VIROLOGICA SINICA, December 2012, 27 (6):360-368
DOI 10.1007/s12250-012-3287-0

Research Article

Specific VpU Codon Changes were Significantly associated with gp120 V3 Tropic Signatures in HIV-1 B-subtype

Salvatore Dimonte[1]✉ Muhammed Babakir-Mina[1,2], Stefano Aquaro[3] and Carlo-Federico Perno[1,4]

1. Department of Experimental Medicine and Surgery, University of Rome "Tor Vergata", Rome 00133, Italy;
2. Slemani Polytechnic University; Iraqi Kurdistan Region;
3. Faculty of Pharmacy, Nutritional and Health Sciences, University of Calabria, Arcavacata di Rende (CS) 87036, Italy;
4. National Institute for Infectious Diseases "L. Spallanzani", Rome 00149, Italy.

After infection and integration steps, HIV-1 transcriptions increase sharply and singly-spliced mRNAs are produced. These encode Env (gp120 and gp41) and auxiliary proteins Vif, Vpr and VpU. The same localization within the unique structure of the mRNAs suggests that the *VpU* sequence prior to the *Env* could affect the Env polyprotein expression.The HIV-1 infection process begins when the gp120 subunit of the envelope glycoprotein complex interacts with its receptor(s) on the target cell. The V3 domain of gp120 is the major determinant of cellular co-receptor binding. According to phenotypic information of HIV-1 isolates, sequences from the VpU to V3 regions (119 in R5- and 120 X4-tropic viruses; one *per* patient) were analysed. The binomial correlation phi coefficient was used to assess covariation among VpU and gp120$_{V3}$ signatures. Subsequently, average linkage hierarchical agglomerative clustering was performed. Beyond the classical V3 signatures (R5-viruses: S11, E25D; X4-viruses: S11KR, E25KRQ), other specific V3 and novel VpU signatures were found to be statistically associated with co-receptor usage. Several statistically significant associations between V3 and VpU mutations were also observed. The dendrogram showed two distinct large clusters: one associated with R5-tropic sequences (bootstrap=0.94), involving: (a) H13NP$_{V3}$, E25D$_{V3}$, S11$_{V3}$, T22A$_{V3}$ and Q61H$_{VpU}$, (b) E25A$_{V3}$ and L12F$_{VpU}$, (c) D44E$_{VpU}$, R18Q$_{V3}$ and D80N$_{VpU}$; and another associated with X4-tropic sequences (bootstrap=0.97), involving: (i) E25I$_{V3}$ and V10A$_{VpU}$, (ii) 0-1insV$_{VpU}$, H13R$_{V3}$, I46L$_{VpU}$, I30M$_{V3}$ and 60-62del$_{VpU}$, (iii) S11KR$_{V3}$ and E25KRQ$_{V3}$. Some of these pairs of mutations were encoded always by one specific codon. These data indicate the possible VpU mutational patterns contributing to regulation of HIV-1 tropism.

HIV; VpU; gp120 V3; Mutations; Tropism; Cluster analysis

The Human Immunodeficiency Virus type 1 (HIV-1) protein U (VpU) is a membrane-associated "auxiliary" protein, approximately 16 kDa, that is unique to HIV-1 and a subset of the related Simian Immunodeficiency Virus (SIV)[3]. This oligomeric integral membrane protein is predicted to have a short luminal N-terminal domain, a single transmembrane (TM) spanning domain at the amino terminus that serves as an uncleaved signal peptide, and two cytoplasmic charged hydrophilic α-helices-interconnected by a flexible loop containing a highly conserved phosphorylated sequence ($D_{52}SGNES_{57}$)-at the carboxyl terminus[35]. Moreover, the TM was found to down-regulate the CD4 receptor[27], while the deletion of the highly negatively charged C-terminal end (with three aspartic residues in the last five amino acids) was observed to be detrimental to *in vitro* VpU-induced CD4 degradation[6].

Received: 2012-09-17, Accepted: 2012-10-24
✉ Corresponding author.
Phone: +96-47-480190092\93, Fax: +39-06-72596039,
Email: salvatore.dimonte@gmail.com

VpU is associated with two primary functions during the HIV-1 life cycle. First, it contributes to viral-induced CD4 receptor down-regulation by mediating the proteasomal degradation of newly synthesized CD4 molecules in the endoplasmic reticulum. Secondly, it enhances the release of viral progeny from infected cells by antagonizing tetherin that is an interferon (IFN)-regulated host restriction factor which directly cross-links virions on the host cell surface[14]. Recently, a novel function of VpU related to stabilization of tumour suppressor p53 was reported: the modulation of the p53-stability correlates positively with apoptosis during late stages of HIV-1 infection[43]. These observations allow us to consider VpU to be an antagonist of the innate immune response to viral infection.

As with all other retroviruses, HIV-1 employs a variety of different overlapping reading frames and splicing events to express a large array of mRNAs (at least 40) and proteins (at least 15) from a single primary transcript. During HIV-1 viral replication, three classes of RNAs are produced: (i) the early doubly spliced approximately 2-kb transcripts encoding Tat, Rev, and Nef; (ii) the late singly spliced mRNAs encoding Vif, Vpr, VpU, and Env products; and (iii) the late approximately 9-kb unspliced mRNAs that are packaged into progeny virions as genomic RNA and which can also serve for the expression of Gag/Pol genes[12, 24]. Specifically, the viral envelope (Env) glycoprotein is translated from about 16 alternatively spliced 4 kb bicistronic mRNA Rev-dependent isoforms that all contain the upstream ORF for VpU. (Rev is a virally encoded sequence-specific RNA-binding protein that shuttles viral mRNAs between the nucleus and the cytoplasm). The biosynthesis of the Env polyprotein - starting about 160 nucleotides downstream from the VpU start codon - implies that the expression of VpU and Env from the same 4-kb spliced mRNA isoforms through leaky scanning translation are coordinated during HIV-1 infection[8, 26, 36].

HIV-1 entry into the host cell is mediated by the Env glycoproteins, gp120 and gp41: the initial binding of gp120 to the cellular CD4 receptor triggers conformational changes in gp120 that promote its following interaction with one of the chemokine co-receptors, usually CCR5 or CXCR4. HIV-1 strains can be phenotypically classified according to their ability to use the CCR5 and/or CXCR4 co-receptor: this binding is based upon the presence of selected amino acids in gp120 (specifically within the V3-loop, but also in other proteic regions) providing greater affinity to CCR5 or CXCR4, and therefore the viral tropism[11, 45].

In this light, the aim of the present study was to genetically characterize HIV-1 B-subtype $gp120_{V3}$ and VpU sequences in terms of co-receptor usage and to define the association of amino acid changes within the V3 and the VpU regions according to CCR5 and/or CXCR4 usage.

MATERIALS AND METHODS

A large cohort of 239 HIV-1 subtype-B sequences containing the specific genomic region from VpU to gp120 V3 domain, all retrieved from Los Alamos Database (one sequence *per* individual) [http://www.hiv.lanl.gov] were analyzed. All analyzed sequences have a pure phenotype and/or co-receptor determinations available (119 R5- and 120 X4-using viruses, respectively). Sequences were all retrieved from infected individuals at all stages of infection and collected from several countries representing all geographic continents (sampling years: 1983-2009). The drug treatment status for the individuals is not available in the Los Alamos Database, but some sequences were retrieved from ART naïve patients (16 R5- and 2 X4-tropic viruses, respectively). As previously reported[11,13], the multiple sequence alignments of VpU and V3 viral segments were performed by using ClustalX and Bioedit software packages [http://www.es.embnet.org http://www.mbio.ncsu.edu]. To confirm the HIV-1 subtype-B, published consensus sequences of pure HIV-1 subtypes (A, B, C, D, F1, F2, G, H, J, and K) were used [http://www.hiv.lanl.gov], and multi-aligned sequences were subjected to phylogenetic inference[11,13]. To analyze VpU and V3 signatures, the frequency of all substitutions in the 81 VpU amino acids and 35 V3 amino acids were calculated. Fisher exact tests were used to determine the differences in frequency between the 2 groups of patients (infected with R5- or X4-using viruses, respectively). The Benjamini- Hochberg method was been used to identify results that were statistically significant in the presence of multiple-hypothesis testing[2]. A false discovery rate of 0.055 was used to determine statistical significance.

To identify significant patterns of pairwise

associations between V3 and VpU amino acid changes, the phi coefficient (φ) and its statistical significance for each pair of mutations were calculated[11, 13, 40].

RESULTS

Consistent with previous observations, the first analysis was to confirm the classical V3 positions 11 and 25. The serine wild-type at position 11 (indicated as S11S) and the E25D mutation were significantly associated with R5-tropic viruses, while the S11KR and E25KRQ mutations were significantly associated with CXCR4 co-receptor usage[9, 15] (Fig.1a).

By evaluating the phenotypic dataset of V3-loop sequences, we tested our previous genotypic results[11]: 8 out of 9 V3 previously predicted mutations (89%) were present with a prevalence significantly higher in R5- than in X4-tropic viruses (the known mutations were E25D, and H13N, H13P, G15A, R18Q, F20L, Y21F and T22A) ($P < 0.05$) (Fig.1a). In addition, the following V3 signatures, R18G, A19T, G24gap, and E25A, were found to be significantly associated with R5-tropic viruses ($P < 0.05$) (Fig.1a).

Moreover, 28 out of 33 previously predicted amino acid changes (85%) were significantly more prevalent in X4- than in R5-tropic viruses (the known mutations were the classical S11KR and E25KRQ, and T2M, N6Y, N7K, T8I, R9K, I12M, H13RSTY, P16Q, A19V, F20IVY, Y21HV, G24EKR, I27V, R31K, Q32K) ($P < 0.05$) (Fig.1a). The following V3 signatures, N5Y, N6IKR, N7Y, T8KR, R9I, I14L, T23I, G24D, E25I, I27gap, and I30M, were found to be significantly associated with X4-tropic viruses ($P < 0.05$) (Fig.1a).

The conserved V3 region encompassing the residues 4 to 7 (P4-N5-N6-N7) has been demonstrated to interact with the CCR5 co-receptor: the binding of this co-receptor is blocked when N7 is replaced by a charged amino acid[20]. As previously observed [11], the phenotypic dataset in the present study indicates that the mutation N7K has been found only in X4-tropic viruses (prevalence:10.8%; $P = 0.0002$) (Fig.1a).

By analyzing the VpU sequences, 23 out of 81 VpU positions were found to be significantly associated with different co-receptor usage for the first time ($P < 0.05$) (Fig.1b). In particular, 11 VpU signatures whose prevalence was significantly higher in R5- than in

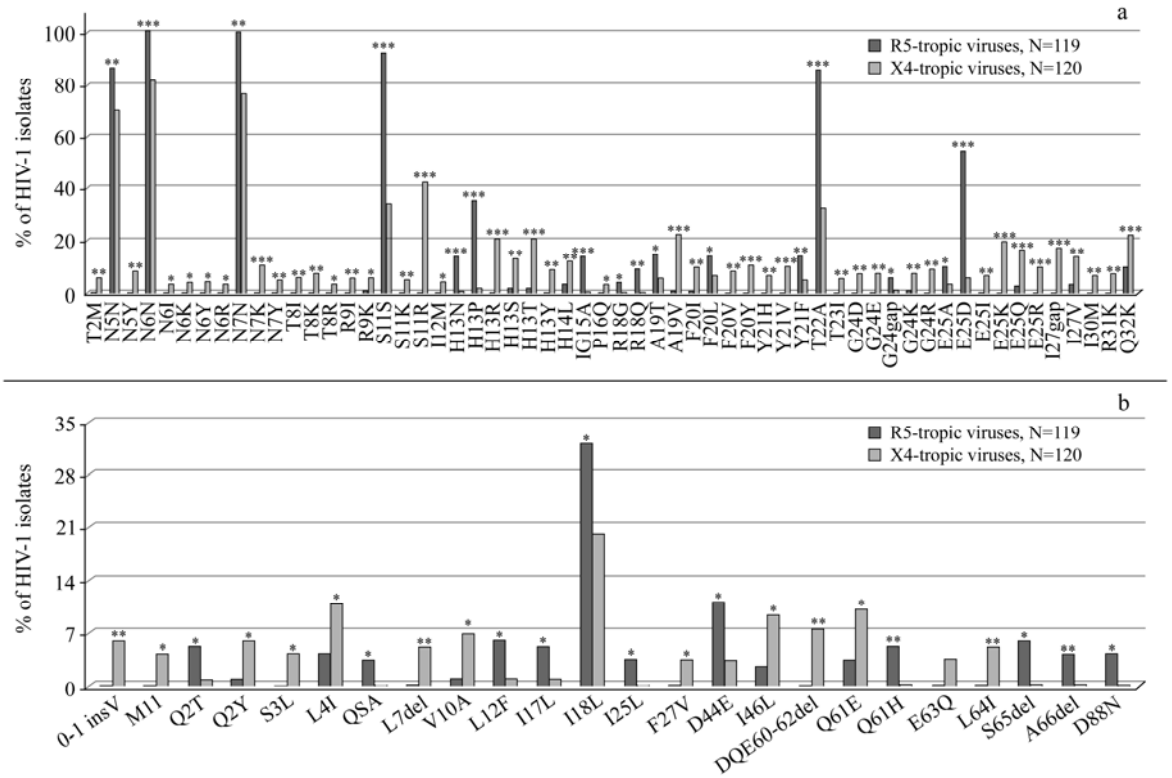

Fig. 1. Frequencies of HIV-1 gp120$_{V3}$ and VpU amino acid changes. Frequencies of V3 (panel "a") and VpU (panel "b") mutations in HIV-1 R5-tropic and HIV-1 X4-tropic isolates. The analysis was performed in sequences derived from 239 patients, 119 reported as R5-tropic and 120 reported as X4-tropic at phenotypic test. Statistically significant differences were assessed by chi-square tests of independence. *P* values were significant at a false-discovery rate of 0.05 following correction for multiple tests. *, $P < 0.05$; **, $P \leqslant 0.01$; ***, $P \leqslant 0.001$.

X4-tropic viruses, were identified: 7 of them had a prevalence > 5% in R5-predicted viruses (Q2T, L12F, I17L, I18L, D44E, Q61H, and S65del) (Fig.1b). Conversely, 13 substitutions whose prevalence was significantly higher in X4- than in R5-viruses were identified, suggesting their association with CXCR4-usage. Among the 13 substitutions, 9 signatures had a prevalence > 5% in X4-tropic viruses (0-1insV, Q2Y, L4I, L7del, V10A, I46L, DQE60-62del, Q61E, and L64I) (Fig. 1b).

Mutation of the VpU start codon occurs at low frequency during PBMC culture of HIV-1 isolates *in vitro* (1.2%; 12 of 967)[10]: recently, it has been shown that the HIV-1 YU-2 clone (highly macrophage-tropic) carries a mutated VpU start-codon. A low frequency (nearly 5%) of VpU sequences in the HIV sequence database contain a mutated *AUG* translation start codon[10]. This mutation failed to confer macrophage replication[33]. Our results are consistent with previous observations: 0-1insV and M1I are signatures retrieved only in X4-tropic viruses, with 5.8% and 4.2% of prevalence, respectively (Fig. 1b).

Several residues associated with different co-receptor usage reside within the VpU TM domain (amino acids 6-28) (L7del, V10A, L12F, I17L, I18L, I25L, and F27V): the fact that all these signatures are localized in the unexposed domain of the protein suggests that they may act as a scaffold in order to maintain the stability of the phospholipids-protein interaction. Notably, the changes are biochemically hydrophobic (non-polar) amino acid substitutions that are not expected to perturb the protein significantly (and so might be conveniently classified as good conservative substitutions): this supports the conservation of the TM domain functionality. Biochemical and biophysical evidence have suggested that this is critical for the VpU oligomerization (most favourably as a pentameric form)[21], as well as for down-regulating the CD4 receptor[27]. Moreover, a single mutation at VpU position 18 by a histidine has been shown to make the virus infection sensitive to the M2 channel blocker Adamantanes[19,31]. In our dataset, the I18H mutation was absent, while I18L was the more prevalent mutation retrieved in all sequences (31.9% in R5- and 20.0% X4-tropic virus, respectively), followed by the polymorphism I18V (1.7% in R5- and 2.5% X4-tropic virus, respectively). Thus, only conservative amino acid changes in this position were found: the local modification of the primary structure plays an important role in secondary and tertiary structures of VpU TM in lipid bilayers and may affect its ability to interact with channel blockers.

Conversely, for the five minor variants observed in the predicted short luminal N-terminal domain (Q2T, Q2Y, S3L, L4I, and Q5A), with the exception of L4I, range of amino acid changes is relatively un conserved (different biochemical functionality)[42]: semi-conservative substitutions (Q2T and Q5A) and non-conservative changes (Q2Y and S3L) were observed. Although the accessory proteins are in general not necessary for viral propagation in tissue culture, several experimental observations suggest that their role *in vivo* is very important[14, 24]. This would suggest that, the short luminal peptide probably has no essential and mandatory role in biochemical conservation and infectivity.

Previous analysis of the VpU cytosolic domain have revealed the presence of two putative trafficking signals (amino acid regions: 29-34 and 63-68) that harbour a degree of amino acid variation among different subtypes[32]. For the first signal region (amino acids 29-34), in the present study the analysis of the prevalence mutations has evidenced a similar variability between R5- and X4-tropic viruses (Fig. 1b). Likewise, for the flexible loop $D_{52}SGNES_{57}$ at the VpU carboxyl terminus, that interconnects the two cytoplasmic charged hydrophilic α-helices, we did not observe the specific tropic-signatures.

On the other hand, in the first charged hydrophilic α-helix the D44E (in R5-tropic viruses) and the I46L (in X4-tropic viruses) were found (Fig. 1b), while several tropic-related positions were found in the second α-helix containing the other trafficking domain (Q61H, S65del, and A66del, for the R5-tropic viruses, and DQE60-62del, Q61E, E63Q, and L64I, for the X4-tropic viruses) (Fig. 1b).

Finally, only the D80N R5-tropic signature was observed in the C-terminal end of VpU (amino acids 76-81) (Fig. 1b). Interestingly, the deletions in several regions of VpU were observed both in R5- and in X4-tropic viruses (residue position 7, 60-62, 65 and 66): on the contrary, deletion of the C-terminal residues would not be allowed because it would be deleterious to the virus, probably resulting in the loss of interaction with CD4 and the inability to induce CD4 degradation[29].

Although the second part of the encoding *VpU* sequence overlaps the 5' end signal sequence of HIV-1

Env (starting from the codon of amino acid residue $N55_{VpU}$)[16], only a third of "tropic-signatures" within this region were observed (Fig.1b). Interestingly, this overlapping region was found to be related to different viral effects. Michalski *et al*. have shown that the signal peptide of HIV-1 Env itself has a direct role in cellular cytotoxicity and the triggering of cell death pathways[30].

At the same time, we underline that various classes of small non-coding RNAs (sncRNAs) are important regulators of gene expression across the divergent types of the organisms, with virally encoded sncRNAs (particularly those of RNA viruses) expressed at very low levels. Althaus and his co-workers have captured almost 900 HIV-1 sncRNAs, some of them distinguishable retrieved in the *VpU*, *Env* and *VpU*/E*nv* overlapping viral regions[1]. In this last cited overlapping region a moderate variability related to the viral tropism was allowed. The HIV-1 encoded sncRNAs vary in length and in their locations on the viral genome and they may have the potential to play roles in HIV-1 replication[1]. Thus, we wanted to emphasize these aspects in relation to the results of VpU association with the first step of viral cycle (the binding between viral envelope proteins and the host cellular surface).

Statistically-significant correlations among V3 and VpU signatures in HIV-1 clade B were identified for the first time (Table 1). Some of these correlations involved the classical V3 positions 11 and 25 (Table 1). Specifically, the $E25KRQ_{V3}$ mutations showed positive correlations with several VpU signatures (0-1insV, I46L, and 60-62del). Similarly, $S11KR_{V3}$ mutations showed positive correlations with $I46L_{VpU}$ and $60\text{-}62del_{VpU}$ signatures (Table 1). All these amino acid variants correlate with CXCR4-usage. Conversely, $S11S_{V3}$ established a positive correlation with only one VpU amino acid change, Q61H ($P = 0.041$; $\varphi = 0.14$), localized in the second trafficking domain contained within the second cytosolic charged hydrophilic α-helix (Table 1). Of note, $S11S_{V3}$ was found in 100% of patients with $Q61H_{VpU}$: the association of these two amino acid residues has always shown codon *cau* for the $Q61H_{VpU}$ and codon *agu* for $S11S_{V3}$, further supporting that these signatures are strongly correlated with each other both at the nucleotide and amino acid level.

However, the classical $E25D_{V3}$ R5-tropic mutation was not established to have a statistically significant correlation with VpU amino acid residues: on the other hand, the classically strong association with $S11S_{V3}$ was confirmed ($P = 3.54e^{-11}$; $\varphi = 0.41$).

Among the positive correlations between V3 and VpU signatures associated with CXCR4-usage, strong correlations were observed for $I30M_{V3}$ with $I46L_{VpU}$ ($P = 3.61e^{-9}$; $\varphi = 0.65$) and with $60\text{-}62del_{VpU}$ ($P = 3.86e^{-11}$; $\varphi = 0.82$) (Table 1). Of note, $I30M_{V3}$ was found in 50% of patients with $I46L_{VpU}$, and in 77.8% of patients with $60\text{-}62del_{VpU}$: in the co-variation frequency the $I46L_{VpU}$ mutation has shown to always have the codon *cua* (for the $I30M_{V3}$ substitution, it is known that the metionine was encoded by only one codon), further supporting the case that these substitutions are strongly correlated with each other both at the nucleotide and amino acid level (Table 1). Moreover, it has been recently shown that $I30_{V3}$ is critically important for efficient HIV-1 entry into macrophages[5, 17]: the M30I in non-M-tropic CXCR4-tropic Envs enhances macrophage entry *via* CXCR4. These reports support a critical role for I30 in promoting efficient X4-mediated HIV-1 entry into macrophages[5, 17].

Another positive correlation between V3 and VpU signatures associated with CXCR4-tropic was the $E25I_{V3}$ with $V10A_{VpU}$ ($P = 0.032$; $\varphi = 0.21$) (Table 1). In an *in vivo* study new V3-genetic mutations modulating co-receptor usage were identified[41], and the $E25I_{V3}$ was listed as a significantly associated minor variant with CXCR4-tropic (with low prevalence). In the present analysis this preliminary observation was confirmed, with a prevalence of 6.7% in X4- vs 0% in R5-tropic viruses, respectively ($P = 0.004$).

In the V3-loop it has been shown that the secondary structure of the GPGX crown (at positions 15-18) is crucial in modulation of HIV-1 subtype co-receptor usage. This motif forms a proteic β-turn that binds to the extracellular loops of the co-receptor[23]. Among the positive correlations between V3 and VpU mutations associated with CCR5-usage, a correlation for $D80N_{VpU}$ with a $R18Q_{V3}$ substitution was observed ($P = 0.018$; $\varphi = 0.25$). The V3 position 18 - along with position 20 - resides in a motif shown to be involved in the binding with two specific glycosphingolipids (GSLs): galactosylceramide and sphingomyelin. This binding has been shown to mediate the attachment of HIV-1 to plasma membrane microdomains (rafts). Several works suggest that GSLs are involved in the entry of a broad range of HIV-1 isolates into cell lines

Table 1. Novel VpU signatures significantly associated with specific $gp120_{V3}$ amino acid changes

VpU signatures	Frequency no. (%)of isolates[a]	Frequency % in X4-tropic viruses[b]	Correlated mutations	Frequency no. (%) of isolates[a]	Covariation frequency no. (%) of isolates[c]	φ[d]	P[e]
0-1insV	6 (2.5)	100	$I30M_{V3}$	8 (3.3)	4 (66.7)	0.56	8.11E-6
			$60\text{-}62del_{VpU}$	9 (3.8)	4 (66.7)	0.53	1.45E-5
			$I46L_{VpU}$	14 (5.9)	4 (66.7)	0.42	1.11E-4
			$H13R_{V3}$	25 (10.5)	4 (66.7)	0.29	0.001
			$E25K/R/Q_{V3}$	59 (24.7)	5 (83.3)	0.22	0.004
			$T22A_{V3}$	141 (59.0)	1 (16.7)	−0.14	0.040
V10A	9 (3.8)	88.9	$E25I_{V3}$	8 (3.3)	2 (22.2)	0.21	0.032
			$T22A_{V3}$	141 (59.0)	2 (22.2)	−0.15	0.032
L12F	8 (3.8)	12.5	$E25A_{V3}$	16 (6.7)	3 (37.5)	0.23	0.012
			$T22A_{V3}$	141 (59.0)	8 (100)	0.15	0.023
D44E	17 (7.1)	23.5	$D80N_{VpU}$	5 (2.1)	2 (11.8)	0.19	0.043
I46L	14 (5.9)	78.6	$60\text{-}62del_{VpU}$	9 (3.8)	9 (64.3)	0.79	3.73E-13
			$I30M_{V3}$	8 (3.3)	7 (50.0)	0.65	3.61E-9
			$H13R_{V3}$	25 (10.5)	10 (71.4)	0.50	2.09E-8
			$E25K/R/Q_{V3}$	59 (24.7)	9 (64.3)	0.23	0.002
			$S11K/R_{V3}$	57 (23.8)	7 (50.0)	0.15	0.046
			$S11S_{V3}$	137 (57.3)	4 (28.6)	−0.15	0.026
			$T22A_{V3}$	141 (59.0)	3 (21.4)	−0.20	0.004
60-62del	9 (3.8)	100	$I30M_{V3}$	8 (3.3)	7 (77.8)	0.82	3.86E-11
			$H13R_{V3}$	25 (10.5)	9 (100)	0.58	3.81E-10
			$E25K/R/Q_{V3}$	59 (24.7)	8 (88.9)	0.29	7.55E-5
			$S11K/R_{V3}$	57 (23.8)	6 (66.7)	0.20	0.007
			$T22A_{V3}$	141 (59.0)	0 (0)	−0.24	2.19E-4
Q61H	6 (2.5)	0	$H13N/P_{V3}$	61 (25.5)	4 (66.7)	0.15	0.040
			$S11S_{V3}$	137 (57.3)	6 (100)	0.14	0.041
D80N	5 (2.1)	0	$R18Q_{V3}$	11 (4.6)	2 (40.0)	0.25	0.018

[a] Frequency was determined in 239 isolates from HIV-1 infected patients; [b] Frequency was determined in 120 HIV-1 isolates reported as X4-tropic at phenotypic test; [c] Percentages were calculated in patients with each specific VpU signature; [d] Positive and negative correlations with $\varphi > 0.10$ and $\varphi < -0.10$, respectively, are shown; [e] P values significant ($P \leqslant 0.05$) after correction for multiple hypothesis testing[2].

expressing CD4, CCR5 and/or CXCR4, and the GSL depletion blocked subsequent viral fusion and infection[44].

The VpU amino acid change L12F, localized in the first part of the TM domain, also established a positive correlation with two V3 mutations, E25A ($P = 0.012$; $\varphi = 0.23$) and T22A ($P = 0.015$; $\varphi = 0.15$) (Table 1). Of note, $T22A_{V3}$ was found in 100% of patients with $L12F_{VpU}$: in the covariation frequency these two mutations has shown to always have the codon *gca* (for the $T22A_{V3}$) and the codon *uuu* (for $L12F_{VpU}$), further underlining that these mutations are strongly correlated with each other both at the nucleotide and amino acid level.

The correlation of signatures in V3 and VpU was also confirmed by hierarchical clustering analysis[11,13,40]. In particular, the topology of the dendrogram suggests the existence of a three sub-clusters associated with R5-tropism composed of five V3 mutations and four VpU signatures: $E25D_{V3}$ and $V10A_{VpU}$, (bootstrap=0.71); $0\text{-}1insV_{VpU}$, $H13R_{V3}$, $I46L_{VpU}$, $I30M_{V3}$, and $60\text{-}62del_{VpU}$ (bootstrap = 0.88); $S11K/R_{V3}$ and $E25D_{V3}$ (bootstrap =0.64). These sub-clusters were part of a large R5-cluster (bootstrap = 0.97) (Fig.2). Similarly, another large cluster was found associated with X4-tropism (bootstrap =0.94). This involves three sub-clusters: $D44E_{VpU}$, $R18Q_{V3}$ and $D80N_{VpU}$, (bootstrap = 0.63); $E25A_{V3}$ and $L12F_{VpU}$, (bootstrap = 0.70); $Q61H_{VpU}$, $T22A_{V3}$, $S11S_{V3}$, $E25D_{V3}$, and $H13N/P_{V3}$ (bootstrap = 1) (Fig. 2).

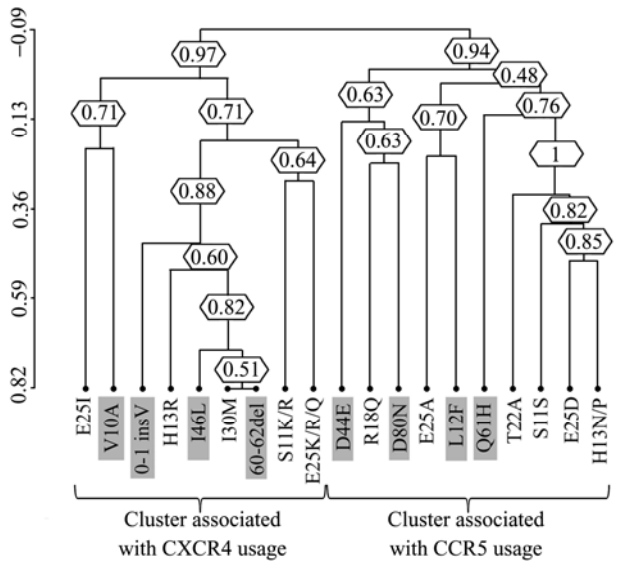

Fig.2. Clusters of correlated V3 and VpU signatures. Dendrogram obtained from average linkage hierarchical agglomerative clustering, showing significant clusters involving V3 and VpU (gray box) signatures. The length of branches reflects distances between mutations in the original distance matrix. Boostrap values, indicating the significance of clusters, are reported in the boxes. The analysis was performed in sequences derived from 239 patients, 119 reported as R5-tropic and 120 reported as X4-tropic at phenotypic test.

DISCUSSION

In addition to co-receptor usage, there may be many different reasons for the correlation among V3 and VpU amino acid changes. Possibly, these associations may have an impact on the HIV-1 pathogenesis. It is known that the CXCR4 phenotype has been associated with progression and increased severity of HIV-1 disease. Recently, Marchal and his co-workers highlighted a novel functional link between the VpU and caspase-dependent apoptosis *via* the activation of the c-Jun N-terminal Kinas pathway[28]. Moreover, VpU-deficient virions have been tightly associated with the cell surface, indicating a possibility that VpU can convert the mode of virus transmission[22]. Consistently, it has been also found that VpU indirectly enhances retroviral release through modification of the cellular environment (numerous reports have shown the high complexity of the relationships between VpU and cellular proteins of the host)[18]. Hence, some of the VpU signatures observed here could also be linked with the phenomena listed above.

In addition, Stephens and his co-workers have suggested that removal of the *VpU* sequence upstream of *Env* resulted in enhanced Env precursor synthesis, with a viral capacity of low pathogenic phenotype for pig-tailed macaques[38], while Casella *et al*. have observed the capacity of VpU to increase the viral susceptibility in infected cells to Fas killing[4]. Moreover, other studies have already indicated the importance of VpU for virus replication in macrophages[10, 25, 34]. These other details underline the complexity of the HIV-1 life cycle *via* viral factors and expression of additional proteins (regulatory and accessory) that can regulate its virulence and pathogenesis. In our analysis, specific variable residues were shown to be correlated with primary tropic amino acid signatures in the gp120-V3-loop that strengthens the case for a role of VpU that is more than a "simple" accessory protein.

For the sake of completeness, previous studies have shown that the VpU and Env proteins are translated from different reading frames from the same messenger RNA (mRNA) species[36]. Thus, the unique structure of these mRNAs suggests that the upstream *VpU* sequences prior to the E*nv* could condition the expression of the Env polyprotein. Occasionally, VpU tropic-properties have described in the past[10, 37], generally associated with intracellular location of this protein and consequently not with the virus entry steps. Nevertheless, the co-presence of *Env* and *VpU* ORFs in the same mRNAs (in late phases of viral infection)[24] should suggest a closer relationship between VpU and viral tropism. The selection in the gp120-V3 domain of mainly R5- and X4-tropic mutations could determine a secondary forced selection of mutations in the VpU sequence. Specifically, since HIV replication depends upon the stability of its RNA genome and its single spliced transcripts within the infected cell, the conservation of some positions could be a direct consequence of the need for the virus to maintain the proper folding of genomic RNA into highly stable RNA paired stems. Hence, after the discovery of VpU as a functional gene[7, 39], we now hypothesize that its gene product may have an important function not only in regulating virus release, but also has significant association with different co-receptor usage and specific V3 mutations.

In summary, this study shows that specific VpU substitutions are significantly associated with different co-receptor usage and with specific V3 mutations, both at the nucleotide and amino acid level. Specific

in vitro studies are needed to exclude an hypothetical statistical false positive result, to confirm that these VpU mutations contribute directly to co-receptor usage and to establish the specific and precise utility of this information.

References

1. Althaus C F, Vongrad V, Niederost B, *et al*. 2012. Tailored enrichment strategy detects low abundant small noncoding RNAs in HIV-1 infected cells. Retrovirology, 9: 27.
2. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B, 57: 289–300.
3. Binette J, Cohen E A. 2004. Recent advances in the understanding of HIV-1 Vpu accessory protein functions. Curr Drug Targets Immune Endocr Metabol Disord, 4: 297–307.
4. Casella C R, Rapaport E L, Finkel T H. 1999. Vpu increases susceptibility of human immunodeficiency virus type 1-infected cells to fas killing. J Virol, 73: 92–100.
5. Cashin K, Roche M, Sterjovski J, *et al*. 2011. Alternative coreceptor requirements for efficient. J Virol, 85: 10699–10709.
6. Chen M Y, Maldarelli F, Karczewski M K, *et al*. 1993. Human immunodeficiency virus type 1 Vpu protein induces degradation of CD4 *in vitro*: the cytoplasmic domain of CD4 contributes to Vpu sensitivity. J Virol, 67: 3877–3884.
7. Cohen E A, Terwilliger E F, Sodroski J G, *et al*. 1988. Identification of a protein encoded by the vpu gene of HIV-1. Nature, 334: 532–534.
8. D'Agostino D M, Felber B K, Harrison J E, *et al*. 1992. The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of gag/pol and vpu/env mRNAs. Mol Cell Biol, 12: 1375–1386.
9. de Jong J J, de Ronde A, Keulen W, *et al*. 1992. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. J Virol, 66: 6777–6780.
10. Dejucq N, Simmons G, Clapham P R. 2000. T-cell line adaptation of human immunodeficiency virus type 1 strain SF162: effects on envelope, vpu and macrophage-tropism. J Gen Virol, 81: 2899–2904.
11. Dimonte S, Mercurio F, Svicher V, *et al*. 2011. Selected amino acid mutations in HIV-1 B subtype gp41 are associated with specific $gp120_{V3}$ signatures in the regulation of co-receptor usage. Retrovirology, 8: 33.
12. Dimonte S, Mercurio F, Svicher V, *et al*. 2012. Genetic and Structural Analysis of HIV-1 Rev Responsive Element Related to V38A and T18A Enfuvirtide Resistance Mutations. Intervirology, 55: 385–390.
13. Dimonte S, Svicher V, Salpini R, *et al*. 2011. HIV-2 A-subtype $gp125_{C2\text{-}V3\text{-}C3}$ mutations and their association with CCR5 and CXCR4 tropism. Arch Virol, 156: 1943–1951.
14. Dube M, Bego M G, Paquay C, *et al*. 2010. Modulation of HIV-1-host interaction: role of the Vpu accessory protein. Retrovirology, 7: 114.
15. Fouchier R A, Groenink M, Kootstra N A, *et al*. 1992. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. J Virol, 66: 3183–3187.
16. Fujita K, Omura S, Silver J. 1997. Rapid degradation of CD4 in cells expressing human immunodeficiency virus type 1 Env and Vpu is blocked by proteasome inhibitors. J Gen Virol, 78 (Pt 3): 619–625.
17. Ghaffari G, Tuttle D L, Briggs D, *et al*. 2005. Complex determinants in human immunodeficiency virus type 1 envelope gp120 mediate CXCR4-dependent infection of macrophages. J Virol, 79: 13250–13261.
18. Gottlinger H G, Dorfman T, Cohen E A, *et al*. 1993. Vpu protein of human immunodeficiency virus type 1 enhances the release of capsids produced by gag gene constructs of widely divergent retroviruses. Proc Natl Acad Sci U S A, 90: 7381–7385.
19. Hout D R, Gomez L M, Pacyniak E, *et al*. 2006. A single amino acid substitution within the transmembrane domain of the human immunodeficiency virus type 1 Vpu protein renders simian-human immunodeficiency virus (SHIV(KU-1bMC33)) susceptible to rimantadine. Virology, 348: 449–461.
20. Huang C C, Lam S N, Acharya P, *et al*. 2007. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. Science, 317: 1930–1934.
21. Hussain A, Das S R, Tanwar C, *et al*. 2007. Oligomerization of the human immunodeficiency virus type 1 (HIV-1) Vpu protein - a genetic, biochemical and biophysical analysis. Virol J, 4: 81.
22. Iwatani Y, Song S K, Wang L, *et al*. 1997. Human immunodeficiency virus type 1 Vpu modifies viral cytopathic effect through augmented virus release. J Gen Virol, 78 (Pt 4): 841–846.
23. Jiang C, Parrish N F, Wilen C B, *et al*. 2011. Primary infection by a human immunodeficiency virus with atypical coreceptor tropism. J Virol, 85: 10669–10681.
24. Karn J, Stoltzfus C M. 2012. Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. Cold Spring Harb Perspect Med, 2: a006916.
25. Kawamura M, Ishizaki T, Ishimoto A, *et al*. 1994. Growth ability of human immunodeficiency virus type 1 auxiliary gene mutants in primary blood macrophage cultures. J Gen Virol, 75 (Pt 9): 2427–2431.
26. Krummheuer J, Johnson A T, Hauber I, *et al*. 2007. A minimal uORF within the HIV-1 vpu leader allows efficient translation initiation at the downstream env AUG. Virology, 363: 261–271.
27. Magadan J G, Bonifacino J S. 2012. Transmembrane domain determinants of CD4 Downregulation by HIV-1 Vpu. J Virol, 86: 757–772.
28. Marchal C, Vinatier G, Sanial M, *et al*. 2012. The HIV-1 Vpu Protein Induces Apoptosis in Drosophila via Activation of JNK Signaling. PLoS One, 7: e34310.
29. Margottin F, Benichou S, Durand H, *et al*. 1996. Interaction between the cytoplasmic domains of HIV-1 Vpu and CD4: role of Vpu residues involved in CD4 interaction and in vitro CD4 degradation. Virology, 223: 381–386.
30. Michalski C J, Li Y, Kang C Y. 2010. Induction of cytopathic effects and apoptosis in Spodoptera frugiperda cells by the HIV-1 Env glycoprotein signal peptide. Virus Genes, 41: 341–350.
31. Park S H, Opella S J. 2007. Conformational changes induced by a single amino acid substitution in the trans-membrane domain of Vpu: implications for HIV-1 susceptibility to channel blocking drugs. Protein Sci, 16: 2205–2215.

32. Peterlin B M, Trono D. 2003. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. Nat Rev Immunol, 3: 97–107.
33. Richards K H, Clapham P R. 2007. Effects of vpu start-codon mutations on human immunodeficiency virus type 1 replication in macrophages. J Gen Virol, 88: 2780–2792.
34. Schubert U, Clouse K A, Strebel K. 1995. Augmentation of virus secretion by the human immunodeficiency virus type 1 Vpu protein is cell type independent and occurs in cultured human primary macrophages and lymphocytes. J Virol, 69: 7699–7711.
35. Schubert U, Henklein P, Boldyreff B, *et al*. 1994. The human immunodeficiency virus type 1 encoded Vpu protein is phosphorylated by casein kinase-2 (CK-2) at positions Ser52 and Ser56 within a predicted alpha-helix-turn-alpha-helix-motif. J Mol Biol, 236: 16–25.
36. Schwartz S, Felber B K, Pavlakis G N. 1992. Mechanism of translation of monocistronic and multicistronic human immunodeficiency virus type 1 mRNAs. Mol Cell Biol, 12: 207–219.
37. Shibata R, Hoggan M D, Broscius C, *et al*. 1995. Isolation and characterization of a syncytium-inducing, macrophage/T-cell line-tropic human immunodeficiency virus type 1 isolate that readily infects chimpanzee cells in vitro and in vivo. J Virol, 69: 4453–4462.
38. Stephens E B, McCormick C, Pacyniak E, *et al*. 2002. Deletion of the vpu sequences prior to the env in a simian-human immunodeficiency virus results in enhanced Env precursor synthesis but is less pathogenic for pig-tailed macaques. Virology, 293: 252–261.
39. Strebel K, Klimkait T, Martin M A. 1988. A novel gene of HIV-1, vpu, and its 16-kilodalton product. Science, 241: 1221–1223.
40. Svicher V, Alteri C, D'Arrigo R, *et al*. 2009. Treatment with the fusion inhibitor enfuvirtide influences the appearance of mutations in the human immunodeficiency virus type 1 regulatory protein rev. Antimicrob Agents Chemother, 53: 2816–2823.
41. Svicher V, Cammarota R, Artese A, *et al*. 2010. New V3-genetic Signatures Modulate Co-receptor Usage *in vivo* and the Interaction with CCR5 N-terminus. [paper 542]. In: Program and abstracts of the 17th Conference on Retroviruses and Opportunistic Infections; February 16-19, 2010; San Francisco, CA (http://retroconference.org/2010/ Abstracts/38958.htm)
42. Taylor W R. 1986. The classification of amino acid conservation. J Theor Biol, 119: 205–218.
43. Verma S, Ali A, Arora S, *et al*. 2011. Inhibition of beta-TrcP-dependent ubiquitination of p53 by HIV-1 Vpu promotes p53-mediated apoptosis in human T cells. Blood, 117: 6600–6607.
44. Waheed A A, Freed E O. 2010. The Role of Lipids in Retrovirus Replication. Viruses, 2: 1146–1180.
45. Wilen C B, Tilton J C, Doms R W. 2012. Molecular mechanisms of HIV entry. Adv Exp Med Biol, 726: 223–242.