Research Article

The Comparison of Genetic Variation in the Envelope Protein Between Various Immunodeficiency Viruses and Equine Infectious Anemia Virus[*]

Qing Yuan[#], Chang Liu[#], Zhipin Liang, Xueqing Chen, Danhong Diao and Xiaohong Kong[**]

Laboratory of Medical Molecular Virology, School of Medicine, Nankai University, Tianjin 300071, China

The envelope protein (Env) of lentiviruses such as HIV, SIV, FIV and EIAV is larger than that of other retroviruses. The Chinese EIAV attenuated vaccine is based on Env and has helped to successfully control this virus, demonstrating that envelope is crucial for vaccine. We compared Env variation of the four kinds of lentiviruses. Phylogenetic analysis showed that the evolutionary relationship of Env between HIV and SIV was the closest and they appeared to descend from a common ancestor, and the relationship of HIV and EIAV was the furthest. EIAV had the shortest Env length and the least number of potential N-linked glycosylation sites (PNGS) as well as glycosylation density compared to various immunodeficiency viruses. However, HIV had the longest Env length and the most PNGS. Moreover, the alignment of HIV and SIV showed that PNGS were primarily distributed within extracellular membrane protein gp120 rather than transmembrane gp41. It implies that the size difference among these viruses is associated with a lentivirus specific function and also the diversity of *env*. There are low levels of modification of glycosylation sites of Env and selection of optimal protective epitopes might be useful for development of an effective vaccine against HIV/AIDS.

The initial step in cell infection by retroviruses, including lentiviruses occurs with the binding of envelope glycoproteins (Env) to target cells. Knowledge of the genetic variation and glycosylation modification of Env is important for understanding the molecular mechanisms underlying virus-cell interactions and for the design of effective vaccines to viruses. Human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV) and equine infectious anemia virus (EIAV) are all members of the lentivirus genus of the retroviridae family, and infection by members of this group of viruses often lead to immunosuppression such as AIDS caused by HIV. The first and widely used lentivirus vaccine is the vaccine against EIAV, which was developed three decades ago[25].

The underlying mechanisms of persistence and pathogenesis

of infection by HIV and EIAV are similar, albeit the clinical manifestations are different[4]. HIV and EIAV share similar biological properties such as genome structure, viral replication and the conformational structure of some viral structural proteins[10]. Similar to infection by HIV, horses infected by EIAV experience a subclinical phase after recurring cycles of fever, thrombocytopenia, and wasting symptoms[24]. EIAV has also been used as a model to study HIV persistence, pathogenesis and immune response[1,5].

Recently, a number of studies have shown that the N-linked glycosylation of lentivirus Env associate closely with viral tropism, infectivity and viral escape against the host immune response[12,13,15,19,23,27]. Similar to other viruses, tche envelope protein of EIAV plays a critical role in mediating receptor binding to target cells, subsequent entry, and induction of host humoral immune response[3,16]. The glycosylation modification endows the Env and the viruses a number of biological functions to enhance the binding and infection ability. The potential N-linked glycosylation sites (PNGS) of the EIAV envelope protein play a pivotal role in EIAV attenuated vaccine. This vaccine is highly effective in protecting the vaccinated animals from infection by wild type strains. In comparison with the virulent strain, the attenuated EIAV strain has been shown to

Springer

have lower numbers of N-linked glycosylation sites[11,26]. Whether less glycosylation of the EIAV Env contributes to its successes as a vaccine remains to be determined. The effectiveness of the vaccine to protect infection by EIAV virulent strains demonstrated this vaccine can serve as a good model to study the induction of protective immune response against other lentiviruses, including HIV.

Therefore, in this study we attempted to compare the envelope protein of various immunodeficiency viruses such as HIV, SIV and FIV with that of EIAV to gather information about the amino acid length, the number of PNGS and their defining characteristics within Env of these viruses. Furthermore, we analyzed Env sequences to investigate the evolution of these viruses from various lineages. The comparative study of Env from various lineages lentiviruses to EIAV highlighted some common evolutionary characteristics that may help to provide some information for the development of an effective HIV vaccine.

MATERIALS AND METHODS
Sequences collection

HIV-1 *env* sequences comprising subtypes A-K downloaded from the Los Alamos HIV Database (http://hiv.lanl.gov/content/index)[9]. In order to improve the accuracy of the phylogenetic analyses, we excluded previously determined recombinant sequences[30]. A total of 48 *env* from subtype HIV-1 A, B, C, D, F, G, H, J, K, HIV-1 group O and HIV-1 group N were finally selected. Envelope gene sequences of SIV, FIV and EIAV were downloaded from NCBI (http://www.ncbi.nlm.nih.gov).

Phylogenetic Analysis

Env sequences were aligned using CLUSTALX[29]. A Neighbor-Joining (NJ) tree was constructed for all the Env sequences which included subtypes A-K of the HIV-1 Env sequences and all of SIV, EIAV and FIV Env sequences by MEGA 4.1 software. We calculated the evolutionary distance through DNAPars DNA parsimony method supplied provided by the BioEdit 7.0.5 software[7,28].

Potential N-linked Glycosylation Sites (PNGS)

Potential N-linked Glycosylation sites (PNGS) were predicted using N-Glycosite from the Los Alamos National HIV database[32]. (http://www. hiv. lanl. gov/content/sequence/GLYCOSITE/glycosite.html). This is a web-based tool that was developed for tracking and quickly assessing the patterns in N-linked glycosylation sites during protein alignments[32].

Statistical Analysis

To compare the amino acid length and the number of PNGS of envelope gene between various immunodeficiency viruses and EIAV, we used a non-parametric Mann-Whitney test. The *p* values of less than or equal to 0.05 were considered significant. All statistical tests were performed with GraphPad Prism

software version 5.0[31].

RESULTS
Phylogenetic linkage of Env between various immunodeficiency viruses and EIAV

We analyzed 93 Env sequences in total which included HIV-1 Env (n=48), SIV Env (n=14), FIV Env (n=15) and EIAV Env (n=16, including one EIAV vaccine strain sequence), respectively. The 48 HIV Env came from three groups of M, O and N. They include subtype A, B, C, D, F, G, H, J, K from group M. We built a phylogenetic tree, where each leaf corresponds to a different strain of virus whose trait values are known. We calculated the evolutionary distance reflecting the evolutionary relationship among different strains. As shown in Fig. 1, the evolutionary distance between HIV and SIV was 0.257, followed by the distance between EIAV and FIV, which was 0.52. Slightly higher distance was displayed between FIV and HIV, as well as between FIV and SIV, which was 0.604 and 0.63, respectively. The furthest distance observed was between EIAV and SIV, which was 0.781. A similar distance was observed between EIAV and HIV, which was 0.775.

Comparison of the amino acids lengths between Env of various immunodeficiency viruses and EIAV

We examined the genetic features of Env including the amino acid length which may influence the protein folding and reflect the diversity of Env.

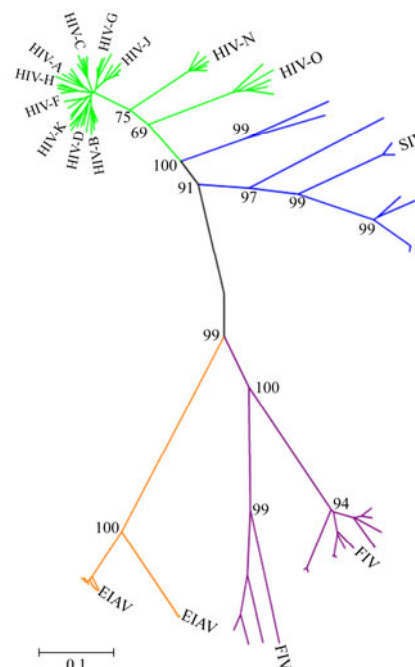We found that the amino acids length of EIAV Env ranges

Fig. 1. A phylogeny of Env from various lineages of immunodeficiency viruses and EIAV. Each kind of immunodeficiency viruses Env sequences were labeled with different colors (HIV, green; SIV, blue; FIV, purple; EIAV, orange). The black line represents the backbone of the Neighbor-Joining tree. The tree was constructed using MEGA version 4.1.

from 561 to 865 (median, 709) and was the shortest among the viruses analyzed. The amino acid length of HIV ranges from 812 to 899 (median, 854) was the longest, followed in turns by SIV (ranging from 832 to 897, median, 852), and FIV (ranging from 810 to 863, median, 853). The results showed that the amino acids lengths of SIV and FIV were both shorter than that of HIV, but the difference was not statistically significant (p>0.05, HIV vs SIV, p>0.05, HIV vs FIV) (Fig. 2D, 2E). Interestingly, the amino acids length of EIAV Env was significantly shorter than that of HIV, FIV and SIV, respectively. The difference was statistically significant (p<0.01) (Table 1 and Fig. 2A, 2B, 2C).

Table 1. Comparison of PNGS, length and relative glycosylation density of Env between various immunodeficiency viruses and EIAV

Virus	Number of samples	Amino acid length of envelope		PNGS		Relative glycosylation density	
		range	median	range	median	range	median
EIAV	16	561–865	709	18–25	21	2.10–3.39	2.85
HIV	48	812–899	854	25–35	30.5	2.93–4.12	3.58
SIV	14	832–897	852	24–28	26	2.17–3.13	3.05
FIV	15	810–863	853	20–29	23	2.34–3.57	2.69

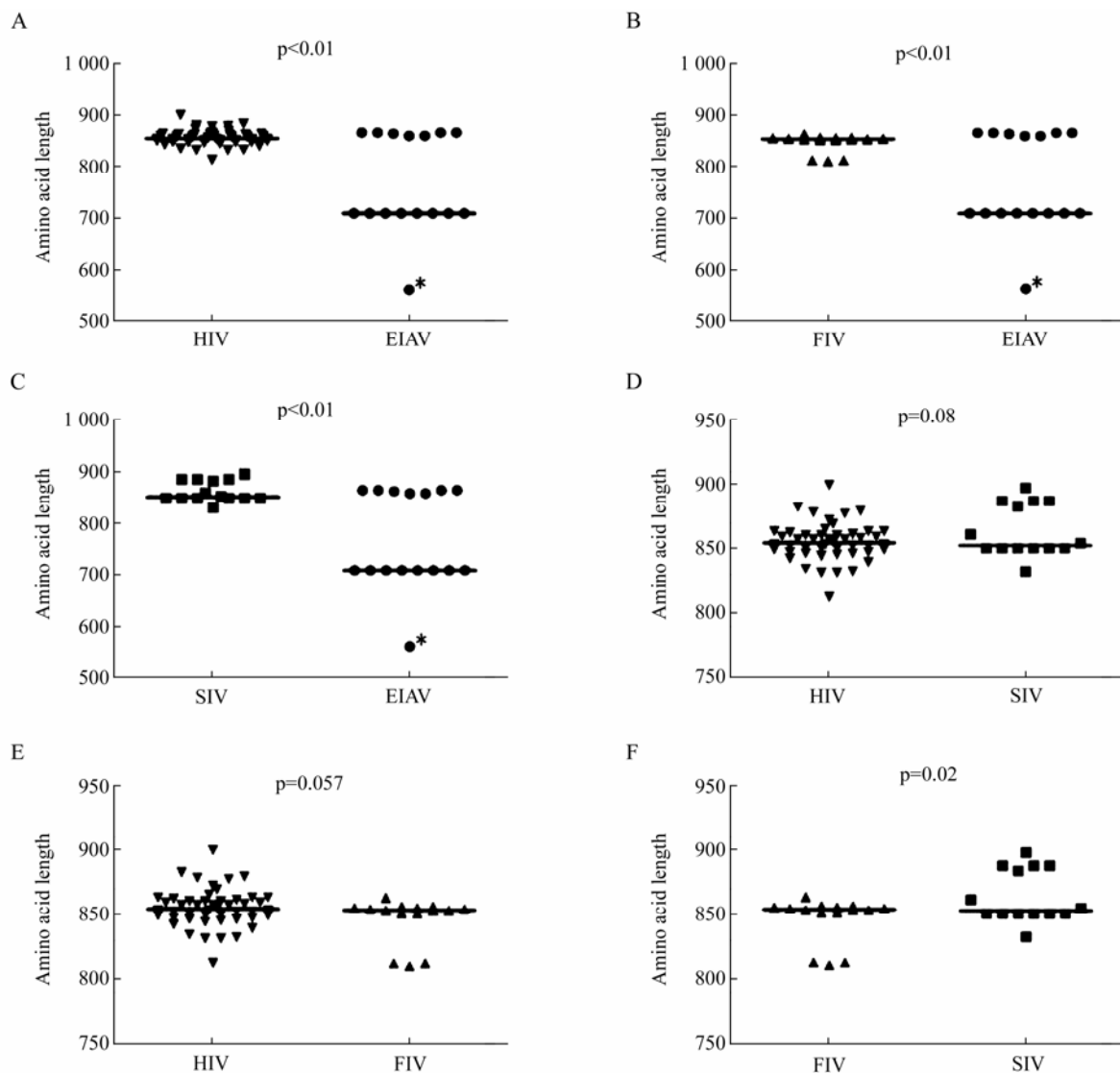Median values for PNGS, length and relative glycosylation density of Env are shown in the table.

Fig. 2. Comparison of Env amino acid length of HIV, SIV, FIV and EIAV. A: Comparison of Env amino acid length between HIV and EIAV. The outlier is labeled with the asterisk (*). B: Comparison of Env amino acid length between FIV and EIAV. C: Comparison of Env amino acid length between SIV and EIAV. D: Comparison of Env amino acid length between SIV and HIV. E: Comparison of Env amino acid length between FIV and HIV. F: Comparison of Env amino acid length between FIV and SIV. The p value for each pair is shown.
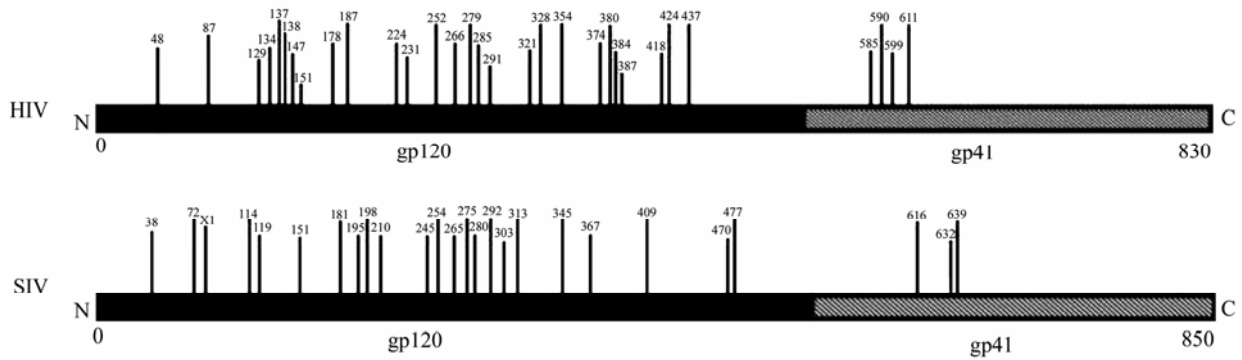
Fig. 3. Potential N-linked glycosylation sites distributed in the extracellular membrane gp120 is more than the transmembrane gp41 for HIV and SIV. (Vertical bars denoted potential N-linked glycosylation sites and the numbers above the vertical bars represented the amino acid positions of N-linked glycosylation sites.)
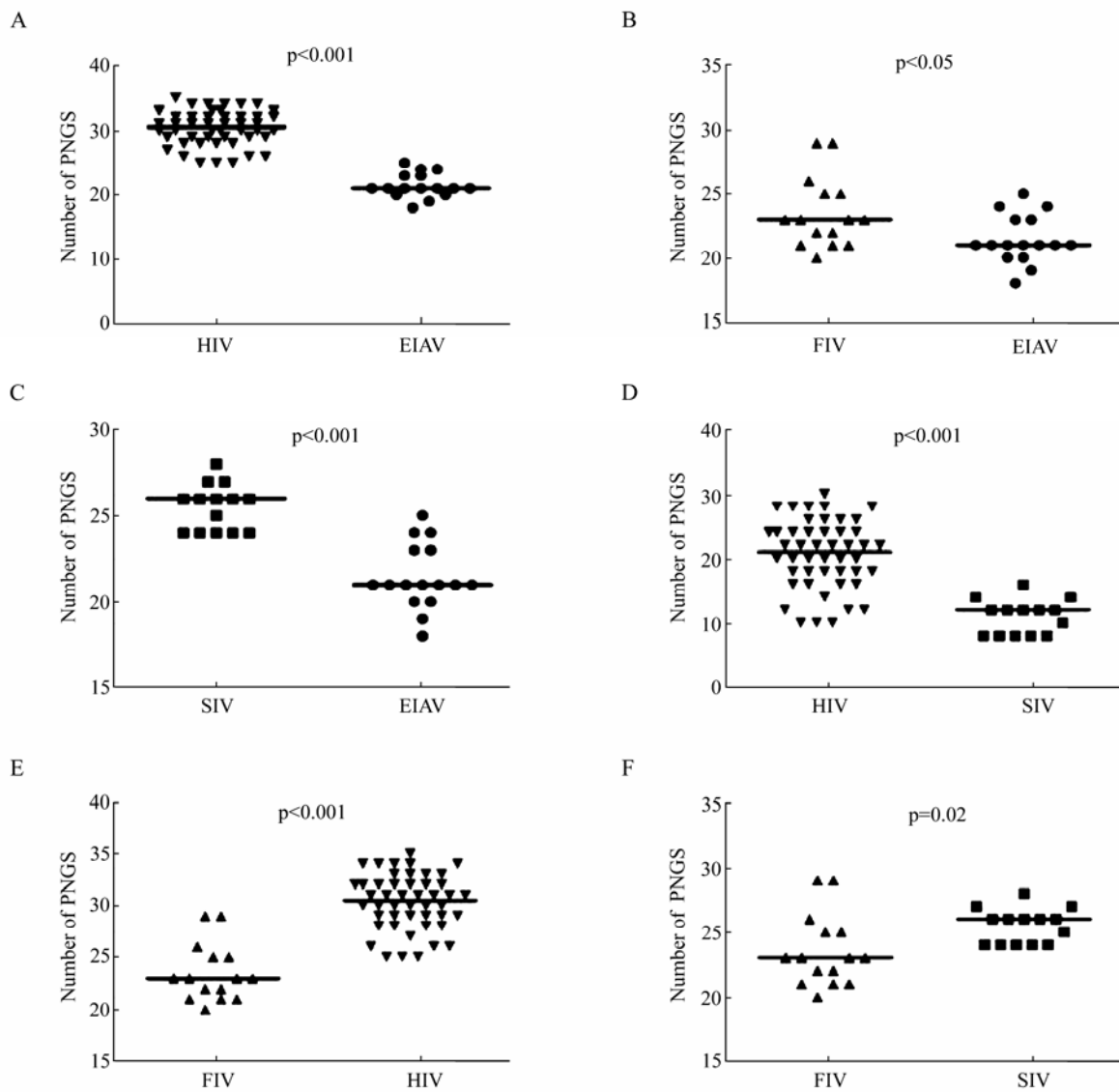

Fig. 4. Comparison of PNGS in Env of HIV, SIV, FIV and EIAV. A: Comparison of PNGS in Env between EIAV and HIV. B: Comparison of PNGS in Env between EIAV and FIV. C: Comparison of PNGS in Env between EIAV and SIV. D: Comparison of PNGS in Env between HIV and SIV. E: Comparison of PNGS in Env between HIV and FIV. F: Comparison of PNGS in Env between SIV and FIV. The *p* value for each pair is shown.
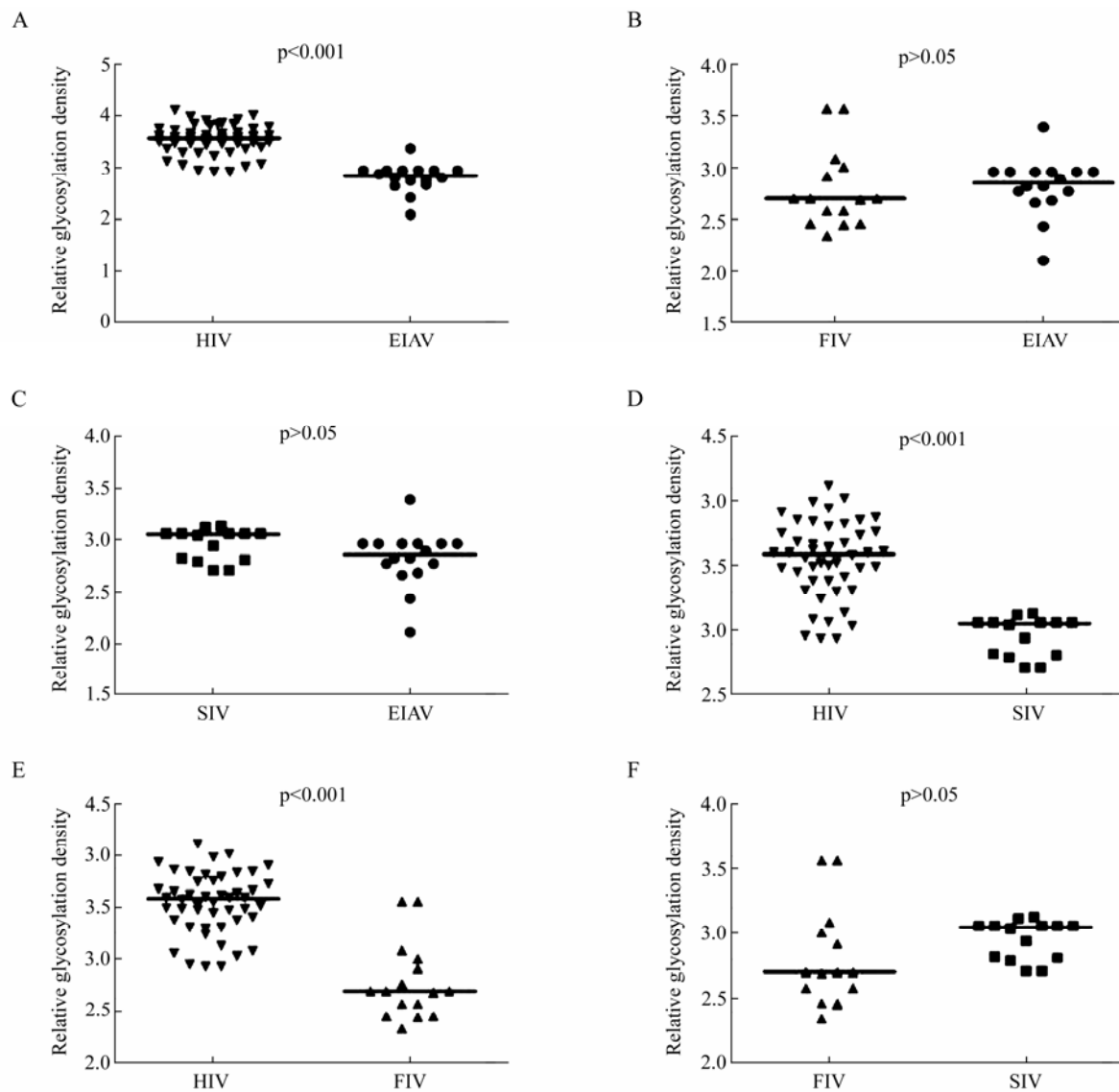
Fig. 5. Comparison of relative glycosylation density of potential N-linked glycosylation sites in Env of HIV, SIV, FIV and EIAV. A: Comparison of relative glycosylation density in Env between EIAV and HIV. B: Comparison of relative glycosylation density in Env between EIAV and FIV. C: Comparison of relative glycosylation density in Env between EIAV and SIV. D: Comparison of relative glycosylation density in Env between HIV and SIV. E: Comparison of relative glycosylation density in Env between HIV and FIV. F: Comparison of relative glycosylation density in Env between SIV and FIV. The *p* value for each pair is shown.

The comparison of the total number and relative glycosylation density of potential N-linked glycosylation sites of Env between various immunodeficiency viruses and EIAV

The glycosylation of Env is also an important factor which affects the function of Env, such as immunogenicity, cell tropism and the replication capacity of the virus. We first compared the consensus Env of SIV and HIV, and we found that the number of PNGS located on the gp120 was more than that on gp41 (Fig. 3). With regard to the number of PNGS, among the different viruses analyzed, the HIV Env exhibited the highest number of PNGS, followed by SIV and FIV. Interestingly, among all the viruses analyzed, EIAV had lowest number of PNGS ($p < 0.05$) (Table 1 and Fig. 4A, 4B and 4C). With regard to the number of PNGS among the three immunodeficiency viruses (HIV, SIV and FIV),

HIV was the most and FIV was the least. There was a reduced number of PNGS in SIV as compared to HIV, which was statistically significant ($p < 0.001$). Similarly, the number of PNGS in FIV Env was less than that observed in HIV and SIV, and the difference was also significant with HIV ($p < 0.001$) but not with SIV ($p = 0.02$) (Fig. 4D, 4E and 4F).

Our results showed that Env quasispecies variety of viruses in Fig. 2. We further compared the relative glycosylation density as normalized to Env amino acids length. Consistent with the total number of glycosylation sites analyzed, the density of HIV glycosylation was the highest among the various lentiviruses analyzed; much higher than those observed with EIAV, with the glycosylation density ranging from 2.10 to 3.39, with SIV or with FIV ($p < 0.01$) (Table 1 and Fig. 5A, 5D, 5E). A comparison

of the other viruses showed that the relative glycosylation density of SIV was slightly higher than that of FIV and EIAV, but there was no significant difference (p>0.05) (Fig. 5C, 5F).

DISCUSSION

Although the clinical manifestations of infections by immunodeficiency viruses and EIAV are different, the underlying pathogenesis are very similar[4]. The success of EIAV vaccination might facilitate the research on effective HIV vaccine design. As a basis toward defining the role of this protein in the pathogenesis, we studied envelope genotype phylogeny of the lentiviruses, including the various immunodeficiency viruses and EIAV. The phylogenetic traits of those viruses analyses included the *env* gene evolutionary distance, amino acid length and the number of PNGS.

It is well established that virus evolution is selected for both the host and the virus itself, and Env is more likely to mutate in the variable regions than other viral genes. Therefore, the evolutionary distance of Env will reflect evolutionary relationship among species of viruses. Our results show that the Env protein in HIV was most closely related to SIV and the greatest distance was observed between EIAV and SIV.

The Env of retroviruses have a number of common features. However, the envelope protein of lentiviruses is significantly larger than that of other retroviruses. The size difference suggests that there may be a lentivirus-specific function associated with Env[22]. Among various immunodeficiency viruses and EIAV analyzed, the Env length of EIAV was the shortest and HIV was the longest. It suggests that the immunodeficiency viruses are more complex than EIAV with respect to the Env amino acid length, and it is possible that such complexity of the Env amino acid sequence length could lead to more sequence variation and diversification for the HIV Env, which is a major obstacle in developing an effective HIV vaccine.

The N-linked glycosylation sites influence the envelope protein to fold correctly and exert the normal functions. More important, the deglycosylation of the envelope proteins is closely related to an attenuated vaccine[18]. Our study has demonstrated that the number of PNGS in immunodeficiency viruses of primates (human and simian) is more than that of feline and equine. Because of the Env protein exists as quasispecies, the comparison of PNGS was normalized to Env amino acid length in order to compare the density of glycosylation. The relative glycosylation density of HIV and SIV was higher than that of FIV and EIAV, especially for HIV. We found that the difference of PNGS between HIV and SIV was mainly on gp120. There have been reports that the surface glycoprotein for the virulence was stronger than transmembrane glycoprotein[20]. Previous studies have shown that the heavy glycosylation of Env is an unique feature of HIV/SIV that is distinct from other enveloped viruses, and it is significantly related to their neutralization resistant

properties[2,14,21]. This study showed that the number of PNGS and the density of glycosylation is also the lowest after normalization with the Env sequence length. Therefore, we speculate that the effective EIAV attenuated vaccine benefits from the less number of PNGS in Env and the less relative glycosylation density. In agreement with, the insufficient immunosuppression of HIV/SIV might be due to heavy glycosylation sites in Env and remove some of the glycans could enable the host to mount the stronger protective immune response against the infection[18]. Currently, the heavily glycosylated structure of Env has been considered as a main contributor to chronic and persistent viral replication, and the pathogenicity of HIV or SIV, primarily because it potentially interferes with the development of an effective host immune response[6,17,18].

Our results showed that there was a significant difference in the number of PNGS between HIV and SIV in spite of similar amino acid length of Env. The number of PNGS in SIV was significantly less than that of HIV. The failure of HIV vaccine in clinical trial might be explained by the difference of PNGS in Env between simian and human. The quasispecies of Env as a result of amino acid length and the number of PNGS variable are an obstacle for successful HIV/AIDS vaccine. It is possible that the successful EIAV Env attenuated vaccine strains possess an advantage over other immunodeficiency viruses because EIAV Env sequences have shorter amino acid length and less number of potential N-linked glycosylation sites (PNGS). The modification of glycosylation of HIV vaccine might be different between human and simian, even though they had been tested effective in simian. In addition, the glycosylation sites are known for play an important role in viral sensitivity to neutralization, pathogenicity and the cell tropism[8].

References

1. **Bogers W M, Cheng-Mayer C, Montelaro R C.** 2000. Developments in preclinical aids vaccine efficacy models. **AIDS**, 14 Suppl 3: S141-S151.

2. **Burton D R, Desrosiers R C, Doms R W, et al.** 2004. Hiv vaccine design and the neutralizing antibody problem. **Nat Immunol**, 5(3): 233-236.

3. **Cavarelli M, Karlsson I, Zanchetta M, et al.** 2008. Hiv-1 with multiple ccr5/cxcr4 chimeric receptor use is predictive of immunological failure in infected children. **PLoS One**, 3(9): e3292.

4. **Clements J E, Zink M C.** 1996. Molecular biology and pathogenesis of animal lentivirus infections. **Clin Microbiol Rev**,

9(1): 100-117.

5. **Craigo J K, Zhang B, Barnes S***, et al.* 2007**.** Envelope variation as a primary determinant of lentiviral vaccine efficacy. **Proc Natl Acad Sci U S A**, 104(38): 15105-15110.

6. **Desrosiers R C.** 2004**.** Prospects for an aids vaccine. **Nat Med**, 10(3): 221-223.

7. **Felsenstein J.** 1996**.** Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. **Methods Enzymol**, 266: 418-427.

8. **Johnson W E, Sanford H, Schwall L***, et al.* 2003**.** Assorted mutations in the envelope gene of simian immunodeficiency virus lead to loss of neutralization resistance against antibodies representing a broad spectrum of specificities. **J Virol**, 77(18): 9993-10003.

9. **Kuiken C, Korber B, Shafer R W.** 2003**.** Hiv sequence databases. **AIDS Rev**, 5(1): 52-61.

10. **Leroux C, Cadore J L, Montelaro R C.** 2004**.** Equine infectious anemia virus (eiav): What has hiv's country cousin got to tell us? **Vet Res**, 35(4): 485-512.

11. **Liang H, He X, Shen R X***, et al.* 2006**.** Combined amino acid mutations occurring in the envelope closely correlate with pathogenicity of eiav. **Arch Virol**, 151(7): 1387-1403.

12. **Losman B, Bolmstedt A, Schonning K***, et al.* 2001**.** Protection of neutralization epitopes in the v3 loop of oligomeric human immunodeficiency virus type 1 glycoprotein 120 by n-linked oligosaccharides in the v1 region. **AIDS Res Hum Retroviruses**, 17(11): 1067-1076.

13. **Losman B, Biller M, Olofsson S***, et al.* 1999**.** The n-linked glycan of the v3 region of hiv-1 gp120 and cxcr4-dependent multiplication of a human immunodeficiency virus type 1 lymphocyte-tropic variant. **FEBS Lett**, 454(1-2): 47-52.

14. **Means R E, Greenough T, Desrosiers R C.** 1997**.** Neutralization sensitivity of cell culture-passaged simian immunodeficiency virus. **J Virol**, 71(10): 7895-7902.

15. **Mellquist J L, Bowman B, Kasturi L***, et al.* 1998**.** Characterization of hiv type 1 gp120 v3 region sequences from ugandan infants. **AIDS Res Hum Retroviruses**, 14(15): 1391-1395.

16. **Miyauchi K, Curran A R, Long Y***, et al.* The membrane-spanning domain of gp41 plays a critical role in intracellular trafficking of the hiv envelope protein. **Retrovirology**, 7: 95.

17. **Mori K, Yasutomi Y, Ohgimoto S***, et al.* 2001**.** Quintuple deglycosylation mutant of simian immunodeficiency virus sivmac239 in rhesus macaques: Robust primary replication, tightly contained chronic infection, and elicitation of potent immunity against the parental wild-type strain. **J Virol**, 75(9): 4023-4028.

18. **Mori K, Sugimoto C, Ohgimoto S***, et al.* 2005**.** Influence of glycosylation on the efficacy of an env-based vaccine against simian immunodeficiency virus sivmac239 in a macaque aids model. **J Virol**, 79(16): 10386-10396.

19. **Ohgimoto S, Shioda T, Mori K***, et al.* 1998**.** Location-specific, unequal contribution of the n glycans in simian immunodeficiency virus gp120 to viral infectivity and removal of multiple glycans without disturbing infectivity. **J Virol**, 72(10): 8365-8370.

20. **Payne S L, Pei X F, Jia B***, et al.* 2004**.** Influence of long terminal repeat and env on the virulence phenotype of equine infectious anemia virus. **J Virol**, 78(5): 2478-2485.

21. **Reitter J N, Means R E, Desrosiers R C.** 1998**.** A role for carbohydrates in immune evasion in aids. **Nat Med**, 4(6): 679-684.

22. **Rice N R, Henderson L E, Sowder R C***, et al.* 1990**.** Synthesis and processing of the transmembrane envelope protein of equine infectious anemia virus. **J Virol**, 64(8): 3770-3778.

23. **Schonning K, Jansson B, Olofsson S***, et al.* 1996**.** Rapid selection for an n-linked oligosaccharide by monoclonal antibodies directed against the v3 loop of human immunodeficiency virus type 1. **J Gen Virol**, 77 (Pt 4): 753-758.

24. **Sellon D C, Fuller F J, McGuire T C.** 1994**.** The immunopathogenesis of equine infectious anemia virus. **Virus Res**, 32(2): 111-138.

25. **Shen R X, Xu Z D, HE Y S***, et al.* 1979**.** Development of a eiav donkey leucocyte attenuated vaccine. **Sci Agr Sin,** 4: 1-15. (in Chinese)

26. **Shen T, Liang H, Tong X***, et al.* 2006**.** Amino acid mutations of the infectious clone from chinese eiav attenuated vaccine resulted in reversion of virulence. **Vaccine**, 24(6): 738-749.

27. **Sjolander S, Bolmstedt A, Akerblom L***, et al.* 1996**.** N-linked glycans in the cd4-binding domain of human immunodeficiency virus type 1 envelope glycoprotein gp160 are essential for the in vivo priming of t cells recognizing an epitope located in their vicinity. **Virology**, 215(2): 124-133.

28. **Sober E.** 2004**.** The contest between parsimony and likelihood. **Syst Biol**, 53(4): 644-653.

29. **Thompson J D, Gibson T J, Plewniak F***, et al.* 1997**.** The clustal_x windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res**, 25(24): 4876-4882.

30. **Tully D C, Wood C.** 2010. Chronology and evolution of the hiv-1 subtype c epidemic in ethiopia. **AIDS**, 24(10): 1577-1582.

31. **Zhang H, Tully D C, Hoffmann F G***, et al.* 2010**.** Restricted genetic diversity of hiv-1 subtype c envelope glycoprotein from perinatally infected zambian infants. **PLoS One**, 5(2): e9294.

32. **Zhang M, Gaschen B, Blay W***, et al.* 2004**.** Tracking global patterns of n-linked glycosylation site variation in highly variable viral glycoproteins: Hiv, siv, and hcv envelopes and influenza hemagglutinin. **Glycobiology**, 14(12): 1229-1246.