronic phase (z test, $p < 0.05$), Fig. 3b and c.

On the other hand, the distribution of epitopes found in both stages of infection, that is, those that were found in both acute and chronic stages with a frequency equal to or greater than 50% in both phases were also analyzed. A total of 334 possible epitopes were counted, of which 72.8% were in p24, and 62.5% of the total

Table 2
Number of possible epitopes in full gag of acute and chronic patients and their prediction of cytotoxic response ordered according to prevalence of HLA I alleles analyzed.

		HLA gag epitopes		Positive CTL score response	
HLA	Frequency	Acute	Chronic	Acute	Chronic
A02	46	18	7	12	4
A24	30	0	1	–	1
B35	24	3	0	3	–
A68	24	6	6	5	1
A03	23	14	5	11	1
B07	18	3	4	1	2
B39	18	1	0	1	–
A01	17	2	3	0	0
B51	17	0	0	–	–
B40	15	4	1	3	0
A31	15	22	9	14	0
B08	12	0	1	–	0
B14	12	0	0	–	–
A11	10	10	4	7	1
B38	7	0	0	–	–
A26	6.8	0	0	–	–
B44	5.8	18	13	7	1
A30	5.8	8	9	5	2
B15	5	20	18	15	13
B18	4.9	2	1	2	0
B57	4.9	10	16	4	10
B48	4.9	0	0	–	–
A33	49	13	5	8	0
A23	3.9	0	1	–	1
B58	3.9	11	10	6	8
B27	2.9	8	1	3	0
B53	2.9	1	0	0	–
A29	1.9	3	2	2	2
A25	1.9	0	0	0	–
B45	1	19	12	17	0
Total		196	129	126	47

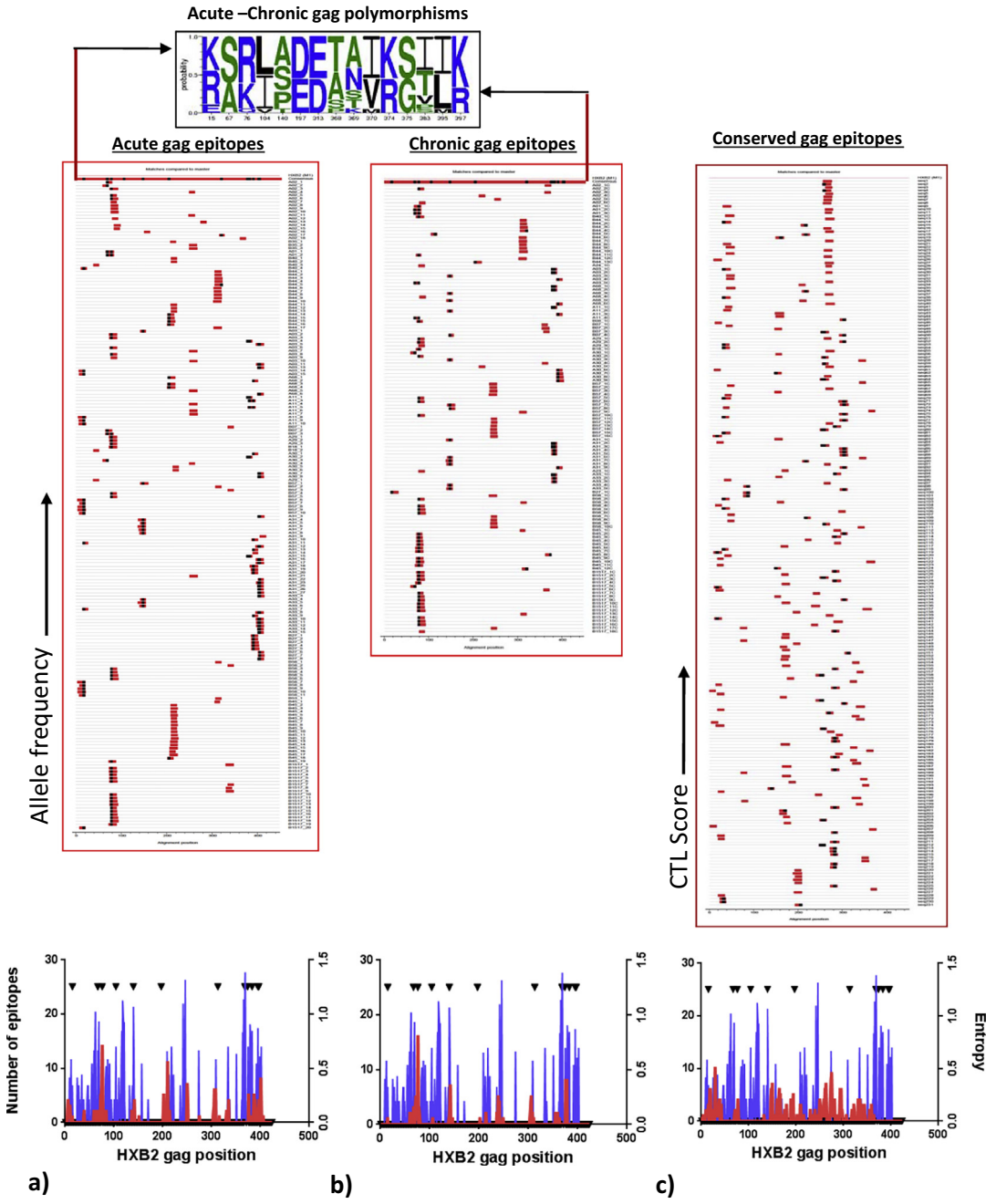

Fig. 2. The alignments and graphs distribution of acute (a), chronic (b) and conserved (c) epitopes in the gag protein sequence accounted for based on the inclusion criteria is shown. In black, (■ in alignments, ▼ in graphics), the distribution (HXB2 sequence position) of 15 acute/chronic polymorphisms. In (c), the most conserved epitopes are represented, where only epitopes with a positive CTL prediction score are shown. In the graphs, the result of the entropy distribution in the analyzed sequences is also observed.

coincided with the consensus HXB2, the opposite of what happens with the possible exclusive epitopes of the acute phase where the frequency of HXB2 epitopes was lower.

Finally we assessed the possible relation between the frequency of alleles in the population and the epitopes found at each stage of infection. There was only a significant difference in the number of epitopes between both stages of infection in alleles with a frequency equal to or greater than 10%. Thus, considering only the peptide-p MHC affinity, most of the reversions found in the acute phase of infection are associated with the higher frequency alleles in the population, Fig. 3d.

3.3. Prediction of cytotoxic CTL response

We performed a prediction of cytotoxic CTL response for all possible epitopes found, both the exclusives of each phase and the conserved. To do this we used the model of the immunological response of the IEBD analysis resource (http://tools.iedb.org) which evaluates each amino acid position in each peptide, especially 4–6 (most associated with interaction with the TCR receptor) and it also evaluates the interaction p MHCI-TCR [42–44].We obtained a score for each peptide analyzed that had a value of -1 to 1, where peptides with scores above zero have a better chance

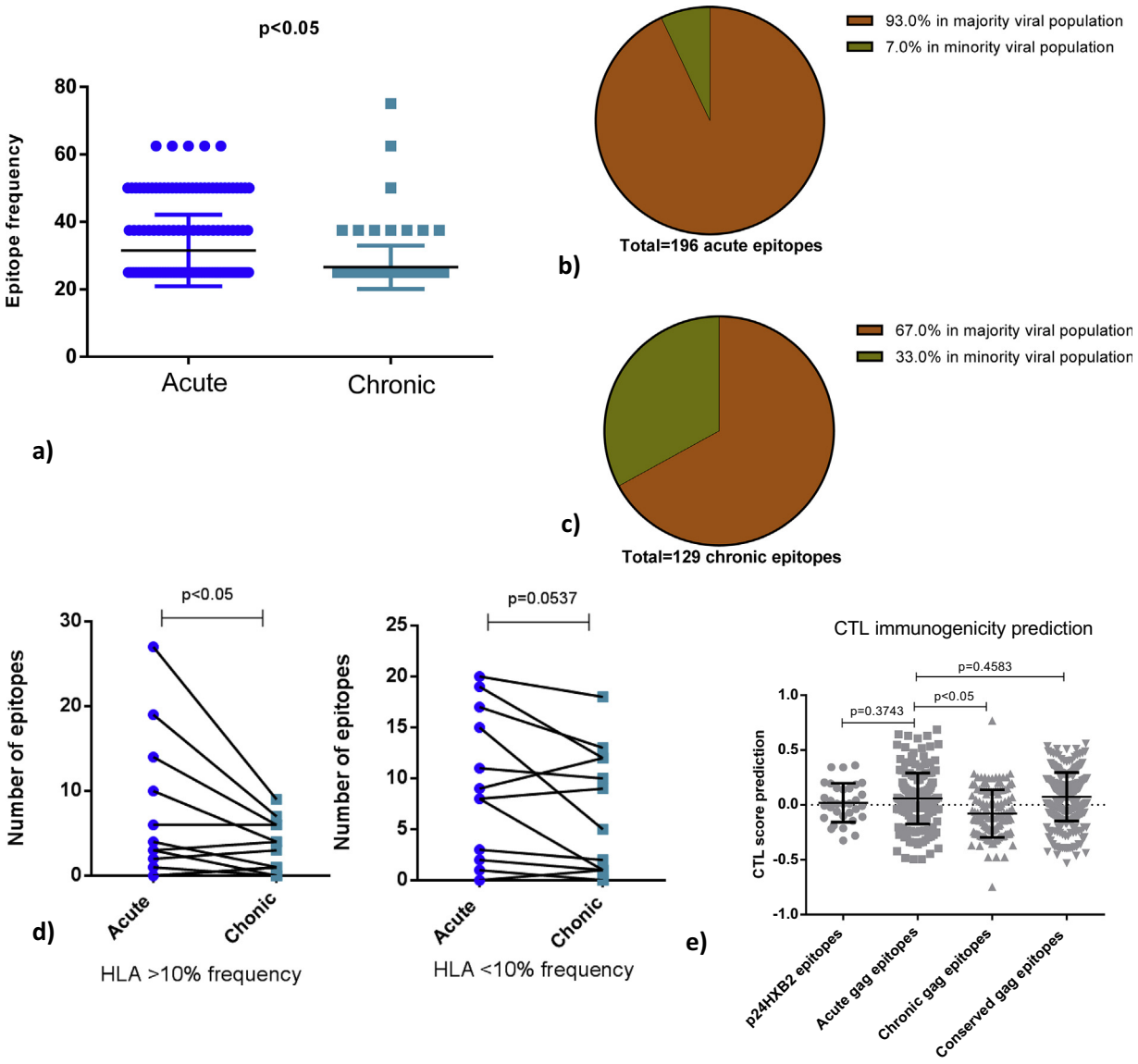

Fig. 3. (a) Prevalence comparison for possible epitopes within each group, acute versus chronic. In (b) and (c) are shown the proportion of intra-patient viral clones (quasispecies) of acute and chronic containing the possible epitopes. (d) Relationship between the frequency of HLA I alleles in our population and the possible epitopes found at each stage of infection based on the affinity for the HLA I molecule. (e) Prediction of cytotoxic CTL response for the possible epitopes of each phase and for conserved epitopes. We used as control p24 gag epitopes consensus with immune response tested by ELISPOT assay (Source: HIV molecular immunology database, www.hiv.lanl.gov).

(about 66%) of generating an immune response, whereas peptides with scores below zero are less likely (about 44%) to generate in vivo/*in vitro* immune responses. The results showed that the possible acute phase epitopes were more likely to generate immune response compared to chronic phase epitopes ($p < 0.01$). We used as control p24 gag epitopes consensus with the immune response tested by ELISPOT assay (Source: HIV molecular immunology database, www.hiv.lanl.gov) [45–48], which showed a slightly positive tendency of score while the possible acute phase epitopes presented a higher score than the chronic epitopes. Finally, we evaluated the possible epitopes preserved in both groups, which showed a higher average score to the rest of the groups ($p < 0.05$), Fig. 3e.

3.4. Acute and chronic sequence analysis

The lower phylogenetic distance of gag sequences previously observed among patients with acute infection suggests the presence of prevalent amino acids in both phases. Therefore, we consider it important to find out if the prevalent epitopes found in each phase of infection were associated with these amino acid positions or are randomly distributed throughout the gag protein. From the amino acid sequences obtained in each phase of infection, we obtained acute and chronic consensus sequences. Then the alignment and comparison of these consensus sequences was made and we were able to determine which amino acids are conserved and which have a significantly unequal prevalence in each phase. We observed 15 amino acid positions with significant variability (frequency difference $\geq$25%, $p < 0.1$) that distinguish both groups (polymorphisms), Fig. 4a and b.

Thus, we found that the amino acids of 12 of the 15 identified polymorphisms are included in 73% of the possible acute phase epitopes (polymorphisms; 1 Lysine, 2 Alanine, 3 Arginine, 5 Alanine, 6 Glutamic acid, 10 Valine, 11 Arginine, 12 Glycine, 13 Isoleucine, 14 Leucine and 15 Arginine). 7.1% of these epitopes included polymorphism 5 and 10 but not amino acids (Alanine and Valine, respectively) because they represent an escape mutation to HLAI. On the other hand, we found that 9 of the 15 amino acids

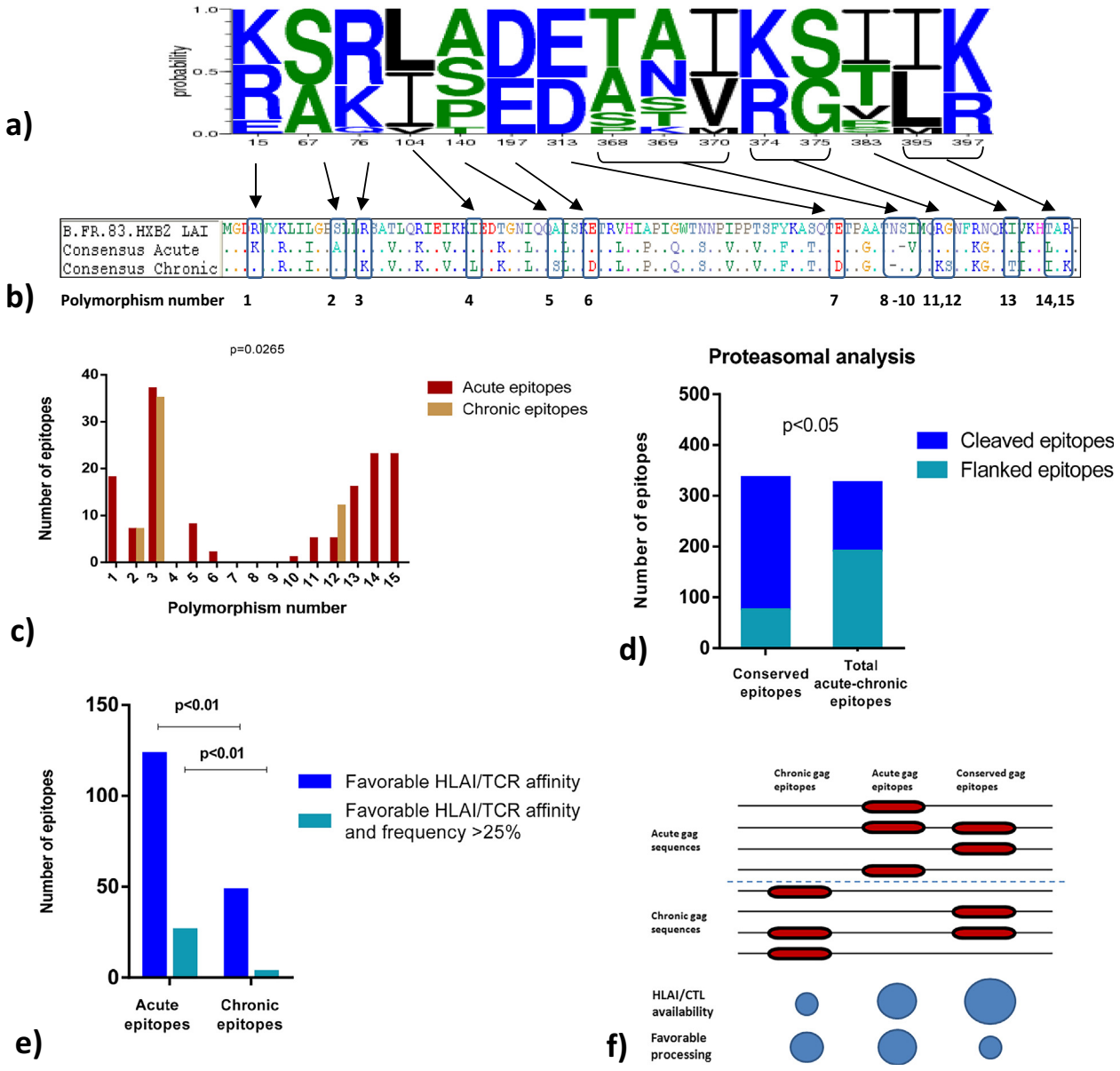

Fig. 4. In (a), the HXB2 sequence position and variability in each polymorphism is observed according to frequency of each amino acid. In (b) the acute and chronic consensus sequences, and prevalent aminoacid in each polymorphism are shown. In (c) the number of acute and chronic epitopes associated with polymorphisms is shown ($p < 0.05$, Paired *t* test). (d) Comparison between cleavage sites of possible acute-chronic epítopes versus conserved possible epitopes. *Cleaved epitopes:* with site of proteosomal cleavage at anchor points to the HLA I molecule or between them. *Flanked epitopes:* with proteasomal cleavage site outside the anchor points. (z test, $p < 0.05$). (e) Comparison of number and frequency of acute and chronic epitopes whit favorable HLAI and TCR affinity (z test, $p < 0.01$). In (f), the results of the variables measured in this research are represented in a schematic summary. It is observed that the possible acute phase epitopes present, in combination, better values compatible with favorable cytotoxic immune response.

associated with the chronic phase (polymorphisms; 1 Arginine, 2 Serine, 3 Lysine, 5 Serine, 6 Aspartic Acid, 10 Isoleucine, 11 Lysine, 12 Serine and 13 Threonine) are associated with 82% of the possible epitopes of chronic phase, but 34.9% of these, (28.7% of the total of the chronic epitopes) included polymorphisms 3, 5, 6, 12 and 13 but not amino acids because they represent escape mutations to HLAI of according to our prediction. Thus, it is possible to infer that while acute consensus polymorphisms are mostly associated with epitopes (67.9%/196 in acute phase vs. 52.7%/129 in chronic phase, $p < 0.01$), which prevails in the consensus polymorphisms of the chronic stage is the escape, Fig. 4c.

Finally, 66.6% of the possible acute phase epitopes associated with consensus polymorphisms were associated with HXB2 consensus polymorphism, however, and as we saw earlier, only 27.5% of these possible epitopes matched with consensus HXB2 in its entirety.

3.5. Prediction of proteasomal activity

The processing of viral antigens by immunoproteasomes and intracellular aminopeptidases as well as the presentation of the epitope is critical for the recognition of pathogen-infected cells by CD8+T lymphocytes [49–51]. On the other hand, mutations surrounding the epitopes they are associated with altered processing and presentation [52].In order to evaluate the cleavage sites in the gag sequences of our study, we used the NetChop 3.1 CBS prediction server (C term 3.0 method, with maximum specificity allowed) [53]. The objective was to analyze the association

between cleavage sites and possible acute, chronic and conserved epitopes. We found that 63% of the possible acute epitopes were flanked by cleavage sites versus 52% of chronic epitopes (non-significant difference with $p > 0.05$) and only 22.5% of the conserved epitopes ($p < 0.05$), Fig. 4d.

3.6. According analysis of the variables under study

So far, the result of epitope predictions in each phase of infection was analyzed separately. Therefore, it was necessary to evaluate, for each acute and chronical epitope, the level of concordance between prediction results for HLA / TCR. We found a favorable HLA/TCR affinity correlation (with HLAI and TCR affinity from 400 nM to 1 nM and 0.01 to 0.99 respectively) of 62.7% (123/196) for acute phase epitopes and 37.2% (48/129) in chronic phase epitopes ($p < 0.01$). In addition, we found that 21.1% (26/123) of the acute phase epitopes with favorable HLA/TCR affinity showed a frequency within the acute phase greater than 25% compared to the equivalents of the chronic phase that was 6.2% (3/48) ($p < 0.01$). These results suggest that, there is a greater correlation of variables associated with immune response within prevalent acute epitopes, Fig. 4e. In addition, Fig. 4f shows a schematic representation of our results.

4. Discussion

Many studies have described the importance of T cell response to counteract viral replication during the acute phase [1,2] and it is also well described how the first escape mutations arise around the first thirty days of infection and how they affect the viral fitness [10,23]. On the other hand, although many studies have detailed the viral dynamics during the first thirty days of infection [38], it is not fully elucidated as the weak specific cytotoxic immune response in this initial period (Fiebig stage I to IV) [3,5] is associated with mutations found in the transmitted/founder virus. In this work, we identify a series of polymorphisms between the acute and chronic stages and analyze their association with prevalent epitopes in each phase of infection in the patients under study.

From cloning and sequencing the gag coding sequence in both groups of patients, we were able to compare the phylogenetic distance, the frequency of possible intrapatient and interpatient epitopes (acute or chronic), the prediction of affinity for the HLA I molecule and the TCR receptor of the possible epitopes found, and in addition, we compared the consensus polymorphisms between both groups and the HXB2 consensus. The results obtained indicated that during the early acute stage of infection, the viral population acquires particular HLA-dependent characteristics that distinguish it from those found in the chronic stage and that these characteristics could be associated with specific cytotoxic immune responses.

First, we found a lower viral phylogenetic distance between patients with acute infection, which suggests that in the absence of specific immune pressure (pre or post-transmission), certain mutations (possibly reversions) are selected which are probably associated with viral fitness rather than factors associated with the host. This lower phylogenetic distance between gag sequences of acute phase, led us to analyze the amino acid positions prevalent in each phase (acute/chronic polymorphisms). We found a significant association between polymorphisms (10/15 66.6% $p < 0.01$) with acute epitopes. In turn, we find that the 67.9% of the acute epitopes were associated whit polymorphisms in comparison with chronic epitopes (52.7%). We saw that this was due to the higher proportion of HLA I A/B escape found in the chronic polymorphisms, whose amino acid positions were included in 28.7% of the chronic epitopes, in comparison with the acute polymorphisms with escape to HLA I A/B whose positions were included in 7.1% of acute epitopes ($p < 0.01$). We deduce that the escape mutations to HLAI (locus A and B) located on polymorphisms 3, 5, 6, 12 and 13, are sufficient to cause the escape in almost one-third (28.7%) of the total potential targets for cytotoxic immune responses on the gag protein analyzed.

We also found that, despite the heterogeneity of HLA I alleles of the patients studied, the acute epitopes had an increased intra- and inter-patient frequency and a lower frequency of possible TCR receptor mutations than those found in chronic stage (this deduction arising from CTL cytotoxic response prediction). During the chronic phase, we believe that the increase in phylogenetic distance observed in comparison with the acute phase occurs due to the progressive adaptation of the viral populations to HLA context in each host, anyway a longitudinal study on the same patients would answer this question better. Interestingly, although most of the possible acute phase epitopes have a HXB2 consensus polymorphism, most of these epitopes do not completely coincide with this consensus. The latter suggests that the design of vaccines based on either epitopes or consensus amino acid positions would not necessarily lead to the best levels of expected immune responses.

Regarding the possible epitopes conserved in both groups, our results on the prediction of proteasomal cleavage sites show that they contain a greater proportion of internal cleavages compared to the acute-chronic epitopes, which reduces their antigenic potential. This may be contributing to the low variability and apparent decreases of escape mutations in these gag segments.

In addition, from the analysis of variability through the measurement of entropy we observed a high association of acute-chronic epitopes with more variable regions, and a significantly lower association with conserved epitopes, which coincides with previous studies [54,55]. Although this association is not fully understood, we can deduce from our results that the variable regions have greater escape / reversion capacity of escape mutations compared to conserved segments, which in turn, showed a favorable prediction of affinity epitope-HLAI/TCR, which is consistent with recent studies [56,57].

Given that the number of patients in this research is limited to estimate population prevalence of epitopes and acute-chronic polymorphisms, it will be necessary to extend this work in the future. In turn, the prediction algorithms used in this work have an intrinsic estimation error of each method (see methods), such as affinity of the different proteins involved with the peptides, the length of the peptides analyzed, among others. In addition, more studies will be needed to determine the efficacy of HLA/TCR binding peptide predictors performed with circulating recombinant forms (CRFs) BF, because, although there are studies that suggest a good correlation of prediction with immune response assays for clade B [58,59], the same result was not found for clade C [60]. Therefore, it will be necessary to check the immunogenicity of the prevalent epitopes in each face with *in vitro* immune response assays in future research.

Overall, results presented here show that, during the acute stage, the virus acquires mutations that indicate certain vulnerability associated with cytotoxic specific immune response, and that certain viral residues, particularly non-consensus, can be transmitted more frequently, or selected during the acute phase, and are found predominantly in positions with immunological relevance (epitopes). This observation could contribute to the development of an effective vaccine capable of reducing the viral reservoir and inducing a functional cure of HIV infection based on prevalent acute epitopes.

Acknowledgments

This research was funded by grants from the Agencia Nacional de Promoción Científica y Tecnológica (PICT 2008-00771 and PICT 2011-0271) to H. Salomon.

We thank all INBIRS colleagues for their interest and direct or indirect collaboration in this research.

We thank all the physicians belonging to the Grupo Argentino de Seroconversión and Fundación Huésped for their clinical contribution and especially to their patients for their good predisposition to participate in this research. Also thanks to Paul Kerston for assistance during manuscript preparation.

We also thank Adrian Perusco, Pablo Galarza and Carina Lopez from the immunogenetic laboratory of the Instituto de Investigaciones Médicas Alfredo Lanari for their collaboration in the histocompatibility analysis. We thank the Universidad Nacional de San Luis (UNSL-FQByF) for their support and collaboration.

We have no conflict of interest to testify.

References

[1] Walker B, McMichael A. The T-cell response to HIV. Cold Spring Harb Perspect Med. 2012;2(11). pii: a007054. doi: 10.1101/cshperspect.a007054.
[2] Goonetilleke N, Liu MK, Salazar-Gonzalez JF, Ferrari G, Giorgi E, Ganusov VV, et al. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. J Exp Med 2009;206(6):1253–72. doi: 10.1084/jem.20090365. Epub 2009 Jun 1.
[3] Dalod M, Dupuis M, Deschemin JC, Goujard C, Deveau C, Meyer L, et al. Weak anti-HIV CD8(+) T-cell effector activity in HIV primary infection. J Clin Invest 1999;104(10):1431–9.
[4] Altfeld M, Rosenberg ES, Shankarappa R, Mukherjee JS, Hecht FM, Eldridge RL, et al. Cellular immune responses and viral diversity in individuals treated during acute and early HIV-1 infection. J Exp Med 2001;193 (2):169–80.
[5] Radebe M, Nair K, Chonco F, Bishop K, Wright JK, van der Stok M, et al. Limited immunogenicity of HIV CD8+ T-cell epitopes in acute Clade C virus infection. J Infect Dis 2011;204(5):768–76. doi: 10.1093/infdis/jir394.
[6] Edwards BH, Bansal A, Sabbaj S, Bakari J, Mulligan MJ, Goepfert PA. Magnitude of functional CD8+ T-cell responses to the gag protein of human immunodeficiency virus type 1 correlates inversely with viral load in plasma. J Virol Mar 2002;76(5):2298–305.
[7] Zuñiga R, Lucchetti A, Galvan P, Sanchez S, Sanchez C, Hernandez A, et al. Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. J Virol 2006;80 (6):3122–5.
[8] Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat Med 2007;13(1):46–53. Epub 2006 Dec 17.
[9] Turk G, Ghiglione Y, Falivene J, Socias ME, Laufer N, Coloccini RS, et al. Early Gag immunodominance of the HIV-specific T-cell response during acute/early infection is associated with higher CD8+ T-cell antiviral activity and correlates with preservation of the CD4+ T-cell compartment. J Virol 2013;87(13):7445–62. doi: 10.1128/JVI.00865-13. Epub 2013 Apr 24.
[10] Martinez-Picado J, Prado JG, Fry EE, Pfafferott K, Leslie A, Chetty S, et al. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. J Virol 2006;80(7):3617–23.
[11] Brumme ZL, Tao I, Szeto S, Brumme CJ, Carlson JM, Chan D, et al. Human leukocyte antigen-specific polymorphisms in HIV-1 Gag and their association with viral load in chronic untreated infection. AIDS 2008;22(11):1277–86. doi: 10.1097/QAD.0b013e3283021a8c.
[12] Sunshine JE, Larsen BB, Maust B, Casey E, Deng W, Chen L, et al. Fitness-balanced escape determines resolution of dynamic founder virus escape processes in HIV-1 Infect J Virol 2015;89(20):10303–18. doi: 10.1128/JVI.01876-15. Epub 2015 Jul 29.
[13] Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM, Feeney M, et al. HIV evolution: CTL escape mutation and reversion after transmission. Nat Med 2004;10(3):282–9. Epub 2004 Feb 8.
[14] Nagaraja Pradeep, Alexander Helen K, Bonhoeffer Sebastian. Influence of recombination on acquisition and reversion of immune escape and compensatory mutations in HIV-1. Epidemics 2016;14:11–25. https://doi.org/10.1016/j.epidem.2015.09.001.
[15] Li B, Gladden AD, Altfeld M, Kaldor JM, Cooper DA, Kelleher AD, et al. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. J Virol 2007;81(1):193–201. Epub 2006 Oct 25.
[16] Duda A, Lee-Turner L, Fox J, Robinson N, Dustan S, Kaye S, et al. HLA-associated clinical progression correlates with epitope reversion rates in early human immunodeficiency virus infection. J Virol. 2009;83(3):1228–39. doi: 10.1128/JVI.01545-08. Epub 2008 Nov 19.
[17] Gounder K, Padayachi N, Mann JK, Radebe M, Mokgoro M, van der Stok M, et al. High frequency of transmitted HIV-1 Gag HLA class I-driven immune escape variants but minimal immune selection over the first year of clade C infection. PLoS One 2015;10(3):e0119886. https://doi.org/10.1371/journal.pone.0119886. eCollection 2015.
[18] Liu Y, McNevin J, Zhao H, Tebit DM, Troyer RM, McSweyn M, et al. Evolution of human immunodeficiency virus type 1 cytotoxic T-lymphocyte epitopes: fitness-balanced escape. J Virol 2007;81(22):12179–88.
[19] Troyer RM, McNevin J, Liu Y, Zhang SC, Krizan RW, Abraha A, et al. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. PLoS Pathog 2009;5(4):e1000365. https://doi.org/10.1371/journal.ppat.1000365.
[20] Yokomaku Y, Miura H, Tomiyama H, Kawana-Tachikawa A, Takiguchi M, Kojima A, et al. 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 2004;78 (3):1324–32.
[21] Cardinaud S, Consiglieri G, Bouziat R, Urrutia A, Graff-Dubois S, Fourati S, et al. CTL escape mediated by proteasomal destruction of an HIV-1 cryptic epitope. PLoS Pathog 2011;7(5):e1002049. https://doi.org/10.1371/journal.ppat.1002049.
[22] Zanini F, Brodin J, Thebo L, Lanz C, Bratt G, Albert J, et al. Population genomics of intrapatient HIV-1 evolution. Elife 2015;4. pii: e11282. doi: 10.7554/eLife.11282.
[23] J Virol 2015 Oct;89(20):10303–18. doi: 10.1128/JVI.01876-15.
[24] Kløverpris HN, Leslie A, Goulder P. Role of HLA adaptation in HIV evolution. Front Immunol 2016;18(6):665. https://doi.org/10.3389/fimmu.2015.00665.
[25] Byrn RA, Kiessling AA. Analysis of human immunodeficiency virus in semen: indications of a genetically distinct virus reservoir. J Reprod Immunol 1998;41 (1–2):161–76.
[26] Anderson JA, Ping LH, Dibben O, Jabara CB, Arney L, Kincer L, et al. Center for HIV/AIDS Vaccine Immunology. HIV-1 Populations in Semen Arise through Multiple Mechanisms. PLoS Pathog 2010;6(8):e1001053.
[27] Boeras DI, Hraber PT, Hurlston M, Evans-Strickfaden T, Bhattacharya T, Giorgi EE, et al. Role of donor genital tract HIV-1 diversity in the transmission bottleneck. Proc Natl Acad Sci U S A. 2011;108(46):E1156–E1163. Published online 2011 Nov 7. doi: 10.1073/pnas.1103764108.
[28] Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS 2003;17 (13):1871–9.
[29] Dilernia D, Chien JT, Monaco DC, Brown MP, Ende Z, Deymier MJ, et al. Multiplexed highly-accurate DNA sequencing of closely-related HIV-1 variants using continuous long reads from single molecule, real-time sequencing. Nucl Acids Res 2015;43(20):e129.
[30] Oh SH, Fleischhauer K, Yang SY. Isoelectric focusing subtypes of HLA-A can be defined by oligonucleotide typing. Tissue Antigens 1993;41:135.
[31] Middleton D, Williams F, Cullen C, Mallon E. Modification of an HLA-B PCR-SSOP typing system leading to improved allele determination. Tissue Antigens 1995;45(4):232–6. Tissue Antigens 45:232.
[32] Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. J Mol Evol, Academic Press, New York; 1965. p. 97–166.
[33] Tajima F, Nei M. Estimation of evolutionary distance between nucleotide sequences. Mol Biol Evol. 1984;1(3):269–85.
[34] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecul Biol Evol 2013;30:2725–9.
[35] Roberts RJ, Carneiro MO, Schatz MC. The advantages of SMRT sequencing. Genome Biol 2013;14(7):405. https://doi.org/10.1186/gb-2013-14-6-405.
[36] Brinzevich D, Young GR, Sebra R, Ayllon J, Maio SM, Deikus G, et al. HIV-1 interacts with human endogenous retrovirus K (HML-2) envelopes derived from human primary lymphocytes. J Virol 2014;88(11):6213–23. doi: 10.1128/JVI.00669-14. Epub 2014 Mar 19.
[37] Dilernia DA, Jones L, Rodriguez S, Turk G, Rubio AE, Pampuro S, et al. HLA-driven convergence of HIV-1 viral subtypes B and F toward the adaptation to immune responses in human populations. PLoS ONE 2008;3(10):e3429. https://doi.org/10.1371/journal.pone.0003429.
[38] Carlson Jonathan M, Schaefer Malinda, Monaco Daniela C, Batorsky Rebecca, Claiborne Daniel T, Prince Jessica, et al. Selection bias at the heterosexual HIV-1 transmission bottleneck. Science 2014;345(6193):1254031. https://doi.org/10.1126/science.1254031.
[39] Gnanakaran S, Bhattacharya T, Daniels M, Keele BF, Hraber PT, Lapedes AS, et al. Recurrent signature patterns in HIV-1 B clade envelope glycoproteins associated with either early or chronic infections. PLoS Pathog 2011;7(9): e1002209.
[40] Chohan B, Lang D, Sagar M, Korber B, Lavreys L, Richardson B, et al. Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1–V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. J Virol 2005;79 (10):6528–31.
[41] Sagar M, Laeyendecker O, Lee S, Gamiel J, Wawer MJ, Gray RH, et al. Selection of HIV variants with signature genotypic characteristics during heterosexual transmission. J Infect Dis 2009;199(4):580–9. https://doi.org/10.1086/596557.
[42] Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 2003;12:1007–17.

[43] Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 2016;32 (4):511–7. doi: http://10.1093/bioinformatics/btv639. Epub 2015 Oct 29.
[44] Larsen MV, Lundegaard C, Lamberth K, Buus S, Brunak S, Lund O, et al. An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. Eur J Immunol 2005;35(8):2295–303.
[45] Nicole Frahma, Brett Bakera, Christian Brandera. Identification and optimal definition of HIV-Derived Cytotoxic T Lymphocyte (CTL) epitopes for the study of CTL escape, functional avidity and viral evolution. HIV Molecul Immunol; 2008.
[46] Anuska L, Frahmb Nicole, Brander Christian. How to Optimally Define Optimal Cytotoxic T Lymphocyte Epitopes in HIV Infection? HIV Molecular Immunology. Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory; 2009.
[47] Brumme Zabrina L, Poon Art FY, Carlson Jonathan M, Walker Bruce D. Identifying HLA-Associated Polymorphisms in HIV-1. HIV Molecular Immunology; 2010. Escape.tex r6004–6892.
[48] Anuska L, Ashley Williams, Alex Olvera, Sandra Silva-Arrieta, Christian Brander. Best-characterized HIV-1 CTL epitopes: the 2013 update. HIV Mol Immunol, 2013, HIV Molecular Immunology. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos.
[49] Kisselev AF, Akopian TN, Woo KM, Goldberg AL. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. J Biol Chem 1999 Feb 5;274(6):3363–71.
[50] Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, van Veelen P, et al. Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I. Immunity 2003 Jan;18(1):97–108.
[51] Hutchinson Sarah, Sims Stuart, O'Hara Geraldine, Silk Jon, Gileadi Uzi, Cerundolo Vincenzo. A dominant role for the immunoproteasome in CD8+ T cell responses to murine cytomegalovirus. PLoS One 2011;6(2):e14646. https://doi.org/10.1371/journal.pone.0014646.
[52] Le Gall S, Stamegna P, Walker BD. Portable flanking sequences modulate CTL epitope processing. J Clin Invest 2007;117(11):3563–75.
[53] Kesmir C, Nussbaum AK, Schild H, Detours V, Brunak S. Prediction of proteasome cleavage motifs by neural networks. Protein Eng 2002 Apr;15 (4):287–96.
[54] Bansal Anju, et al. CD8 T-cell responses in early HIV-1 infection are skewed towards high entropy peptides. AIDS 2005;19(3):241–250.
[55] Ferrari G, Korber B, Goonetilleke N, Liu MK, Turnbull EL, Salazar-Gonzalez JF, et al. Relationship between functional profile of HIV-1 specific CD8 T cells and epitope variability with the selection of escape mutants in acute HIV-1 infection. PLoS Pathog 2011;7(2):e1001273. doi: 10.1371/journal.ppat.1001273.
[56] Ahmed T, Borthwick NJ, Gilmour J, Hayes P, Dorrell L, Hanke T. Control of HIV-1 replication in vitro by vaccine-induced human CD8(+) T cells through conserved subdominant Pol epitopes. Vaccine 2016;34(9):1215–24. https://doi.org/10.1016/j.vaccine.2015.12.021. Epub 2016 Jan 16.
[57] Borthwick N, Lin Z, Akahoshi T, Llano A, Silva-Arrieta S, Ahmed T, et al. Novel, in-natural-infection subdominant HIV-1 CD8+ T-cell epitopes revealed in human recipients of conserved-region T-cell vaccines. PLoS One 2017;12(4): e0176418. doi: http://10.1371/journal.pone.0176418. eCollection 2017.
[58] Turk Gabriela, Gherardi María Magdalena, Laufer Natalia, Saracco Mónica, Luzzi Renata, Cox Josephine H, et al. Magnitude, breadth, and functional profile of T-cell responses during human immunodeficiency virus primary infection with B and BF viral variants. J Virol 2008;82(6):2853–66.
[59] Park H, Chung YS, Yoon CH, Lee SH, Kim SS, Kang C, et al. Presentation of available CTL epitopes that induction of cell-mediated immune response against HIV-1 Koran clade B strain using computational technology. HIV Med 2016;17(6):460–6. doi: 10.1111/hiv.12316. Epub 2015 Oct 15.
[60] Ngandu NG, Bredell H, Gray CM, Williamson C, Seoighe C, HIVNET028 Study Team. CTL response to HIV type 1 subtype C is poorly predicted by known epitope motifs. AIDS Res Hum Retrovirus 2007;23(8):1033–41.