



## RESEARCH ARTICLE

# The V1 region of gp120 is preferentially selected during SIV/HIV transmission and is indispensable for envelope function and virus infection

Yanpeng Li<sup>1</sup>, Ulf Dittmer<sup>2</sup>, Yan Wang<sup>1</sup>, Jiping Song<sup>1</sup>, Binlian Sun<sup>1✉</sup>, Rongge Yang<sup>1✉</sup>

1. Research Group of HIV Molecular Epidemiology and Virology, Center for Emerging Infectious Diseases, State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China

2. Institute for Virology, University Hospital Essen, University of Duisburg-Essen, 45122 Essen, Germany

**A transmission bottleneck occurs during each human immunodeficiency virus (HIV) transmission event, which allows only a few viruses to establish new infection. However, the genetic characteristics of the transmitted viruses that are preferentially selected have not been fully elucidated. Here, we analyzed amino acid changes in the envelope protein during simian immunodeficiency virus (SIV)/HIV deep transmission history and current HIV evolution within the last 15–20 years. Our results confirmed that the V1V2 region of gp120 protein, particularly V1, was preferentially selected. A shorter V1 region was preferred during transmission history, while during epidemic, HIV may evolve to an expanded V1 region gradually and thus escape immune recognition. We then constructed different HIV-1 V1 mutants using different HIV-1 subtypes to elucidate the role of the V1 region in envelope function. We found that the V1 region, although highly variable, was indispensable for virus entry and infection, probably because V1 deletion mutants exhibited impaired processing of gp160 into mature gp120 and gp41. Additionally, the V1 region affected Env incorporation. These results indicated that the V1 region played a critical role in HIV transmission and infection.**

**KEYWORDS** simian immunodeficiency virus (SIV); human immunodeficiency virus (HIV); transmission; selection; V1 loop; Env function

## INTRODUCTION

The human immunodeficiency virus (HIV) type 1 (HIV-1) envelope (Env) protein consists of a trimer of heterodimers of the gp120 surface protein and the gp41 trans-membrane protein (Dalglish et al., 1984; Klasse,

2012; Cardaci et al., 2013). HIV-1 gp120 is responsible for binding both the CD4 receptor and coreceptors (CCR5 or CXCR4) on the target cell, whereas gp41, together with gp120, mediates fusion of the viral and host cell membranes for virus entry (Alkhatib et al., 1996; Deng et al., 1996; Feng et al., 1996). The gp120 monomer has five highly conserved regions (C1–C5) and five variable regions (V1–V5) (Starcich et al., 1986; Willey et al., 1986). Due to the error-prone nature of HIV reverse transcriptase and extensive selection pressure from the host immune system, the HIV genome has undergone many mutations, particularly in the variable loop regions of the *env* gene, to escape immune recognition. Thus, after disease progression, viral diversity inside the host is extremely high. However, transmission of HIV involves a

Received: 11 January 2016, Accepted: 28 March 2016,

Published online: 19 April 2016

✉Correspondence:

Rongge Yang, Phone: +86-27-87198736, Fax: +86-27-87198736,

Email: ryang@wh.iov.cn

ORCID: 0000-0002-6333-7941

Binlian Sun, Phone: +86-27-87198736, Fax: +86-27-87198736,

Email: sunbl@wh.iov.cn

ORCID: 0000-0002-3060-9867

population bottleneck, in which a limited number of variants from a genetically diverse viral quasi species in the transmitting partner establishes productive infection in the newly infected partner (Salazar-Gonzalez et al., 2009; Shaw and Hunter, 2012; Joseph and Swanstrom, 2014). Many researchers have attempted to elucidate the specific genetic characteristics of these transmitted viruses that are preferentially selected during each transmission event (Keele et al., 2008; Sagar et al., 2009; Boeras et al., 2011; Wilen et al., 2011; Frange et al., 2013; Parrish et al., 2013; Del Prete et al., 2014; Gonzalez et al., 2015); such information is expected to provide important insights into vaccine development.

During the course of HIV infection, the V1V2 region exhibits the most diversity, with many substitutions, deletions, and insertions, and the statistically longer V1V2 region of chronically infected viruses is believed to be the result of escape from the immune response (Sagar et al., 2006; Curlin et al., 2010; Liao et al., 2013a; Nie et al., 2014). In addition, these longer V1V2 viruses are more resistant to specific neutralizing antibodies, and the adaptation of V1V2 is also found at the population level during HIV epidemics, with contemporary viruses having modestly longer V1V2 regions and more glycosylated sites than historical viruses (Bunnik et al., 2010; van Gils et al., 2011). Recent studies indicate that this genetic signature may have conferred increased neutralization resistance to Env over the course of the HIV epidemic (van Gils et al., 2011). However, although not exclusive, current data suggest that newly transmitted viruses encode shorter and less glycosylated variable loop regions, particularly for the V1V2 region, compared with that encoding Env acquired from chronically infected partners (Chohan et al., 2005; Liao et al., 2013b). Additionally, some studies have indicated that these genetic signatures exhibit enhanced fusion and virion incorporation capacity, which can be advantageous during transmission (Cavrois et al., 2014). Moreover, in addition to its potential role in transmission and immune escape, V1V2 also influences virus entry and is a new target for broadly neutralizing antibodies (bNabs) (O'Connell et al., 2014).

Both HIV and its precursor SIVcpz originated from cross-species transmission events, and many studies have investigated the transmission origin of these viruses (Worobey et al., 2010; Sharp and Hahn, 2011; Sharp et al., 2013). Unlike HIV, ancient SIVs have undergone adaptive evolution over thousands of years. Thus, analysis of the changes in gp120 genotypes during the transmission history between SIVs and HIV may provide insights into the V1V2 signature during transmission and subsequent adaptation. Subsequent functional studies may improve our understanding of the role of the V1V2 signature in HIV infection and transmission.

Accordingly, in this study, we analyzed gp120 geno-

types from SIV and HIV sequences and confirmed that the V1V2 region, particularly the V1 region, may play an important role in HIV transmission or function. Next, based on the sequence characteristics, we constructed various Env mutants from different HIV-1 subtypes in order to elucidate the role of the V1 region in Env function.

## MATERIALS AND METHODS

### Cells, plasmids, and primers

293T and TZM-bl cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco, Grand Island, USA) supplemented with 1% glutamine and 10% fetal bovine serum (Gibco). The 293CD4R5X4 dual tropism cell lines were established by transfection with a CCR5-expressing plasmid (Vigene, Rockville, USA), selected with puromycin, and maintained in DMEM containing 0.7  $\mu\text{g}/\text{mL}$  puromycin. All other conditions were as described above. All cell cultures were grown at 37 °C in an atmosphere containing 5% CO<sub>2</sub>. Dual split protein (DSP) assays were used to determine the fusion capacity of the HIV-1 isolates. Stable cell lines expressing DSP1–7 and DSP8–11 were a kind gift from Prof. Matsuda of the Institute of Biophysics, CAS (Kondo et al., 2010; Ishikawa et al., 2012; Wang et al., 2014).

Construction of  $\Delta\text{V1}$  deletion mutants ( $\Delta\text{V1}$ ) for different HIV-1 subtypes was carried out as follows. Wild-type *env* genes of five HIV-1 subtypes cloned from blood samples of patients with HIV were inserted into pcDNA 3.1D/V5-His-TOPO (Invitrogen, Carlsbad, USA) as a template for mutagenesis, and different V1 mutants were constructed using corresponding primers (Table 1). The subtype B V1 loop was also replaced by Gly-Gly-Gly-Gly-Ser (GGGGS) linkers of different lengths to generate different V1 loop replaced mutants. All mutants were confirmed by sequencing of the entire *env* gene.

All the primers used in this study are listed in Table 1.

### Virus production and infection

HIV-1 stocks encoding the luciferase reporter gene were produced by transfecting the HIV-1 proviral constructs pNL4-3.LucE and pcDNA 3.1 (containing Env) into 293T cells using Lipo2000 reagents (Life Tech., Grand Island, USA). One day after transfection, the cell medium was refreshed, and cells were incubated for another 24 h. The supernatants were then harvested and centrifuged to remove cell debris, followed by passing through a 0.45- $\mu\text{m}$  filter and quantification using an HIV-1 p24 enzyme-linked immunosorbent assay (ELISA; Advanced BioScience Laboratories, Rockville, USA). 293CD4R5X4 cell lines plated 1 day earlier were infected with normalized virions at specific concentrations (1 ng p24/10<sup>4</sup> cells). DEAE-dextran was added to each in-

Table 1. Sequences of primers used in this study

Primers	Sequences (5'–3')	Primers	Sequences (5'–3')
B $\Delta$ V1-F	AAGGGCGAGATCAAGAACTGCAGCT	B' $\Delta$ V1-F	GATATGAAAACTGCTCTTTCAATCTC
B $\Delta$ V1-R	GCACTTCAGGCTCACGCACAGGGG	B' $\Delta$ V1-R	AGTGCAATTTAAAGTAACACAGAGTGG
B1GS-F	CGGCAGCAAGGGCGAGATCAAGAAC TGCAGCT	01_AE $\Delta$ V1-F	GAAGTAAGAACTGTTCTTTAATGTG
B1GS-R	CCGCCACCGCACTTCAGGCTCACGC ACAGGGG	01_AE $\Delta$ V1-R	GGAACAATTTAAAGTAACACAGAGAGG
B2GS-F	GGCGGTGGCGGCTCTAAGGGCGAGA TCAAGAACTGCAGCT	07_BC $\Delta$ V1-F	GAAATGAAAAATTGCTCTTTCAATACA
B2GS-R	GCTGCCGCCGCCACCGCACTTCAGG CTCACGCACAGGGG	07_BC $\Delta$ V1-R	ACATTCTAAAGTGACACAGAGTGGGGT
B3GS-F	GGCGGTGGCGGCTCTCGGCAGCAAG GGCGAGATCAAGAAC	08_BC $\Delta$ V1-F	GGAGTAAAAAATTGCTCTTTCAATGCA
B3GS-R	GCTGCCGCCGCCACCCCGCCACCGC ACTTCAGGCTCACGC	08_BC $\Delta$ V1-R	ACAGTTTAAAGTGACACAGAGTGGGGT
B4GS-F	GGCGGTGGCGGCTCTGGCGGTGGC GGCTCTAAGGGCGAGA	--	--
B4GS-R	GCTGCCGCCGCCACCGCTGCCGCCG CCACCGCACTTCAGG	--	--
B5GS-F	GCGGCAGCGGTGGCGGCGGCAGCG GCGGTGGCGGCTCTGGCGGT	--	--
B5GS-R	CGCCACCGCACTTCAGGCTCACGCA CAGGGGGTTCAGCTTCACG	--	--
B6GS-F	GGCGGTGGCGGCTCTCGGCAGCAAG GGCGAGATCAAGAAC	--	--
B6GS-R	GCTGCCGCCGCCACCCCGCCACCGC ACTTCAGGCTCACGC	--	--
B7GS-F	GGCGGTGGCGGCTCTGGCGGTGGC GGCTCTAAGGGCGAGA	--	--
B7GS-R	GCTGCCGCCGCCACCGCTGCCGCCG CCACCGCACTTCAGG	--	--

fection well at a final concentration of 15  $\mu$ g/mL. The cells were harvested 24 or 48 h later, washed twice with 1 $\times$  phosphate-buffered saline (PBS; Gibco), and then used in luciferase reporter assays. Relative infectivity was calculated by dividing the Log<sub>10</sub> (RLU of the mutant virus) by the Log<sub>10</sub> (RLU of the wild-type virus).

### Western blot analysis

Proteins isolated from cells and HIV virions were obtained from different experiments.

For Env expression assays, 293T cells (> 90% confluence in 6-well culture plates; Corning, USA) were transfected with 2  $\mu$ g of each Env expression plasmid, incubated at 37 °C for 24 h, washed with PBS, and centrifuged at 800 $\times$  g for 5 min. Harvested cells were lysed using RIPA lysis buffer, and the lysates were then subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Gp160 and gp120 were detected using sheep anti-gp120 antibodies (Aalto Bio, Iceland). Other antibodies used for western blotting were as follows: anti-GAPDH monoclonal antibodies (MAbs; Beyotime, Guangzhou, China), horseradish peroxidase (HRP)-linked anti-mouse/rabbit secondary antibodies (Cell Signaling Technology, Danvers, USA), and anti-

HIV-1 p24 MAbs (Santa Cruz Biotechnology, Santa Cruz, USA).

For Env incorporation into virions, virus (equivalent to 100 ng p24) was precipitated using 5 $\times$  PEG-it virus precipitation solution (SBI, Shanghai, China), lysed using P0013, and boiled in a defined amount of sample buffer (containing 1/19 volume of  $\beta$ -ME; Bio-Rad, Hercules, USA). The prepared protein samples were used for western blotting, as described above.

### Infectivity assay

The infectivity of each pseudovirus containing different Env mutants was determined using luciferase reporter assays. Briefly, cells infected with HIV-1 reporter pseudoviruses were harvested, washed twice with 1 $\times$  PBS, and then assayed using the Luciferase Assay System (Promega, Madison, USA) following the manufacturer's instructions. Relative infectivity was then calculated by dividing the Log<sub>10</sub> (RLU of the mutant virus) by the Log<sub>10</sub> (RLU of the wild-type virus).

### Membrane fusion assay

For fusion assays, cell lines stably expressing DSP were

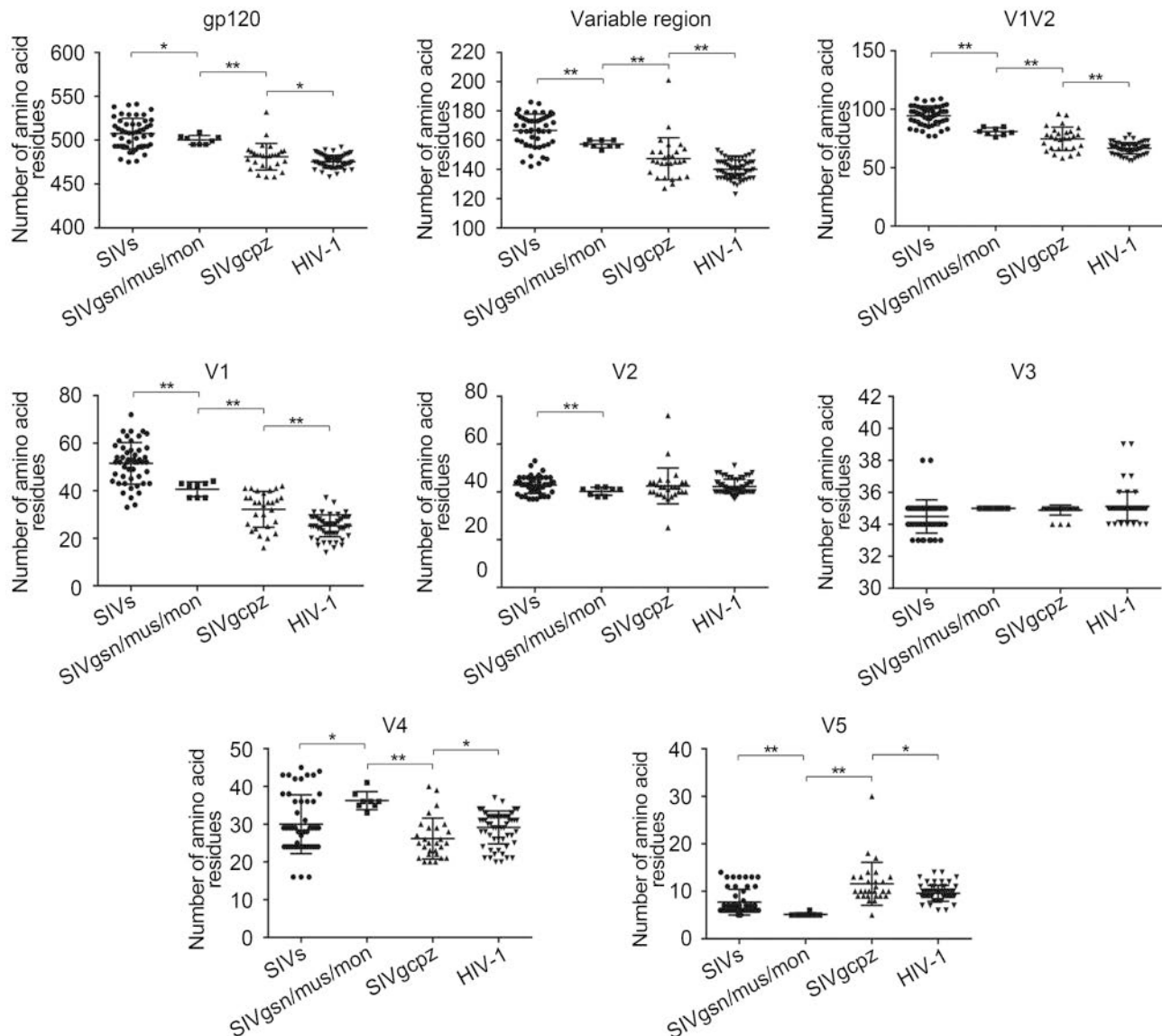
utilized. Stable cell lines expressing DSP8–11 were transfected with expression vectors of interest for Env constructs in triplets. At 48 h post-transfection, 293CD4/DSP1–7 cells, a stable cell line expressing the CD4 molecule and DSP1–7, were cocultured with transfected 293FT/DSP8–11 cells at 37 °C in fresh medium. After about 4–6 h of incubation, membrane fusion was analyzed using fluorescence microscopy (Ishikawa et al., 2012).

The 293CD4X4R5 cell line was also used to assay membrane fusion. Briefly, 293CD4X4R5 cells in 12- or 24-well plates were transfected with different Env expression plasmids. After 24 h of incubation at 37 °C, membrane fusion was directly visualized by microscopy.

### Sequence analysis

All SIV and HIV *env* sequences used in this study were downloaded from the LANL HIV sequence database. For comparison of *env* in the SIV and HIV-1/2 transmission history, *env* sequences were grouped into four datasets (SIVs, SIVgsn/mon/mus, SIVcpz, and HIV-1) for SIV/HIV-1 and three datasets (SIVsmm, SIVmac, and HIV-2) for SIV/HIV-2, according to known transmission history. Unlike HIV, SIVs have only been shown to have a few variations in sequences; therefore, we used as many sequences as possible based on the HIV Sequence Compendium 2014. Additionally, for HIV-1 and HIV-2 (Figure 1), because many different HIV sequences have been identified to date, we only chose representative standard sequences for each group and subtype accord-

A



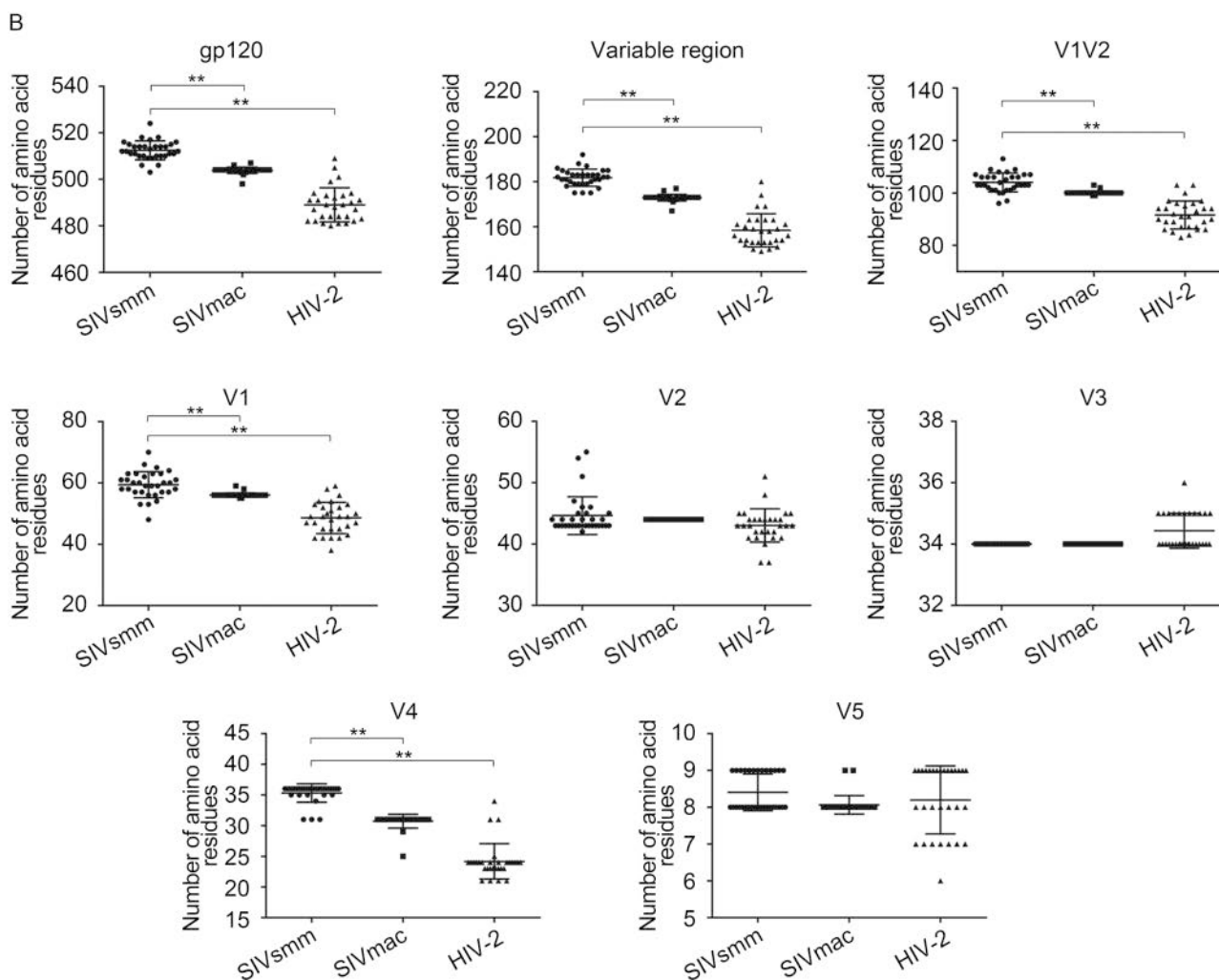


Figure 1. Sequence changes in gp120 protein during the SIV and HIV-1/2 transmission history. Amino acid numbers in each region of the gp120 protein were calculated using online tools for SIV/HIV-1 (A) and SIV/HIV-2 (B). Student's *t*-tests were performed to determine the significance of differences between the two groups; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ .

ing to the HIV Sequence Compendium 2014. For comparison of HIV-1 *env* sequences between different epidemic periods (Figure 2), we attempted to compare changes in the lengths of the V1V2 region during a 15–20-year evolution period. Thus, each sequence was selected based on its sampling date. Sequences having 15–20 years of sampling (defined here as historical and contemporary sequences) were chosen. Only one sequence for each patient was included, and we adjusted the sampling dates of the historical and contemporary sequences to ensure that there were sufficient numbers of sequences for each comparison. More information regarding the sequences used in this study is given in Sup-

plementary Tables S1–S3.

Variable loop characteristics were determined using the online tools (<http://hivtools.publichealth.uga.edu/N-Glyco> and [http://www.hiv.lanl.gov/content/sequence/VAR\\_REG\\_CHAR](http://www.hiv.lanl.gov/content/sequence/VAR_REG_CHAR)).

### Statistical analysis

Statistical analysis was carried out using GraphPad Prism v6 (La Jolla, CA 92037 USA). The data from the luciferase reporter assays are presented as the means  $\pm$  standard deviations (SDs) of three independent experiments. For comparisons of two datasets, Student's *t*-tests were performed, and differences with *P* values of less than 0.05 were considered significant.

## RESULTS

### The shorter V1 region was preferentially selected during SIV and HIV-1/2 transmission history

Existing SIV sequences, which are ancestors of HIV, may serve as “fossils” for analysis of *env* gene evolution during SIV/HIV transmission history. In this study, we collected two groups of SIV and HIV sequences, which showed clear transmission events, and we analyzed changes in gp120 sequence characteristics during each transmission history. When we focused on the SIV gp120 proteins, major differences were observed as compared with that of HIV-1 (Figure 1A). Generally, most SIVs encoded much longer gp120 protein (507.5 amino acids [aa]) than that encoded by HIV-1 (475.4 aa). However, SIVcpz, the precursor of HIV-1, encoded a gp120 protein having almost the same length (481.1 aa) as that of HIV-1. Interestingly, the longer gp120 in most SIVs could be explained by the increased length of the V1 region (51.5 aa). Even SIVcpz, which had almost the same length as gp120, encoded a much longer V1 region (32.2 aa) than HIV-1 (25.3 aa). Furthermore, SIVcpz was the result of recombination between SIVrcm and an unknown SIV from the SIVgsn/mon/mus lineage, which could account for the Env region of SIVcpz. We found that the V1 segment of viruses from this lineage was shorter (40.6 aa) than that of other ancient SIVs (51.5 aa) but longer than that of SIVcpz (32.2 aa). Additionally, we observed the same phenomenon in SIV/HIV-2 transmission history (Figure 1B). SIVsmm, which was the ancestor of SIVmac and HIV-2, had a much longer V1 region (59.4, 56.1, and 48.6 aa for SIVsmm, SIVmac, and HIV-2 respectively). This phenomenon, together with the

observed preferential selection for the shorter V1 region during current HIV-1 transmission, indicated that the V1 region may play an essential role in transmission and selection preferences (Dang et al., 2014).

### The V1 loop length increased slightly over a period of 15–20 years

Although the transmission bottleneck selects for viruses with compact loops and fewer N-glycosylation sites, the HIV-1 subtype B has been reported to escape from immune recognition by ongoing insertions in the V1V2 loop, yielding a longer V1V2 loop after 10–20 years of evolution. Therefore, we analyzed many available sequences from six subtypes and circulating recombinant forms (CRFs) that had a 15–20-year time span (Figure 2). A slight increase (0.5–3.0 amino acids for different subtypes) in V1 length was observed for all subtypes, with significance observed for subtypes B and C owing to the relatively high numbers of available sequences for these subtypes. In contrast, no clear increase was observed for the V2 region. Consistent with previous results, these data indicated that there was ongoing genetic drift in the V1 region to acquire more amino acids, which was believed to mediate escape from neutralizing antibody (Nab) recognition (Bunnik et al., 2010; van Gils et al., 2011).

### The function of the V1 region could not be replaced by amino acid linkers of different lengths

Next, by comparing sequence changes during SIV/HIV transmission, we found that the V1 region may play an important role in transmission selection, particularly for the length of the V1 region. In order to determine wheth-

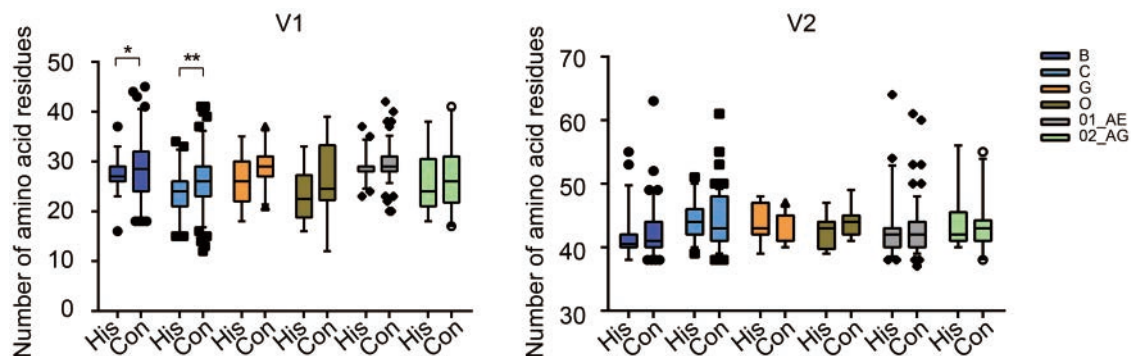


Figure 2. Changes in the sequence of the V1V2 region after 15–20 years of evolution for different HIV-1 subtypes. Amino acid numbers in the V1 and V2 regions were calculated for subtypes B, C, G, group O, CRF01\_AE, and 02\_AG, for which there were sufficient numbers of sequences with sampling date spaced apart by 15–20 years. Student’s *t*-tests were performed to determine the significance of differences between the two groups; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ . #His: historical sequences; Con: contemporary sequences.

er the length of the V1 region affected Env function, we used Env of HXB2 as backbone and established a series of mutations in which the V1 region was deleted and replaced with different lengths of amino acids (1× GGGGS to 7× GGGGS). First, we used a membrane fusion assay system to detect the functions of wild-type and mutant Env. The results showed that none of the Env constructs (V1 deleted or replaced with different linkers) could initiate target cell membrane fusion (Figure 3A). Then, we produced pseudoviruses containing these different Env constructs and performed infectivity assays. Similar to the above findings, compared with the wild-type virus, all pseudoviruses lost infectivity (Figure 3B). These results indicated that the V1 region was indispensable for Env function. Although there did not appear to be any

conservative sequences in this highly variable region, simply replacing the disorganized amino acids could impair Env function.

### V1 mutants may influence normal Env function by impairing gp160 processing

To determine whether different mutations affected Env expression and incorporation, which could influence virus infectivity, we analyzed the gp120 expression and incorporation capacity of subtype B (HXB2) by western blotting (Figure 3C, 3D). The results showed that most V1 mutants expressed similar levels of precursor gp160, whereas  $\Delta$ V1 and 1G4S expressed lower levels of gp160. However, compared with the wild-type virus, the level of

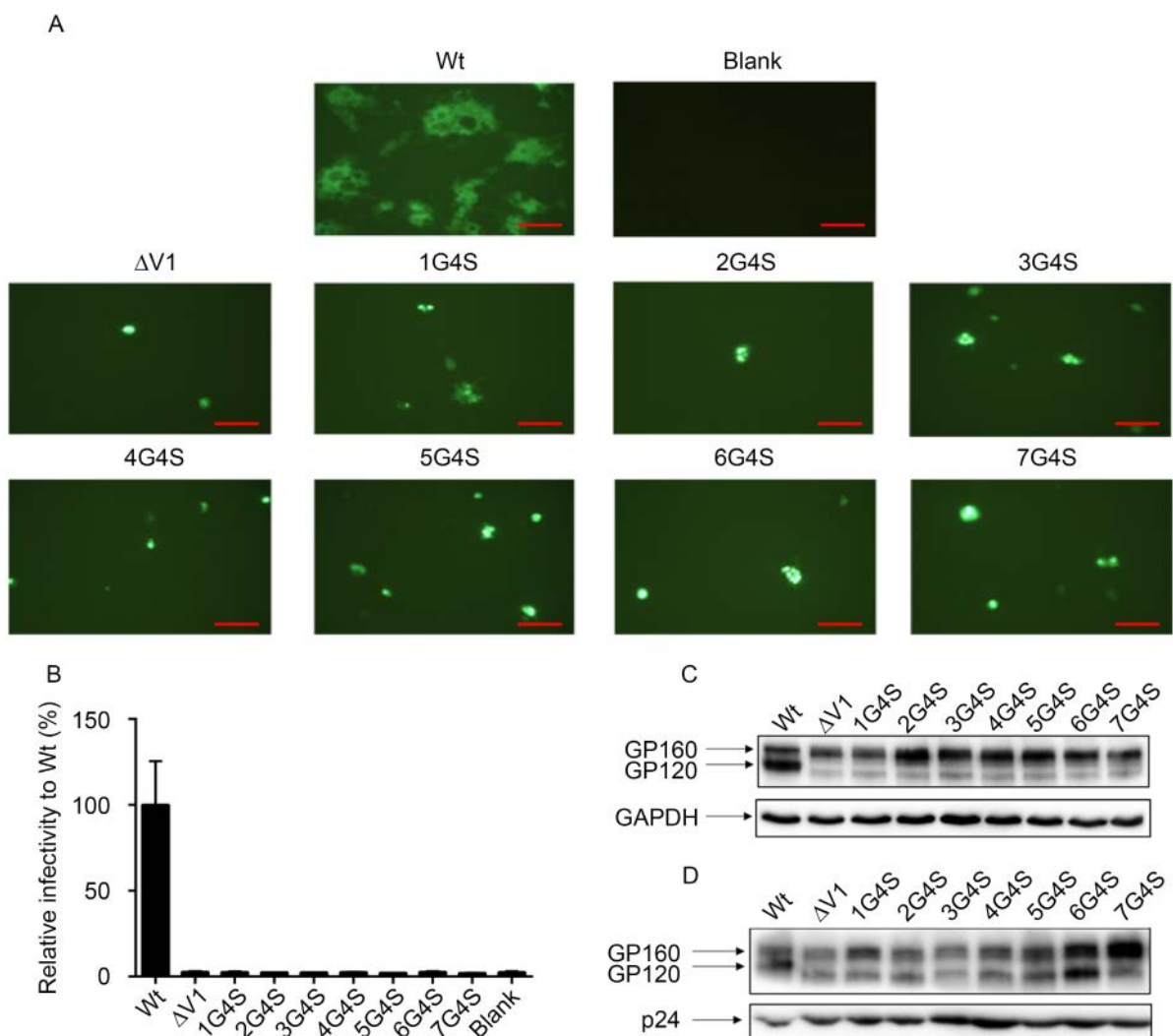


Figure 3. The V1 loop was essential for Env function. The effects of different lengths of V1 amino acids on membrane fusion capacity (A) and virus infectivity (B) were analyzed. For infectivity assays, all mutants were normalized to the wild-type Env infectivity (100%). Western blot analysis of gp160 and gp120 expression in 293T cells and incorporation into the virus for wild-type viruses and different mutant viruses are shown in (C) and (D), respectively. Scale bars (red line) represent 100  $\mu$ m.

gp120 (relative to total gp160 expression, i.e., gp120 / gp120 + gp160) in the mutant Env-expressing cells (Figure 3C) and pseudoviruses (Figure 3D) was dramatically reduced, suggesting that V1 deletion or replacement impaired the processing of gp160. Gp160 processing into mature gp120 and gp41 is a prerequisite for normal Env function; thus, loss of function of these mutants may be due to impairment of gp160 processing capacity by the changes in the V1 sequence. These results suggested that specific V1 sequences may be critical for gp160 processing in HXB2.

### The V1 loop may be necessary for Env function in different HIV-1 subtypes

In order to determine whether the observed role of the

V1 region was independent of the specific clone or subtype, we amplified the *env* gene from four different subtypes and CRFs (B', 01\_AE, 07\_BC, and 08\_BC) isolated from patient blood samples. These genes were then cloned into the pcDNA 3.1-TOPO plasmid, and corresponding V1-deleted Env clones were constructed. Using these clones, we first examined membrane fusion capacity in 293CD4R5X4 cells. The results showed that compared with the wild-type construct, all V1-deleted Env proteins for each subtype exhibited loss of fusion capacity (Figure 4A). Next, we performed infectivity assays for pseudoviruses for each subtype (Figure 4B). The results also showed that the V1 loop was indispensable for virus infection. Next, we analyzed the expression and incorporation of Env (Figure 4C).

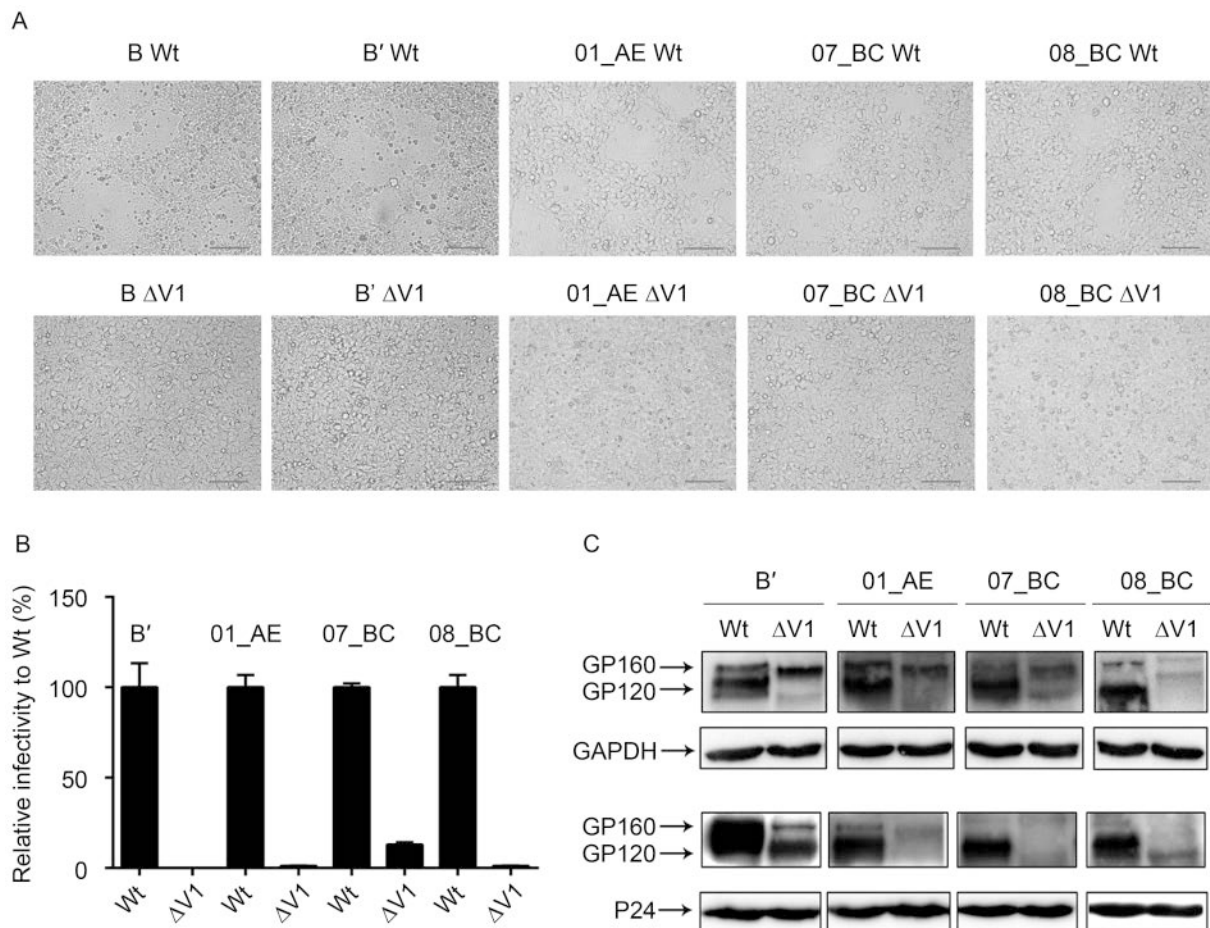


Figure 4. Confirmation of V1 loop function in four subtypes of HIV-1. Wild-type and V1-deleted Env proteins of subtypes B', 01\_AE, 07\_BC, and 08\_BC were analyzed to determine fusion capacity in 293TCD4R5X4 cell lines. Membrane fusion was visualized at 24 h after transfection (A). Relative infectivity of viruses harboring different V1-deleted Env proteins normalized to that of the wild-type virus (100%) was determined by luciferase reporter assay (B). Western blot analysis of gp160 and gp120 expression in 293T cells and incorporation into the virus for wild-type and V1-deleted Env (C). Scale bars (red line) represent 50  $\mu$ m.

Interestingly, we also observed similar results in that, relative to the wild-type Env, all V1-deleted Env mutants exhibited greatly reduced gp160 processing capacity in cell lysates, although some mutants had weak Env expression relative to that of the wild-type virus (Figure 4C, upper panel). Consistent with this, all mutant viruses exhibited reduced or lack of incorporated gp120 (Figure 4C, lower panel), suggesting that Env incorporation was impaired in these V1 mutants. However, these results may also be explained by weak Env expression and maturation.

Taken together, our findings confirmed that the V1 loop was indispensable for Env function, possibly due to impaired Env processing and incorporation. Importantly, this function was observed in different HIV-1 subtypes.

## DISCUSSION

The transmission bottleneck plays an important role in HIV transmission and evolution, and identification of the preferentially selected phenotype will improve our understanding of HIV evolution and facilitate the development of vaccines or drugs that target specific viruses. Many studies have shown that the V1V2 loop of gp120 is a key region in NAb recognition and may act as a selective target in transmission and evolution, particularly because of the length and N-glycan modifications of this region (Chohan et al., 2005; Sagar et al., 2006; Samleerat et al., 2008; van Gils et al., 2011). However, these studies have mainly focused on a few subtypes, and the results have not been consistent due to limitations in available datasets. Therefore, in this study, we first used SIV and HIV gp120 sequences that had a clear transmission history and used these sequences to analyze changes in the sequence length during this transmission and evolution flow. Our results provided important insights into the evolution and role of the V1 region of the HIV sequence and may facilitate the development of novel HIV vaccines.

In this study, we found that shorter V1V2 regions, particularly for V1, were selected during cross-species transmission. Moreover, this phenomenon was observed for both HIV-1 and HIV-2, indicating that selection bias of the V1 region may not be a coincidence. In fact, some studies have shown that natural SIV transmission in the wild is also associated with a genetic bottleneck similar to that observed for HIV transmission (Stone et al., 2010). According to current data, we speculate that the transmission bottleneck and subsequent adaptive evolution may account for the changes in V1V2 length observed in different cross-species transmission events. Some unknown SIVs gave rise to SIV<sub>gsn/mon/mus</sub>, and the sequences were then transmitted to SIV<sub>cpz</sub> and subsequently to HIV-1; in accordance with the selection bi-

as, the observed V1V2 length decreased during this transmission history. Additionally, the extensive transmission of HIV-1 among human populations has shown a similar selection bias. Notably, in addition to the V1 region, the V4 region in HIV-1/SIV transmission history has also exhibited changes, as shown in Figure 1A, with selection of longer or shorter V4 regions. The V4 region in the HIV-2/SIV transmission history showed changes similar to those observed for V1 (i.e., a shorter V4 sequence was selected). Little is known about V4; however, V4 is thought to be critical for Env entry into specific HIV strains (Li et al., 2014), and recent studies using both sequences from the HIV database and transmission clusters have found a negative correlation between the V1V2 and V4V5 regions in terms of amino acid lengths during transmission. Notably, this correlation was lost following seroconversion (Pollakis et al., 2015), indicating that the relationship between these regions may be complex. Based on our current findings, V4 may also play an important role in transmission. However, additional studies are needed to support this hypothesis.

The relative long V1V2 region and additional glycosylation sites serve to protect the virus from the humoral immune response when the virus transmits to a new host. Because the naïve host lacks an HIV-specific immune response, these escape features are no longer necessary, and any fitness cost associated with them may hinder transmission. Although the transmission bottleneck selects viruses with compact loops and fewer N-glycosylation sites, after years of evolution, the virus has adapted to the humoral immune response by enhancing the mask of its envelope. As a result, since the zoonotic transmission from SIV<sub>cpz</sub>, HIV has been evolving to adapt to the human host, and one of the escape features that HIV has acquired is extension of the V1 region.

The population-wide evolution rate of HIV is low compared with the dramatic sequence changes that occur in each host (Curlin et al., 2010), and changes in the V1 loop at the population level were not obvious, as we showed in Figure 2. This could be due to selection bias, resulting in a step back in evolution to consensus (Carlson et al., 2014). Moreover, HIV has a relatively short evolutionary time compared with that of SIV. Most SIVs encode a much longer V1V2 region, and infection by SIVs in natural hosts is nonpathogenic, with the exception of SIV<sub>cpz</sub>, which shows symptoms similar to those of acquired immunodeficiency syndrome (AIDS) in the natural host. The genome of SIV<sub>cpz</sub> encodes a V1V2 region that is longer than that of HIV but shorter than that of other ancient SIVs. This may be the result of its relatively moderate evolutionary history compared with that of other SIVs and HIV. Moreover, because HIV does not have a relatively long history, we may expect that given a much deeper coevolution between HIV and its human

hosts, HIV can evolve with a more adaptive genome, and one of the features of this adaptive genome could be an even longer V1V2 region than that observed in SIV today.

Based on our results, we next constructed different V1 mutants in order to elucidate the function of V1. Although V1 appeared to be comprised of disorganized, highly variable amino acids, this region may function in synergy with other parts of Env. Additionally, consistent with previous studies (Hamoudi et al., 2013; Yuan et al., 2013), replacing or changing the V1 region may also impair Env function, supporting the sophisticated evolutionary mechanism of the Env protein. Moreover, deletion of the V1 region indicated that this region was essential for membrane fusion and virus infection, and these findings were confirmed in all subtypes examined in this study. Furthermore, the V1 region may influence Env processing and incorporation, as was shown by western blot analysis. In all subtypes examined in this study, the V1 region may influence Env processing, thereby leading to impairment of Env function. In some subtypes (i.e., 07\_BC, 08\_BC, and 01\_AE), in addition to impaired Env processing, we also observed reduced Env incorporation, which may be due to the impaired expression of Env, resulting in even lower packaging capacity. Unlike HXB2, the other subtypes used in this study were cloned from blood in our laboratory, and the expression capacity was not sufficiently high; this may affect the overall expression of Env and subsequent packaging of the virus. However, the V1 mutation in these subtypes may also directly influence Env incorporation, leading to failed infection, as observed in this study. Additional studies are needed in order to fully elucidate the mechanisms involved in these processes. Env protein is first synthesized as precursor gp160, which is cleaved into mature gp120 and gp41 by host enzymes (Wang et al., 2008). Mutations in the V1 region may alter the conformation of the cleavage sites, resulting in weak processing efficiency and low gp120 incorporation. Thus, the V1 region is indispensable for Env function and virus infectivity.

In summary, we found, for the first time, that the gp120 V1 region was preferentially selected during SIV/HIV transmission and evolution and was indispensable for HIV infection, suggesting that the V1 region may serve as an attractive target for designing new HIV prevention tools such as vaccines and drugs. However, like other studies, there are still a lot limitations of our study. For example, if it is possible we can use more sequences of all the subtypes, and divide them into more epidemic groups, this may give us a gradual change of V1 region following evolution; and in order to see the influence of V1 length on envelope function, instead of the linkers we used here, we can exchange the V1 region

between different clones. Besides, *env* characteristic changes in real human transmission pairs may provide more useful evidence to our results. Further studies of the sensitivity of V1 mutants and glycan modifications in this region to sCD4 and different NAbs are needed in order to determine how V1 amino acids and glycans affect *env* function and escape from immune recognition.

## ACKNOWLEDGMENTS

This work was supported by the International Science & Technology Cooperation Program of China (2011DFA31030) and Deutsche Forschungsgemeinschaft (Transregio TRR60), National Natural Science Foundation of China (No. 81461130019), the Key Laboratory on Emerging Infectious Diseases and Biosafety in Wuhan and The Core Facility and Technical Support, Wuhan Institute of Virology. And we would like to thank Prof. Matsuda from Institute of Biophysics, CAS, for cell lines expressing DSP1–7 and DSP8–11.

## COMPLIANCE WITH ETHICS GUIDELINES

The authors declare that they have no conflict of interest. This article does not contain any studies with human or animal subjects performed by any of the authors.

## AUTHOR CONTRIBUTIONS

YPL, BLS and RGY conceived and designed the experiments; YPL and JPS performed the experiments; YPL and YW analyzed the data; YPL, BLS and Ulf Dittmer wrote and edited the paper. All authors have read and approved the final manuscript.

Supplementary tables are available on the website of *Virologica Sinica*: [www.virosin.org](http://www.virosin.org); [link.springer.com/journal/12250](http://link.springer.com/journal/12250).

## REFERENCES

- Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, Murphy PM, Berger EA. 1996. CC CKR5: A RANTES, MIP-1 alpha, MIP-1 beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science*, 272: 1955–1958.
- Boeras DI, Hraber PT, Hurlston M, Evans-Strickfaden T, Bhattacharya T, Giorgi EE, Mulenga J, Karita E, Korber BT, Allen S, Hart CE, Derdeyn CA, Hunter E. 2011. Role of donor genital tract HIV-1 diversity in the transmission bottleneck. *Proc Natl Acad Sci U S A*, 108: E1156–E1163.
- Bunnik EM, Euler Z, Welkers MR, Boeser-Nunnink BD, Grijnsen ML, Prins JM, Schuitemaker H. 2010. Adaptation of HIV-1 envelope gp120 to humoral immunity at a population level. *Nat Med*, 16: 995–997.
- Cardaci S, Soster M, Bussolino F, Marchio S. 2013. The V1/V2 loop of HIV-1 gp120 is necessary for Tat binding and con-

- sequent modulation of virus entry. *FEBS Lett*, 587: 2943–2951.
- Carlson JM, Schaefer M, Monaco DC, Batorsky R, Claiborne DT, Prince J, Deymier MJ, Ende ZS, Klatt NR, DeZiel CE, Lin TH, Peng J, Seese AM, Shapiro R, Frater J, Ndung'u T, Tang J, Goepfert P, Gilmour J, Price MA, Kilembe W, Heckerman D, Goulder PJ, Allen TM, Allen S, Hunter E. 2014. HIV transmission. Selection bias at the heterosexual HIV-1 transmission bottleneck. *Science*, 345: 1254031.
- Cavrois M, Neideman J, Santiago ML, Derdeyn CA, Hunter E, Greene WC. 2014. Enhanced fusion and virion incorporation for HIV-1 subtype C envelope glycoproteins with compact V1/V2 domains. *J Virol*, 88: 2083–2094.
- Chohan B, Lang D, Sagar M, Korber B, Lavreys L, Richardson B, Overbaugh J. 2005. 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*, 79: 6528–6531.
- Curlin ME, Zioni R, Hawes SE, Liu Y, Deng W, Gottlieb GS, Zhu T, Mullins JI. 2010. HIV-1 envelope subregion length variation during disease progression. *PLoS Pathog*, 6: e1001228.
- Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA. 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*, 312: 763–767.
- Dang S, Wang Y, Budeus B, Verheyen J, Yang R, Hoffmann D. 2014. Differential selection in HIV-1 gp120 between subtype B and East Asian variant B'. *Virol Sin*, 29: 40–47.
- Del Prete GQ, Ailers B, Moldt B, Keele BF, Estes JD, Rodriguez A, Sampias M, Oswald K, Fast R, Trubey CM, Chertova E, Smedley J, LaBranche CC, Montefiori DC, Burton DR, Shaw GM, Markowitz M, Piatak M, Jr., KewalRamani VN, Bieniasz PD, Lifson JD, Hatzioannou T. 2014. Selection of unadapted, pathogenic SHIVs encoding newly transmitted HIV-1 envelope proteins. *Cell Host Microbe*, 16: 412–418.
- Deng HK, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhardt M, DiMarzio P, Marmon S, Sutton RE, Hill CM, Davis CB, Peiper SC, Schall TJ, Littman DR, Landau NR. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*, 381: 661–666.
- Feng Y, Broder CC, Kennedy PE, Berger EA. 1996. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, 272: 872–877.
- Frango P, Meyer L, Jung M, Goujard C, Zucman D, Abel S, Hochedez P, Gousset M, Gascuel O, Rouzioux C, Chaix ML, Group APCS. 2013. Sexually-transmitted/founder HIV-1 cannot be directly predicted from plasma or PBMC-derived viral quasispecies in the transmitting partner. *PLoS One*, 8: e69144.
- Gonzalez MW, DeVico AL, Lewis GK, Spouge JL. 2015. Conserved molecular signatures in gp120 are associated with the genetic bottleneck during simian immunodeficiency virus (SIV), SIV-human immunodeficiency virus (SHIV), and HIV type 1 (HIV-1) transmission. *J Virol*, 89: 3619–3629.
- Hamoudi M, Simon-Loriere E, Gasser R, Negroni M. 2013. Genetic diversity of the highly variable V1 region interferes with Human Immunodeficiency Virus type 1 envelope functionality. *Retrovirology*, 10: 114.
- Ishikawa H, Meng FX, Kondo N, Iwamoto A, Matsuda Z. 2012. Generation of a dual-functional split-reporter protein for monitoring membrane fusion using self-associating split GFP. *Protein Eng Des Sel*, 25: 813–820.
- Joseph SB, Swanstrom R. 2014. HIV/AIDS. A fitness bottleneck in HIV-1 transmission. *Science*, 345: 136–137.
- Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson JA, Ping LH, Swanstrom R, Tomaras GD, Blattner WA, Goepfert PA, Kilby JM, Saag MS, Delwart EL, Busch MP, Cohen MS, Montefiori DC, Haynes BF, Gaschen B, Athreya GS, Lee HY, Wood N, Seoighe C, Perelson AS, Bhattacharya T, Korber BT, Hahn BH, Shaw GM. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A*, 105: 7552–7557.
- Klasse PJ. 2012. The molecular basis of HIV entry. *Cell Microbiol*, 14: 1183–1192.
- Kondo N, Miyauchi K, Meng FX, Iwamoto A, Matsuda Z. 2010. Conformational Changes of the HIV-1 Envelope Protein during Membrane Fusion Are Inhibited by the Replacement of Its Membrane-spanning Domain. *J Biol Chem*, 285: 14681–14688.
- Li Y, Yang D, Wang JY, Yao Y, Zhang WZ, Wang LJ, Cheng DC, Yang FK, Zhang FM, Zhuang M, Ling H. 2014. Critical amino acids within the human immunodeficiency virus type 1 envelope glycoprotein V4 N- and C-terminals contribute to virus entry. *PLoS One*, 9: e86083.
- Liao HX, Lynch R, Zhou T, Gao F, Alam SM, Boyd SD, Fire AZ, Roskin KM, Schramm CA, Zhang Z, Zhu J, Shapiro L, Program NCS, Mullikin JC, Gnanakaran S, Hraber P, Wiehe K, Kelsoe G, Yang G, Xia SM, Montefiori DC, Parks R, Lloyd KE, Scearce RM, Soderberg KA, Cohen M, Kamanga G, Louder MK, Tran LM, Chen Y, Cai F, Chen S, Moquin S, Du X, Joyce MG, Srivatsan S, Zhang B, Zheng A, Shaw GM, Hahn BH, Kepler TB, Korber BT, Kwong PD, Mascola JR, Haynes BF. 2013a. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature*, 496: 469–476.
- Liao HX, Tsao CY, Alam SM, Muldoon M, Vandergrift N, Ma BJ, Lu XZ, Sutherland LL, Scearce RM, Bowman C, Parks R, Chen HY, Blinn JH, Lapedes A, Watson S, Xia SM, Foulger A, Hahn BH, Shaw GM, Swanstrom R, Montefiori DC, Gao F, Haynes BF, Korber B. 2013b. Antigenicity and Immunogenicity of Transmitted/Founder, Consensus, and Chronic Envelope Glycoproteins of Human Immunodeficiency Virus Type 1. *J Virol*, 87: 4185–4201.
- Nie J, Zhao J, Chen Q, Huang W, Wang Y. 2014. Three amino acid residues in the envelope of human immunodeficiency virus type 1 CRF07\_BC regulate viral neutralization susceptibility to the human monoclonal neutralizing antibody IgG1b12. *Virol Sin*, 29: 299–307.
- O'Connell RJ, Kim JH, Excler JL. 2014. The HIV-1 gp120 V1V2 loop: structure, function and importance for vaccine development. *Expert Rev Vaccines*, 13: 1489–1500.
- Parrish NF, Gao F, Li H, Giorgi EE, Barbian HJ, Parrish EH, Zajic L, Iyer SS, Decker JM, Kumar A, Hora B, Berg A, Cai FP, Hopper J, Denny TN, Ding HT, Ochsenauber C, Kappes JC, Galimidi RP, West AP, Bjorkman PJ, Wilen CB, Doms RW, O'Brien M, Bhardwaj N, Borrow P, Haynes BF, Muldoon M, Theiler JP, Korber B, Shaw GM, Hahn BH. 2013. Phenotypic properties of transmitted founder HIV-1. *Proc Natl Acad Sci U S A*, 110: 6626–6633.
- Pollakis G, Baan E, van Werkhoven MB, Berkhout B, Bakker M, Jurriaans S, Paxton WA. 2015. Association between gp120 envelope V1V2 and V4V5 variable loop profiles in a defined HIV-1 transmission cluster. *AIDS*, 29: 1161–1171.
- Sagar M, Laeyendecker O, Lee S, Gamiel J, Wawer MJ, Gray RH, Serwadda D, Sewankambo NK, Shepherd JC, Toma J, Huang W, Quinn TC. 2009. Selection of HIV variants with signature genotypic characteristics during heterosexual transmission. *J Infect Dis*, 199: 580–589.
- Sagar M, Wu X, Lee S, Overbaugh J. 2006. Human immunodeficiency

- ciency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. *J Virol*, 80: 9586–9598.
- Salazar-Gonzalez JF, Salazar MG, Keele BF, Learn GH, Giorgi EE, Li H, Decker JM, Wang S, Baalwa J, Kraus MH, Parrish NF, Shaw KS, Guffey MB, Bar KJ, Davis KL, Ochsenbauer-Jambor C, Kappes JC, Saag MS, Cohen MS, Mulenga J, Derdeyn CA, Allen S, Hunter E, Markowitz M, Hraber P, Perelson AS, Bhattacharya T, Haynes BF, Korber BT, Hahn BH, Shaw GM. 2009. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J Exp Med*, 206: 1273–1289.
- Samleerat T, Braibant M, Jourdain G, Moreau A, Ngo-Giang-Huong N, Leechanachai P, Hemvuttiphon J, Hinjiranandana T, Changchit T, Warachit B, Suraseranivong V, Lallemand M, Barin F. 2008. Characteristics of HIV type 1 (HIV-1) glycoprotein 120 env sequences in mother-infant pairs infected with HIV-1 subtype CRF01\_AE. *J Infect Dis*, 198: 868–876.
- Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*, 1: a006841.
- Sharp PM, Rayner JC, Hahn BH. 2013. Evolution. Great apes and zoonoses. *Science*, 340: 284–286.
- Shaw GM, Hunter E. 2012. HIV transmission. *Cold Spring Harb Perspect Med*, 2: a006965.
- Starcich BR, Hahn BH, Shaw GM, Meneely PD, Modrow S, Wolf H, Parks ES, Parks WP, Josephs SF, Gallo RC, Wongstaal F. 1986. Identification and Characterization of Conserved and Variable Regions in the Envelope Gene of Htlv-Iii Lav, the Retrovirus of Aids. *Cell*, 45: 637–648.
- Stone M, Keele BF, Ma ZM, Bailes E, Dutra J, Hahn BH, Shaw GM, Miller CJ. 2010. A limited number of simian immunodeficiency virus (SIV) env variants are transmitted to rhesus macaques vaginally inoculated with SIVmac251. *J Virol*, 84: 7083–7095.
- van Gils MJ, Bunnik EM, Boeser-Nunnink BD, Burger JA, Terlouw-Klein M, Verwer N, Schuitemaker H. 2011. Longer V1V2 region with increased number of potential N-linked glycosylation sites in the HIV-1 envelope glycoprotein protects against HIV-specific neutralizing antibodies. *J Virol*, 85: 6986–6995.
- Wang H, Li X, Nakane S, Liu S, Ishikawa H, Iwamoto A, Matsuda Z. 2014. Co-Expression of Foreign Proteins Tethered to HIV-1 Envelope Glycoprotein on the Cell Surface by Introducing an Intervening Second Membrane-Spanning Domain. *Plos One*, 9: e96790.
- Wang J, Sen J, Rong L, Caffrey M. 2008. Role of the HIV gp120 conserved domain 1 in processing and viral entry. *J Biol Chem*, 283: 32644–32649.
- Wilén CB, Parrish NF, Pfaff JM, Decker JM, Henning EA, Haim H, Petersen JE, Wojcechowskyj JA, Sodroski J, Haynes BF, Montefiori DC, Tilton JC, Shaw GM, Hahn BH, Doms RW. 2011. Phenotypic and immunologic comparison of clade B transmitted/founder and chronic HIV-1 envelope glycoproteins. *J Virol*, 85: 8514–8527.
- Willey RL, Rutledge RA, Dias S, Folks T, Theodore T, Buckler CE, Martin MA. 1986. Identification of Conserved and Divergent Domains within the Envelope Gene of the Acquired-Immunodeficiency-Syndrome Retrovirus. *Proc Natl Acad Sci U S A*, 83: 5038–5042.
- Worobey M, Telfer P, Souquiere S, Hunter M, Coleman CA, Metzger MJ, Reed P, Makuwa M, Hearn G, Honarvar S, Roques P, Apetrei C, Kazanji M, Marx PA. 2010. Island biogeography reveals the deep history of SIV. *Science*, 329: 1487.
- Yuan T, Li J, Zhang MY. 2013. HIV-1 envelope glycoprotein variable loops are indispensable for envelope structural integrity and virus entry. *PLoS One*, 8: e69789.