Virology 535 (2019) 266–271

Contents lists available at ScienceDirect

Virology

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

Variability OF HIV-1 V2 env domain for integrin binding: Clinical correlates

Claudia Fabrizio[a,b], Davide F. Bavaro[a], Luigia Scudeller[c], Luciana Lepore[a], Flavia Balena[a], Antonella Lagioia[a], Gioacchino Angarano[a], Laura Monno[a], Annalisa Saracino[a,*]

[a] Clinic of Infectious Diseases, University of Bari, University Hospital Policlinico, Piazza Giulio Cesare n. 11, 70124, Bari, Italy
[b] Malattie Infettive e Tropicali Ospedale Oncologico San Giuseppe Moscati, Taranto, Italy
[c] Scientific Direction, Clinical Epidemiology Unit, IRCCS San Matteo Foundation, Viale Camillo Golgi, 19, 27100, Pavia, Italy

ARTICLE INFO

Keywords:
HIV-1
Gp-120
V2 tripeptide
α4β7
Integrins

ABSTRACT

The HIV V2$^{179-181}$ (HXB2 numbering) tripeptide mediates binding to α4β7 integrin, which is responsible for GALT homing. Our study aimed to assess V2 variability in naive HIV-1 infected patients and its association with clinical and viro-immunological features.

Gp120 sequences were obtained from 322 subjects; length, potential N-linked glycosylation sites (PNGs), net-charge (NC) and $^{179-181}$tripeptide α4β7-binding-motif of V2 were evaluated.

At multivariate analysis, lower V2 length and higher NC correlated with low CD4 cells; no association was found with PNGs. A greater variability pertained positions 162–163, 164–167, 169, 175–179, 187, 194 and 195 in B sequences, and 163 and 177 in X4 tropic viruses.

LDV was the most common tripeptide. Asp^{180} was highly conserved; Leu^{179} was more frequently observed in non-B and in recent infections compared to others, while Val^{181} was found in recent infections and in MSM. Further studies to deeply explore the clinical significance of these associations are warranted.

1. Introduction

The HIV-1 V1V2 region of gp120 is one of the most variable within the whole HIV-1 sequence. Although this region is not essential for viral entry, its absence may be responsible for a higher viral susceptibility to antibody neutralization (Rao et al., 2013).

Whilst the exact role of V1 is still unclear, the V2 loop is known to be involved in several important functions, including the CD4 binding, the *env* trimer assembly, and the chemokine binding (Rao et al., 2013); it also impacts the co-receptorial viral tropism, based on the presence of some potential N-linked glycosylation sites (PNGs) and on the interaction with the V3 region (Monno et al., 2011). It has also been suggested that both V2 length and PNGs may influence the transmission fitness of founder HIV viruses (Cicala et al., 2011).

Moreover, V2 encompasses the 179–181 (according to HXB2 numbering) tripeptide which represents the binding site for the gut homing integrin α4β7: indeed, this interplay promotes the migration of infected lymphocytes towards the Gut-Associated Lymphoid Tissue (GALT) in the early phase of HIV-1 infection, and contributes to the creation of virologic synapse, which in turn facilitates the spreading of infection (Cicala et al., 2011). The role of this interaction in chronic phases of the infection and after the initiation of antiretroviral therapy is not yet studied.

The remarkable advances in the field of novel therapeutic strategies based on immunologic tools are drawing a great deal of attention on both V2 region and α4β7 integrin, as well as on their linkage. In fact, the outcomes of the RV144 vaccine trial suggested a correlation between the protective efficacy of the vaccine and the induction of antibodies directed against HIV-1 V2 region (including the α4β7 binding site) (Rerks-Ngarm et al., 2009). On the other hand, Byrareddy et al., (2016) demonstrated that the combined use of antiretroviral therapy and monoclonal antibodies anti-α4β7 allows to obtain viral control and to partially reconstitute the immune system functionality in non-human primates, also after the withdrawal of the whole therapeutic regimen. These evidences are even more noteworthy in the light of the availability of an anti-α4β7 monoclonal antibody, vedolizumab, currently approved mainly for treatment of inflammatory bowel diseases (Entyvio, Takeda Pharm America, 2014). In addition, V2 region also encompasses other positions (namely, codons 160, 166, 169) which have been shown to represent key sites of immune pressure (Rao et al., 2013).

These premises enlighten the multiple, potential benefits of an improved knowledge about the features of V2 region, particularly regarding the α4β7 binding site. Recently, a peculiar affinity has been

* Corresponding author.
E-mail address: annalisa.saracino@uniba.it (A. Saracino).

https://doi.org/10.1016/j.virol.2019.07.019
Received 20 May 2019; Received in revised form 18 July 2019; Accepted 23 July 2019
Available online 24 July 2019

described between viruses encoding specific motifs at positions 179–181 within V2 region and the expression of integrin α4β7 (Sivro et al., 2018).

Furthermore, Hait et al. (2015) hypothesized an intriguing scenario in which HIV-1 strains containing specific 179–181 tripeptides are spreading worldwide, in line with the increasing diffusion of certain viral subtypes, thus speculating that they may provide some epidemiological advantage.

Therefore, our study aimed to assess the variability of V2 region in naive HIV-1 infected subjects and the potential association of its characteristics with clinical and viro-immunological features of the study population.

2. Materials and methods

2.1. Samples

We retrospectively evaluated 322 gp120 sequences from as many subjects newly diagnosed with HIV-1 infection between 2009 and 2017. Sequences were obtained from plasma samples coincident with HIV diagnosis.

2.2. Gp120 sequencing

For gp120 sequencing, a nested RT-PCR protocol was adopted to amplify a 1320-bp region encompassing almost the entire gp120 (codon 32 of C1 to V5) (Saracino et al., 2009; Monno et al., 2011). Only one PCR product per sample was subjected to standard population sequencing. Sequences were analyzed with Seqscape software v2.5 (Applied Biosystems, Foster City, CA). Nucleotide mixtures were considered if the second highest peak in the electropherogram was > 25%. Ambiguous codes were solved for all possible permutations, and sequences were submitted to the HIV-1 sequence quality-control (QC) tool (http://www.hiv.lanl.gov/content/sequence/QC/index.html), resulting in high-quality sequences. The QC tool also furnishes results from several other Web-based tools, including the Gene Cutter tool (http://www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html), which clips predefined coding regions and generates nucleotide and protein alignments of the cut regions using HXB2 (K03455) as the reference sequence.

After extracting hypervariable loops from the gp120 sequences, V2 regions were retrospectively evaluated for length, potential N-linked glycosylation sites (PNGs), net charge (NC). V2 length an NC were calculated by online free software Innovagen Peptide Property calculator (https://pepcalc.com/); PNGs were evaluated by N-Glycosites predictors online software (https://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html) produced by Los Alamos National Laboratory.

A detailed analysis of the tripeptide motif at positions 179–181 was also performed, as well as of codons 160, 166, 169. V2 variability was evaluated by means of the online entropy tool (https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html); this tool allows to quantify amino acid variability at individual positions, considering the number of possible amino acids replaced and their frequency.

The extracted V3 sequences were submitted to the geno2pheno$_{[\text{CORECEPTOR}]}$ tool to obtain a prediction of coreceptor usage with a false positive rate (FPR) of 10%. Sequences with FPR < 10% and > 10% were classified as obtained from either X4 or R5 viruses, respectively.

Gp120 sequences used in this study included those published previously (accession numbers: JF896816 - JF896819; JF896824 - JF896851; JF896853; JF896855 - JF896865; JF896869 - JF896874), and new submissions (MK732761- MK732853).

2.3. Subtype assignment

For HIV subtype assignment sequences derived from standard genotyping of HIV-1 protease and reverse transcriptase performed according to current guidelines (EACS Guidelines, 2018) were run on the REGA HIV-1 Subtyping Tool - Version 3.0 [http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/]. For the purpose of this study, sequences were classified as derived from either B or non-B HIV strains.

2.4. Study population features and clinical correlates

Baseline clinical and viro-immunological characteristics of the study population were retrieved from our Clinic database. Acute infection was defined based on detectable plasma HIV-1 RNA in the setting of a negative/indeterminate HIV-antibody test. For patients with unknown data of seroconversion to HIV positivity, duration of HIV infection was estimated according to the frequency of ambiguous nucleotides in RT/PR: e.g., a proportion of ambiguity ≤0.2% signifies a recent (less than 1 year) infection (Zheng et al., 2013; Kouyos et al., 2011) A possible association was sought for between the main features of V2 region and:

- viral features (HIV subtype; R5/X4 co-receptor tropism; duration of HIV infection).
- patients' variables (sex; age; risk factor for HIV-1 transmission; stage at diagnosis; baseline viro-immunological data, including HIV-RNA, CD4 and CD8 cell count and percentage, CD4/CD8 ratio).

2.5. Statistical analysis

Descriptive statistics were produced for demographic, clinical and laboratory characteristics of cases. Mean and standard deviation (SD) were obtained for normally distributed variables, and median and interquartile range (IQR) for non-normally distributed variables, number and percentages for categorical variables. Groups were compared with parametric or nonparametric tests (Kruskal Wallis or Mann Whitney Test where appropriate), according to data distribution, for continuous variables, and with Pearson's $\chi 2$ test (Fisher exact test where appropriate) for categorical variables. To assess the association between amino acid positions and laboratory and clinical parameters, univariate and multivariate linear regression models were applied; 2-tailed tests were used; a P-value < 0.05 was considered statistically significant. For the evaluation of variability at each amino acid position by the ENTROPY-TWO tool, a p value < 0.005 was considered significant.

2.6. Ethics

The research did not require a formal approval from the ethics committee according to the Italian law since it was performed as an observational retrospective study in the context of normal clinical routines (art.1, leg. decree 211/2003). However, the study was conducted in accordance with the Declaration of Helsinki and national and institutional standards. All patients provided informed consent for the use of their data for research purposes. In any case, data were previously anonymized, according to the requirements set by Italian Data protection Code (leg. Decree 196/2003).

3. Results

3.1. Features of the study population

A total of 322 patients were included, mostly males (274 subjects, 85.1%), with a median age of 35.5 years (IQR 27.5–44.4), 53% men who had sex with men (MSM), all naive to antiretroviral therapy. A recent HIV-1 infection (< 1 year) was estimated in 148 subjects (46%), including 19 (5.9%) acute infections; 47 patients (14.6%) were AIDS-

Table 1
Features of the study population.

Features	Total	B subtype	Non–b subtype	P value
N of pts (%)	322	214 (66)	108 (34)	
Males, n (%)	274 (85)	189 (88)	85 (79)	**0.03**
Age, median (IQR)	35.5 (27.5–44.5)	38 (29–46)	29 (25–40)	< **0.001**
AIDS, n (%)	47 (14.9)	35 (16.5)	12 (11.5)	0.31
Risk factor, n (%)				
Heterosexual	123 (38)	74 (34)	49 (45)	
MSM	171 (53)	116 (55)	55 (51)	
Unknown/other	28 (9)	24 (11)	4 (4)	**0.03**
R5 coreceptor tropism, n (%)	265 (82)	165 (77)	100 (93)	< **0.001**
CD4 n, median (IQR)	371 (172–525)	353 (161–525)	411 (213–576)	0.27
CD4%, median (IQR)	21 (12–28)	20 (11–28)	21 (13–28)	0.69
HIV-RNA (cps/ml) median (IQR)	59,000 (12,000–210,000)	78,000 (17,000–280,000)	29,000 (4700–120,000)	< **0.001**
V2 Length (median aa, IQR)	41 (39–45)	41 (39–43)	43 (40–48)	< **0.001**
PNGs (median, IQR)	2.12 (2.12–3)	2.12 (2.12–3.5)	3.5 (2–3.5)	< **0.001**
Net Charge (median, IQR)	0.9 (−0.1–1.9)	0.1 (- 0.1–1)	1 (0–2)	< **0.001**

Legend: N = number; IQR = interquartile range; aa = amino acids; MSM = men who have sex with men; PNGs = potential N-linked glycosylation sites.

presenters. The majority of sequences were from patients infected with subtype B HIV strains (214, 66.5%) and R5 clinical isolates were the most common variants (265, 82.3%) (Table 1).

3.2. Characteristics of V2 region

3.2.1. Analysis of length, net charge and PGNs

The median V2 lenght was 41 (39–45) amino acids; the median net charge (NC) was 0.9 (−0.1 – 1.9); and the median number of potential N-linked glycosylation sites (PNGs) was 2 (2–3). A significant diversity was observed between B and non-B sequences in length and number of PNGs. (Table 1).

Univariate and multivariate analysis (Table 2) were performed to evaluate possible correlations between these V2 features and clinical and viro-immunological variables.

A lower V2 length correlated with CD4 cell count < 200 cell/mmc ($p = 0.001$), X4 tropism ($p = 0.016$) and AIDS condition ($p = 0.038$) at univariate analysis; however, only the correlation with CD4 cell count was maintained in multivariate analysis ($p = 0.01$).

Conversely, a direct association was found between higher V2 net charge and either an AIDS diagnosis ($p = 0.016$) and a low CD4 cell count (< 200 cells/μL) ($p = 0.03$); again, the latter association was also confirmed at multivariate analysis ($p = 0.045$).

Finally, univariate analysis indicated that a lower number of PNGs correlated with a CD4 cell count < 200/μL ($p = 0.017$) and AIDS stage at diagnosis ($p = 0.028$), however none of these associations was proved in multivariate analysis.

A PNG at position 160–162 (N160) was present in 88.8% and 92.6% of B and non-B sequences ($p = 0.28$) respectively, whereas N186 and N188 were seen in 13.5% vs 30.5% (33/108) ($p < 0.001$) and 47.2% vs 12.9% (14/108) ($p < 0.001$) of B vs non-B sequences.

3.2.2. Analysis of amino acid variability

Variability at each amino acid position was examined by using the online entropy tool. The ENTROPY-TWO tool, which permits a graphic demonstration of variability, compares two sets of aligned sequences to determine if there is greater variability in one set relative to the other. The alignments of B and non-B, and of R5 and X4 sequence groups were compared.

An unexpected greater sequence variability was seen in B sequences compared to sequences derived from non-B variants (Fig. 1a). A statistically significant ($p < 0.005$) greater variability in favour of B sequences pertained positions 162 and 163 (HXB2 numbering), 164–167, 169, 175–179, 187, 194 and 195. At these positions, amino acid substitutions in B sequences with respect to the HXB2 strain were S162T (88.2%), G167D (72.2%), F175L (79.7%), D187 N (31.1%) and T194I (75%). Notwithstanding the high variability at positions 166 and 169, an Arg^{166} and a Val^{169} were conserved in 49.1% and 25.5% of B sequences, respectively. S162T, G167D, F175L, D187 N and T194I were also detected respectively in 97.2%, 96.3%, 94.4%, 38% and 89.8% of non-B sequences; these latter also carried an Arg^{166} in 87% of cases, whereas Lys^{169} replaced in 54.6% of non-B sequences.

A significant entropy difference in favour of non-B sequences was referred to positions 173 and 190 alone where a Tyr^{173} and a Ser^{190} in HXB2 strain were replaced with a His^{170} and a Glu^{190} in 47.2% and 45.4% of sequences, respectively. On the contrary, an Asn^{160} was conserved in both B and non-B strains (92.9% vs 90.2%, $p = 0.47$), and in R5 and X4 variants (92.4% vs 87.5%; $p = 0.24$).

A statistically significant ($p < 0.005$) greater variability in favour of X4 sequences was observed at positions 163 and 177 (Fig. 1b), nevertheless a Thr^{163} and a Tyr^{177} were commonly detected in both R5 and X4 sequences. Even though Arg^{166} was retained in consensus sequences of both R5 and X4 strains, univariate and multivariate analysis of codon 166 demonstrated that Arg^{166} was related with R5-tropic strains ($p = 0.041$ and 0.032, respectively). A Lys^{169} was the most frequent amino acid in R5 sequences (35.5%); whereas an Ile^{169} was detected in 21.4% of X4 sequences.

Regarding the α4β7 binding region spanning positions 179–181,

Table 2
Correlations between V2 features and clinical-immunovirological variables.

V2 Features		Length			Net Charge			PNGs	
Clinical charateristics	Coefficient	Standard error	p-value	Coefficient	Standard error	p-value	Coefficient	Standard error	p-value
Estimate date of infection (< 1 year)	0.069	0.458	0.879	−0,003	0081	0,973	0011	0,105	0.918
< 200 $CD4^+$ cell count	−1.477	0.575	**0.010**	0,205	0102	**0.045**	−0,214	0131	0.104
X4 coreceptor tropism	−1027	0,542	0.059	−0,175	0096	0.071	0,038	0124	0.760
Homosexual risk factor	0,303	0428	0.479	−0,006	0076	0.932	0,030	0098	0.760
AIDS stage	−0,054	0718	0.940	0,117	0128	0.360	−0,123	0164	0.454

Legend: PNGs = potential N-linked glycosylation sites.

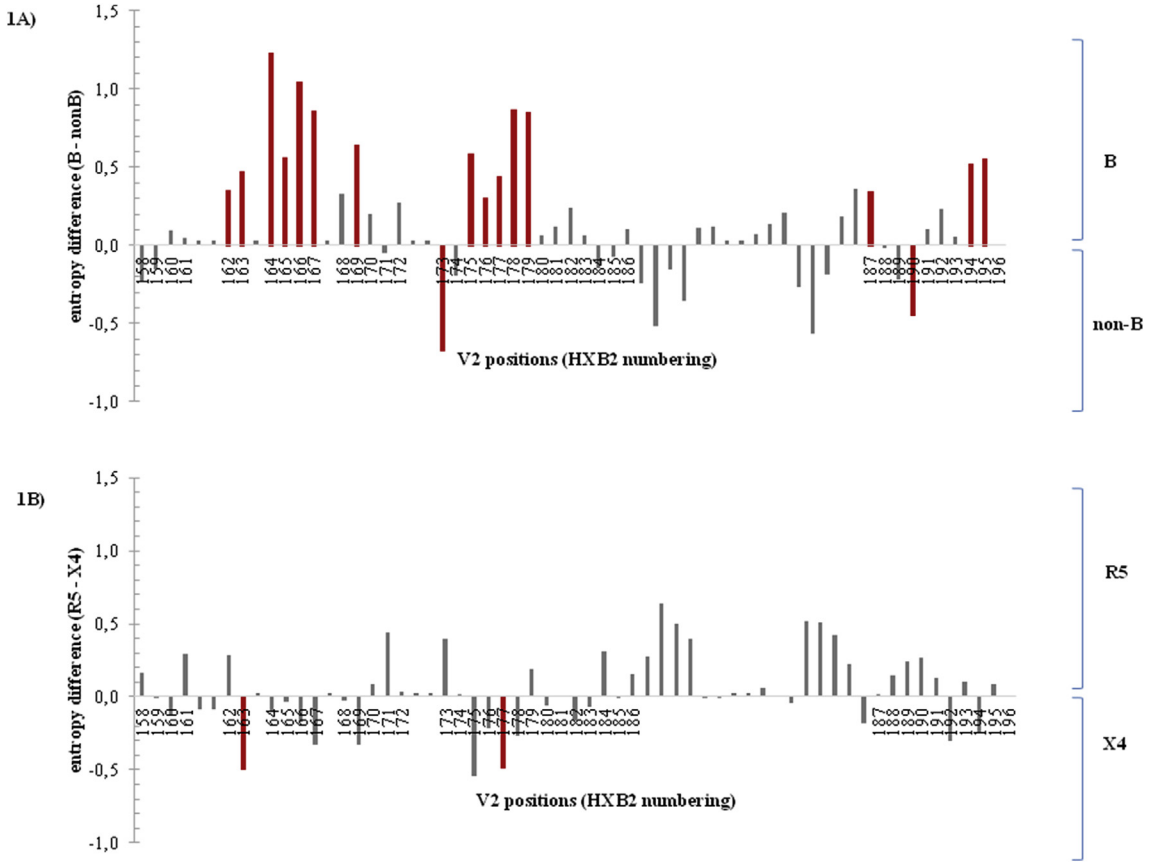

Fig. 1. Variability of amino acid sequence of V2 region according to viral subtype (B vs non-B, Fig. 1A) and predicted co-receptor tropism (R5 vs X4, Fig. 1B).
Legend: Variability at each amino acid position is examined by using the online ENTROPY-TWO tool which compares two sets of aligned sequences. The length of bars quantifies the variability difference between the two groups; the direction of bars indicates the population in which the greater variability is observed. Bars are red when the difference is statistically significant ($p < 0.005$).

LDV was the most common amino acid triplet (41.0%), followed by LDI (24.5%). Position 180 was highly conserved with an aspartic acid (D) retained by 99.3% of B and non-B sequences, and only two B sequences bearing either a Tyr (R5 variant) or an His (X4 variant).

A leucin^{179} (Leu^{179}) was found in 60.7% and 81.4% of B and non-B sequences ($p < 0.001$), and in 67.2% and 70.2% of R5 and X4 sequences ($p = 0.66$) respectively. Subjects infected with strains having a Leu^{179} were more likely to have a recent infection (76% vs 24%; $p = 0.004$); however, this association was not confirmed at multivariate analysis and no other correlation was found between clinical-virological data and variants with or without a Leu^{179}.

The reference HXB2 strain of subtype B HIV-1 also bears an isoleucin^{181} (Ile^{181}); accordingly, 71.4% of all sequences carrying an Ile^{181} pertained to B variants versus 28.6% of non-B sequences ($p = 0.13$), whereas 21.4% vs 78.5% were derived from X4 and R5 strains, respectively ($p = 0.16$). At the same codon, 62.8% and 37.2% of sequences with a Valin (Val) were derived from B and non-B strains, respectively ($p = 0.11$).

Noteworthy, strains with Val^{181} were more likely to be derived from MSM (67% vs 33%, $p < 0.001$) and with an earlier stage at diagnosis ($p = 0.004$); in comparison to subjects viral strains carrying different variants. Multivariate analysis confirmed the correlation between a Val^{181} with MSM risk factor ($p < 0.001$). Fig. 2 shows entropy differences between sequences from MSM and those derived from patients with risk factor for HIV infection other than homosexuality. A Val^{169} was replaced with a Lys^{169} in 27.7% of sequences from MSM and in 38.3% of those obtained from subjects with other risk factors for HIV infection ($p = 0.03$).

4. Discussion

The HIV-1 gp120 V2 region is drawing a great deal of attention due to the increasing evidence of its role in immune response demonstrated in the course of vaccine trials and analysis of interaction with bNAbs and M-Abs. Several aspects are currently being investigated, ranging from the role of the V1V2 conformation (Wibmer et al., 2018), to the development of anti-V2 antibodies (Van Eeden et al., 2018), and to the interaction between V2 region and the α4β7 integrin (Lertjuthaporn et al., 2018). This field is acquiring an even greater importance in view of the controversial results of studies evaluating the effectiveness of molecules interacting with V2 region, and in particular with the $^{179-181}$tripeptide, for the control of HIV infection. The V2$^{179-181}$tripeptide resembles the binding of α4β7 with its natural ligand MAdCAM (Lertjuthaporn et al., 2018). Therefore, antibodies hindering the α4β7-MAdCAM binding (such as vedolizumab) might as well compromise the interaction between α4β7 integrin and HIV-1 V2 (Byrareddy et al., (2016)). Indeed, the observations of Byrareddy et al., (2016) have been questioned by another study on rhesus macaques which did not validate the efficacy of administration of cART plus anti-α4β7 antibodies in controlling SIV replication after treatment interruption (Di Mascio et al., 2018). Similarly, Ling and collaborators showed that vedolizumab was unable to prevent or to control HIV-1 infection in vitro and in humanized mice (Ling et al., 2019). In light of this background, our purpose was to describe the features of HIV-1 gp120 V2 region in sequences collected from clinical isolates from 2009 to 2017 and to identify potential correlations with clinical variables.

Based on available literature (Cicala et al., 2011), the interaction between the V2 region and the gut homing integrin α4β7 is particularly

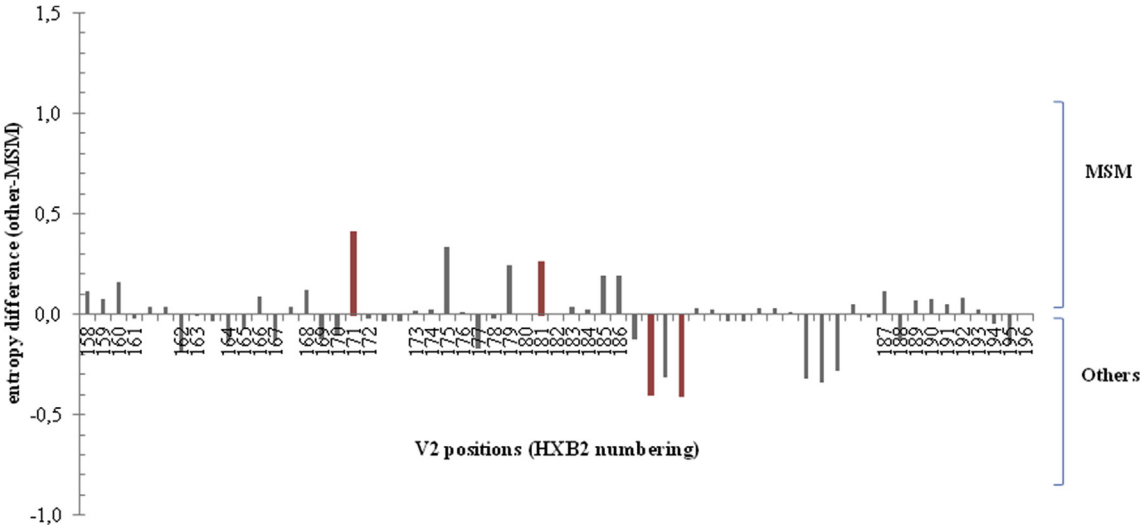

Fig. 2. Variability of amino acid sequence of V2 region according to way of transmission (MSM vs others).
Legend: Variability at each amino acid position is examined by using the online ENTROPY-TWO tool which compares two sets of aligned sequences. The length of bars quantifies the variability difference between the two groups; the direction of bars indicates the population in which the greater variability is observed. Bars are red when the difference is statistically significant ($p < 0.005$). MSM: men who have sex with men.

important in the early phases of HIV-1 infection, when it promotes the migration of infected lymphocytes towards the GALT thus contributing to its spreading. Moreover, according to our knowledge, no data are available regarding the impact of antiretroviral therapy on this binding. For these reasons, we preferred to include in the analysis antiretroviral naïve patients, and to analyse data also according the timing of HIV infection, in order to understand if some associations with viro-immunological and clinical features could be evidenced in the early phase (acute/recent) rather than in the chronic phase of the infection. No differences, however, were observed in V2 characteristics according to the duration of infection in our analysis.

As expected, the V2 loop LDI/V mimotope was the most frequent tripeptide among both B and non-B subtypes in our sample, with a higher diversity among B strains. This result agrees with those of a study of Hait et al., (2015), evaluating the features of the 179–181 triplet among sequences from South-American subjects plus sequences retrieved from the Los Alamos Database. According to this study, LDV motif was more frequent among B strains than LDI, as it was in our sample. Moreover, Hait et al. (2015) also showed that strains containing the LDI motif are spreading worldwide, thus hypothesizing a role of this tripeptide in providing some epidemiological advantage. Based on these data, we could speculate that the steady conservation of the LDI/V motif among non-B subtypes may unravel a more efficient circulation of these strains, in line with the current wider spreading of non-B viruses. On the other hand, however, studies regarding the RV144 trial showed that amino acids different from Ile^{181} were related with a higher efficiency in escaping the immune response induced by the vaccine (Rolland et al., 2012).

It has been shown recently that viruses carrying a proline or a serine at position 179 were particularly frequent in subjects presenting a high amount of CD4 cells expressing high levels of α4β7 (Sivro et al., 2018). This may suggest a more efficient interaction between certain HIV-1 strains and α4β7 integrin, resulting in a favoured infection by viruses presenting specific 179–181 motifs, in subjects with a high expression of α4β7 integrin on target cells. In our population, however, multivariate analysis of potential clinical correlates with amino acid at position 179 did not show any significant association and the majority of strains had a leucine at this codon.

Over a half of our sequences overall carried a valine at position 181, with a higher conservation among non-B subtypes. Moreover, multivariate analysis demonstrated a significant correlation of Val^{181} and also of Val^{169} with MSM route of transmission. It is unclear if this correlation may represent an example of "viral signature" related to a specific route of transmission (Hait et al., 2015): to this regard, it could be helpful to analyse viral sequences obtained from the GALT in MSM who acquired HIV infection.

Viruses encoding shortened V1/V2 with fewer N-linked glycosylation sites are associated with increased transmission fitness (Cicala et al., 2011; Nawaz et al., 2011); however, it is unclear if it is true for all viral subtypes and in all ways of transmission. No evidence of a relation between V2 net charge and transmission fitness was found in previous studies (Ashokkumar et al, 2017). In our population, a lower V2 length and a higher NC correlated with severe immune impairment (CD4 cell count < 200 cells/μl. This result is quite unexpected and should be further investigated in clinical studies. In fact, it appears to contradict previous observations suggesting that an increase in V2 lenght could be associated with disease progression (Kemal et al., 2008). On the contrary, no association was observed in our study between patient characteristics and number of PNGs within the V2 region. Removal of some PNGs especially when combined, may reduce viral infectivity, and also binding to neutralizing and non-neutralizing antibodies (Rathore et al., 2017). To this regard, one of the most important PNG is N160 (Rao et al., 2013), which was present in 89% and 93% of subtype B and non-B sequences in our series. This proportion was higher than in a previous report (Wang et al., 2013) systematically describing all HIV-1 Env PNGs, where N160 was detected in 75% of subtype B strains. Although the level of preservation of N160 in our study population was relatively high, the not negligible proportion of strains that may be resistant to the action of neutralizing antibodies should be taken into account when evaluating potential targets for MAb-based therapies; in fact, N160 removal abolishes the activity of neutralizing antibodies. Noteworthy, in our series the proportion of subtype B strains lacking N160 was higher than that of non-B subtypes. This finding is probably coupled with the greater variability of V2 sequences (in terms of entropy coefficient) observed in our series among subtype B strains compared with non-B variants. Whether this finding would signify a lower susceptibility of B subtype to neutralizing antibodies (Bouvin-Pley et al., 2013) remains to be confirmed.

Moreover, position 169 held the highest variability within the entire region in our series, expecially in B subtypes. As codon 169 represents a signature position for immune response based on several studies (including results from the RV144 trial (Rao et al., 2013; Rolland et al.,

2012; Lertjuthaporn et al., 2018), such observed diversity among clinical strains should be taken into account as a challenging factor.

It has also been reported that sequence variation in the V1/V2 loop might play an important role in viral tropism (Stamatatos et al., 1998; Ogert et al., 2001); however, in our analysis, neither V2 length or NC or PNGs appeared to independently influence co-receptor tropism, even if it should be acknowledged that it was predicted only based on genotypic methods analyzing the V3-loop sequence, and not actually measured by phenotypic assays. However, we found that Arg^{166} was predictive for R5 strains, in contrast with previous observations (Thielen et al., 2010; Monno et al., 2011). A possible explanation for this result can rely on the different viral characteristics of our population including subjects bearing both B and non-B strains, whereas the above mentioned studies were referred to subtype B isolates; in fact, in our study series, the majority of strains with an Arg^{166} were non-B strains.

5. Conclusions

Our results show a certain variability in V2 structure, including the α4β7 binding tripeptide, described as highly conserved, particularly in B subtypes, which could partially explain controversial results deriving from recent studies evaluating the effectiveness of molecules interacting with V2 region for the control of HIV infection. Moreover, the possible association of V2 features with some viro-immunological characteristics would require further studies.

Declaration of interest

Authors do not have any conflict of interest to declare.

Funding

This work was partially supported by un unrestricted grant of Gilead Sciences Srl, Italy, n.02810 (Fellowship Program 2017).

Acknowledgements

This study was partially presented at 25th Conference on Retroviruses and Opportunistic Infections (Boston, USA, March 4–7 2018; poster abstract n. 184).

References

Ashokkumar, M., Nesakumar, M., Cheedarla, N., et al., 2017 Aug. Molecular characteristics of the envelope of vertically transmitted HIV-1 strains from infants with HIV infection. AIDS Res. Hum. Retrovir. 33 (8), 796–806. https://doi.org/10.1089/AID.2016.0260.

Bouvin-Pley, M., Morgand, M., Moreau, A., et al., 2013. Evidence for a continuous drift of the HIV-1 species towards higher resistance to neutralizing antibodies over the course of the epidemic. PLoS Pathog. 9, e1003477.

Byrareddy, S.N., Arthos, J., Cicala, C., et al., 2016 Oct 14. Sustained virologic control in SIV+ macaques after antiretroviral and α4β7 antibody therapy. Science 354 (6309), 197–202.

Cicala, C., Arthos, J., Fauci, A.S., 2011. HIV-1 envelope, integrins and co-receptor use in mucosal transmission of HIV. J. Transl. Med. 9 (1), S2.

Di Mascio, M., Lifson, Srinivasula, S., et al., 2018. Evaluation of an antibody to Alpha4Beta7 in the control of SIV infection. Abstract TUAA0206LB. 22nd international AIDS conference abstract supplement. J. Int. AIDS Soc. 21 (S6), e25148.

Entyvio [prescribing Information]. Takeda Pharmaceuticals America, Inc, Deerfield, IL. https://general.takedapharm.com/ENTYVIOPI, Accessed date: 2 January 2018.

European AIDS Clinical Society, October 2018. Guidelines version 9.1. Available at: http://www.eacsociety.org/files/2018_guidelines-9.1-english.pdf.

Hait, S.H., Soares, E.A., Sprinz, E., et al., 2015. Worldwide genetic features of HIV-1 Env α4β7 binding motif: the Local Dissemination Impact of the LDI tripeptide. J. Acquir. Immune Defic. Syndr. 70 (5), 463 1999.

Kemal, K.S., Beattie, T., Dong, T., et al., 2008 Jun 1. Transition from long-term non-progression to HIV-1 disease associated with escape from cellular immune control. J. Acquir. Immune Defic. Syndr. 48 (2), 119–126. https://doi.org/10.1097/QAI.0b013e31816b6abd.

Kouyos, R., von Wyl, V., Yerly, S., et al., 2011. Ambiguous nucleotide calls from population based sequencing of HIV-1 are a marker for viral diversity and the age of infection. Clin. Infect. Dis. 52 (4), 532–539.

Lertjuthaporn, S., Cicala, C., Van Ryk, D., et al., 2018. Select gp120 V2 domain specific antibodies derived from HIV and SIV infection and vaccination inhibit gp120 binding to α4β7. PLoS Pathog. 14 (8), e1007278.

Ling, L., Wu, T., To, K.K.W., et al., 2019. Vedolizumab-mediated integrin α4β7 blockade does not control HIV-1SF162 rebound after combination antiretroviral therapy interruption in humanized mice. AIDS 33 (4), F1–F12.

Monno, L., Saracino, A., Scudeller, L., et al., 2011. Impact of mutations outside the V3 region on coreceptor tropism phenotypically assessed in patients infected with HIV-1 subtype B. Antimicrob. Agents Chemother. 55 (11), 5078–5084.

Nawaz, F., Cicala, C., Van Ryk, D., et al., 2011. The genotype of early-transmitting HIV gp120s promotes α (4) β(7)-reactivity, revealing α (4) β(7) +/CD4+ T cells as key targets in mucosal transmission. PLoS Pathog. 7 (2), e1001301. https://doi.org/10.1371/journal.ppat.1001301.

Ogert, R.A., Lee, M.K., Ross, W., et al., 2001 Jul. N-linked glycosylation sites adjacent to and within the V1/V2 and the V3 loops of dualtropic human immunodeficiency virus type 1 isolate DH12 gp120 affect coreceptor usage and cellular tropism. J. Virol. 75 (13), 5998–6006.

Rao, M., Peachman, K., Kim, J., et al., 2013. HIV-1 variable loop 2 and its importance in HIV-1 infection and vaccine development. Curr. HIV Res. 11 (5), 427–438.

Rathore, U., Saha, P., Kesavardhana, S., et al., 2017. Glycosylation of the core of the HIV-1 envelope subunit protein gp120 is not required for native trimer formation or viral infectivity. J. Biol. Chem. 292 (24), 10197–10219.

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., et al., Dec 3; 2009. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N. Engl. J. Med. 361 (23), 2209–2220.

Rolland, M., Edlefsen, P.T., Larsen, B.B., et al., 2012. Increased HIV-1 vaccine efficacy against viruses with genetic signatures in Env V2. Nature 490 (7420), 417.

Saracino, A., Monno, L., Cibelli, D.C., et al., 2009. Co-receptor switch during HAART is independent of virological success. J. Med. Virol. 81 (12), 2036–2044.

Sivro, A., Schuetz, A., Sheward, D., et al., 2018. Integrin α4β7 expression on peripheral blood CD4+ T cells predicts HIV acquisition and disease progression outcomes. Sci. Transl. Med. 10 (425) eaam6354.

Stamatatos, L., Wiskerchen, M., Cheng-Mayer, C., 1998. Effect of major deletions in the V1 and V2 loops of a macrophage-tropic HIV type 1 isolate on viral envelope structure, cell entry, and replication. AIDS Res. Hum. Retrovir. 14 (13), 1129–1139 85.

Thielen, A., Sichtig, N., Kaiser, R., et al., 2010. Improved prediction of HIV-1 coreceptor usage with sequence information from the second hypervariable loop of gp120. J. Infect. Dis. 202, 1435–1443.

Van Eeden, C., Wibmer, C.K., Scheepers, C., et al., 2018. V2-Directed vaccine-like antibodies from HIV-1 infection identify an additional K169-binding light chain motif with broad ADCC activity. Cell Rep. 25 (11), 3123–3135.

Vedolizumab- Summary of product characteristics. Available at: https://www.ema.europa.eu/en/documents/product-information/entyvio-epar-product-information_en.pdf..

Wang, W., Nie, J., Prochnow, C., et al., 2013. A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. Retrovirology 10 (1), 14.

Wibmer, C.K., Richardson, S.I., Yolitz, J., et al., 2018. Common helical V1V2 conformations of HIV-1 Envelope expose the α4β7 binding site on intact virions. Nat. Commun. 9 (1), 4489.

Zheng, D.P., Rodrigues, M., Bile, E., al, al, 2013 Oct 17. Molecular characterization of ambiguous mutations in HIV-1 polymerase gene: implication for monitoring HIV infection status and drug resistance. PLoS One 8 (10).