The Journal of Microbiology, February 2009, p. 85-90 Vol. 47, No. 1
DOI 10.1007/s12275-008-0240-x

Timing and Evolution of the Most Recent Common Ancestor of the Korean Clade HIV Subtype B Based on *Nef* and *Vif* Sequences

Mi-Suk Kim[1], So-Young Jang[1], Chan-Seung Park[1], Keon-Myung Lee[2], Dong-Hun Lee[1], and Chan-Hee Lee[1]*

[1]School of Life Sciences, and [2]School of Electrical and Computer Engineering, Chungbuk National University, Chungbuk 361-763, Republic of Korea

(Received September 25, 2008 / Accepted December 5, 2008)

Molecular phylogenetic studies of the HIV-1 isolated from Koreans have suggested the presence of the so-called "Korean clade", which can be defined as a cluster free of foreign isolates. The Korean clade accounts for more than 60% of Korean isolates and exerts characteristic amino acid sequences. Thus, it is merited to estimate when this Korean clade first emerged in order to understand the evolutionary pattern of the Korean clade. We analyzed and reconstructed the most recent common ancestor (MRCA) sequences from *nef* (n=229) and *vif* (n=179) Korean clade sequences. Linear regression analyses of sequence divergence estimates were plotted against sampling years to infer the year in which there was zero divergence from the MRCA sequences. MRCA sequences suggested the Korean clade was first emerged around 1984, before the first detection of HIV-1 in Korea in 1985. Further studies on synonymous and nonsynonymous substitution rates suggested positive selection event for the Korean clade, while other subtype B had undergone negative to neutral evolution.

***Keywords*:** HIV-1, Korean clade, common ancestor

Human immunodeficiency virus (HIV) is the causative agent for AIDS (acquired immunodeficiency syndrome), first identified in 1981. Since then AIDS became one of the leading cause of mortality around the world. Although the incidence of HIV infection in Korea is relatively low compared to other countries, public health authorities are worried about the rapid increase in the number of new HIV infection in recent years. HIV infection in Korea was first detected in 1984 from foreigner, and the first Korean diagnosed as HIV-positive in 1985 was infected abroad. Major source of HIV infection in Korea during the 1980's was from abroad, either from foreigners or Koreans infected in foreign countries. Since then, domestic infection through sexual contact, whether homosexual or heterosexual, appeared to be the major source of HIV infection in Korea.

Molecular epidemiologic studies using gene sequences of HIV isolated from Koreans identified that HIV-1 subtype B accounted for ~80% of all Korean isolates (Lee *et al*., 2003; Park *et al*., 2006). Furthermore, majority of the HIV-1 subtype B isolated from Koreans seem to be clustered together, forming a "Korean clade" where foreign isolates are completely excluded (Kang *et al*., 1998). The presence of the Korean clade has been supported by further studies based on molecular phylogenetic analyses of *pol* (Sung *et al*., 2001), *nef* (Kang *et al*., 1998; Lee *et al*., 2003), *env* (Kim *et al*., 1999a, 1999b), and *vif* (Park *et al*., 2006) genes. The Korean clade, a subcluster of HIV-1 subtype B, accounted for more than 60% of all HIV isolated from Koreans and exhibited unique nucleotide and amino acid sequences that distinguished the Korean clade from others (Park *et al*., 2006, 2008)

HIV-1 is classified into three major groups, M, N, and O, and group M is further divided into several subtypes whose regional distribution is well known. Using the most recent common ancestor (MRCA) sequences inferred from known sampling date and assuming a constant rate of evolution, it was suggested that M group was first established around 1931 (Korber *et al*., 2000). Emergence of subtypes from M group ancestor is thought to have occurred later, late 1960s to early 1970s for subtype B (Korber *et al*., 2000; Salemi *et al*., 2001; Lukashov and Goudsmit, 2002) and mid to late 1960s for subtype C (Travers *et al*., 2004). More studies have established the date when the MRCA of HIV first appeared in specific regions worldwide. However, there has not been such attempt for HIV epidemic in Korea. Considering the importance and uniqueness of the Korean clade, we attempted in this study to understand the time when the Korean clade was first emerged by inferring MRCA sequences of *nef* and *vif* genes of HIV-1 isolated from Korean patients. Furthermore, evolutionary consideration was attempted by calculating the synonymous and nonsynonymous distances.

Materials and Methods

Sequence information and phylogeny

All available nucleotide sequences of *nef* (n=422) and *vif* (n=233) genes of HIV-1 isolated from Korean patients were retrieved from NCBI GenBank database as described before (Park *et al*., 2006, 2008). Relevant foreign sequences with more than 90% homology with Korean isolates were obtained from NCBI GenBank database by BLAST search and reference sequences were obtained from Los Alamos

* To whom correspondence should be addressed.
(Tel) 82-43-261-2304; (Fax) 82-43-273-2451
(E-mail) chlee@chu.ac.kr

HIV Database (http://hiv-web.lanl.gov). Ninety-four foreign and 41 reference sequences for *nef* gene, and 65 foreign and 44 reference sequences for *vif* gene were included in this study. Sequences were multiple-aligned with CLUSTAL X (ver. 1.83, Thomson *et al*., 1997) followed by manual editing. The resulting out-files were used for construction of phylogenetic trees by neighbor-joining and maximum-likelihood methods using PHYLIP package (version 3.6, http://evolution.genetics.washington.edu/phylip.html). Bayesian (BA) trees were constructed using MrBayes program (version 3.0, http://mrbayes.csit.fsu.edu/).

Reconstruction of the most recent common ancestor (MRCA) sequences

The MRCA of the Korean clade subtype B (K_CB) were inferred using maximum likelihood (ML) or Bayesian (BA) method. Dnamlk program in PHYLIP package (ver. 3.6a, Felsenstein, 1993) with the options of transition/transversion ratio=2.0 and no rate variation among sites was used for ML-based reconstruction of the MRCA. Other parameters were set as defaults. The root sequences thus generated were the reconstructed MRCA sequences. MrBayes program was used for MRCA reconstruction by heuristic Bayesian method. Bayesian method does not allow us to get the sequences of the root, which is the MRCA, inclusion of an outgroup sequence is required. In this study, subtype B reference HXB2 sequences were included as an outgroup. If more than two ancestor sequences were generated by MrBayes, the one with the highest likelihood value was chosen as the MRCA. The resulting MRCA sequences were aligned according to HXB2 sequences.

Dating method

Individual sequences with known sampling years between 1991 and 2004 were aligned with MRCA sequences. And the genetic divergence of the K_CB sequences from the MRCA sequences were estimated by calculating synonymous distances (dS) using codeml program from the PAML (Phylogenetic Analysis Maximum Likelihood) package (Yang, 1997). Correlation between the divergence values and the sampling years was estimated with linear regression analysis using SigmaPlot (ver. 8.0) by plotting sampling year at x-axis and dS at y-axis. The resulting graphs contain the following information: regression line between time and distance (solid line), 95% confidence intervals (dotted line), regression coefficient (R), and slope of the regression line. The year at which the regression line crossed the x-axis (x-intercept) was regarded as the year when the MRCA emerged.

Determination of synonymous and nonsynonymous distances

Nucleotide sequences were multiple-aligned with CLUSTAL X. Aligned sequences were compared with SeqAid program developed by us to get genetic distances among sequences belonging to different groups. Then, mean and standard deviation values for each pair of comparison were calculated. Statistical analysis was performed with SPSS (ver. 10) in order to get statistical significance of the data when needed. For determination of synonymous and nonsynonymous distances, aligned sequences and ML trees were analyzed using the codonml program from the PAML package. Three maximum likelihood models were used. Invariant model assumes that all codons fall into a single category of sites with a fixed value of ω, ratio of nonsynonymous to synonymous distances (Goldman and Yang, 1994). Neutral model allows two categories of sites, one with fixed value of $\omega1=1$ and another with negative selection sites ($\omega2=0$) (Nielsen and Yang, 1998). Positive selection model incorporates an additional category of positively selected sites where $\omega3>1$ (Nielsen and Yang, 1998).

Results

Timing the emergence of the Korean clade

In order to estimate when the Korean clade first emerged, the MRCA sequences were reconstructed from 229 *nef* and 179 *vif* gene sequences of the Korean clade of subtype B (K_CB) with known sampling date (Table 1). MRCA sequences were inferred by Bayesian (BA) or maximum likelihood (ML) method, and the genetic distances were calculated as synonymous distances from the MRCA sequences to each of the K_CB sequences. Linear regression analyses from *nef* sequences indicated a date of 1984 for the timing of the origin of the K_CB by both BA and ML methods (Fig. 1). The 95% confidence intervals (CI) ranged from 1972 to 1989. There was a significant correlation between genetic divergence and time ($P<0.005$). Similar results were obtained with *vif* gene data, although BA method estimated 1983 as the emergence time of the K_CB (Fig. 2). The 95% CIs ranged from 1960 to 1989. The rates of genetic divergence for *nef* and *vif* genes were 0.0037 and 0.0015, consistent with higher mutation rate of *nef* than *vif* gene.

Table 1. Yearly distribution of the number of sequences analyzed in this study

Isolation year	Subtype B			
	Korean clade		Non-Korean clade	
	Nef	Vif	Nef	Vif
1991	4	5	0	0
1992	0	0	0	0
1993	11	4	7	2
1994	9	11	5	1
2002	6	10	1	3
2002	22	6	3	0
1997	8	14	2	3
1998	0	0	0	0
1999	15	6	5	3
2000	25	2	5	0
2001	24	19	8	0
2002	48	16	9	3
2003	51	69	10	13
2004	6	17	8	9
Not known	35	1	8	0
Total	264	180	71	37

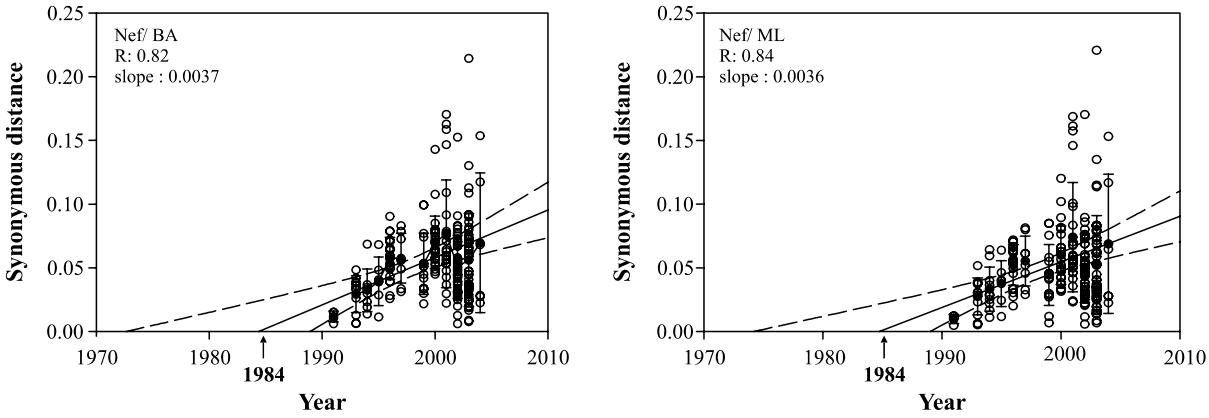

Fig. 1. Synonymous distances of the Korean clade *nef* sequences to the reconstructed most recent common ancestor sequence in relation to sampling years. The regression analysis was performed for 229 individual sequences (open circles) as well as for the mean synonymous distances per sampling year (closed circles with error bars). Regression lines (solid line) and 95% confidence intervals (dotted line) are also shown. Left, Bayesian (BA) method; Right, maximum-likelihood (ML) method.

Evolutionary consideration

In order to understand whether the Korean clade evolved similarly or not with non-Korean clade (NK_CB), we analyzed synonymous and nonsynonymous distances among all nucleotide sequences belonged to K_CB or NK_CB. We included sequences with unknown sampling date since the information of sampling date was not necessary (Table 1). Three maximum likelihood models, invariant, neutral, and positive selection, were compared to find which provided the best fit to the data. Likelihood ratio tests (LRT) among the three models were performed by χ^2-test with degree of freedom 2. As shown in Table 2, positive selection model had better fits than invariant or neutral model to the K_CB or NK_CB data of *nef* gene ($P<0.001$), and the data were analyzed according to positive selection model. The mean synonymous distances (dS) and nonsynonymous distances (dN) for K_CB were 0.63% and 0.75%, respectively. The mean dS and dN for NK_CB were 1.13% and 1.05%, respectively. The ratio of dN/dS (ω) was 1.19 for K_CB and 0.93 for NK_CB (Table 2). Similar results were obtained when *vif* gene data were analyzed. Positive selection model was again the best fit among the three models and estimated the dN/dS ratio 1.17 for K_CB and 0.66 for NK_CB (Table 2). Positive selection model also predicts the proportion of neutral codons (p1), negatively selected codons (p2), and positively selected codons (p3). Comparison of K_CB and NK_CB indicates that p2 was higher in NK_CB while p1 and p3 were higher in K_CB (Table 2). It seems that the Korean clade had undergone slightly positive evolution compared with non-Korean clade subtype B who had undergone slightly negative evolution.

Discussion

In this study we suggest that the Korean clade first emerged around the year 1984, before the first detection of HIV-1

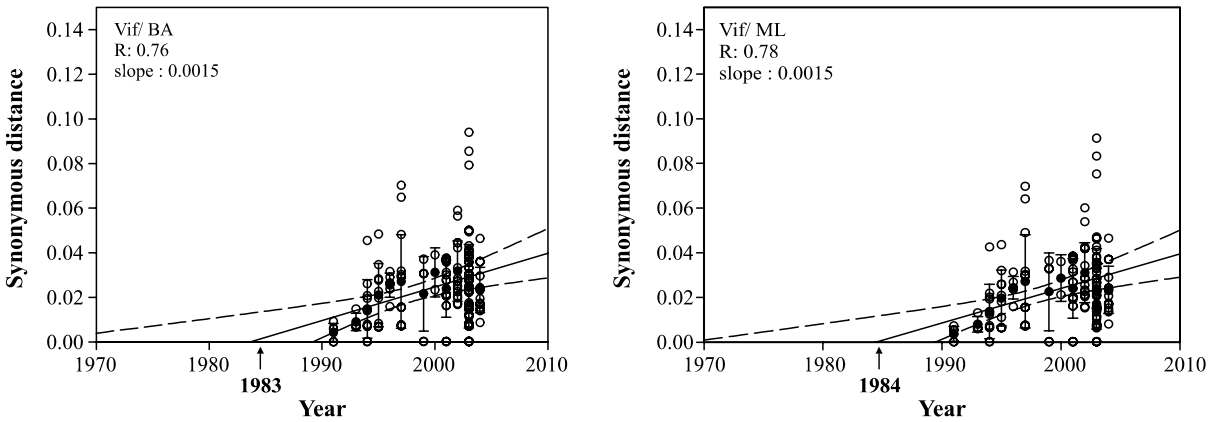

Fig. 2. Synonymous distances of the Korean clade *vif* sequences to the reconstructed most recent common ancestor sequence in relation to sampling years. The regression analysis was performed for all 179 sequences (open circles) as well as for the mean synonymous distances per sampling year (closed circles with error bars). Regression lines (solid line) and 95% confidence intervals (dotted line) are also shown. Left, Bayesian (BA) method; Right, maximum-likelihood (ML) method.

Table 2. Maximum likelihood estimates of selection pressures on HIV-1 *nef*, *vif*, and *pol* sequences from K_CB and NK_CB

Model	Taxa	Nef		Vif	
		K_CB	NK_CB	K_CB	NK_CB
		264	71	180	37
(1) Invariant	κ	4.37	3.01	4.07	3.80
	ω	0.57	0.59	0.86	0.50
	ln *L*	-5403.70	-4856.50	-6571.75	-2490.35
(2) Neutral	κ	4.61	3.08	3.48	3.69
	p1	0.79	0.64	0.57	0.44
	p2	0.21	0.36	0.43	0.56
	ω	0.79	0.64	0.57	0.44
	ln *L*	-5336.33	-4762.57	-6402.33	-2440.03
(3) Positive Selection	κ	4.93	3.37	4.21	4.12
	p1	0.64	0.54	0.48	0.41
	p2	0.21	0.35	0.40	0.53
	p3	0.14	0.11	0.13	0.05
	ω3	3.76	3.61	5.43	4.90
	ω	**1.19**	**0.93**	**1.17**	**0.66**
	ln *L*	-5234.47	-4718.29	-6210.94	-2422.43
LRT (3&1)	LR	338.46	276.42	721.62	135.84
	P value	<0.001	<0.001	<0.001	<0.001
LRT (3&2)	LR	203.72	88.67	382.78	35.20
	P value	<0.001	<0.001	<0.001	<0.001

κ, Transition/transversion ratio
p1, Proportion of neutral codons
p2, Proportion of negatively selected codons
p3, Proportion of positively selected codons
ω3, dN/dS ratio at p3 sites
ω, overall dN/dS ratio
ln *L*, Maximum likelihood ratio
LRT, Likelihood Ratio Test

in Korea. The most recent common ancestor (MRCA) of the Korean clade might have appeared abroad during that time and entered Korea no later than 1991 at which time the first data on the Korean clade was available.

Different studies with different genes or sequence data sets estimated the date of the origin of the HIV-1 subtype B to range from 1967 to 1976 (Korber *et al*., 2000; Salemi *et al*., 2001; Lukashov and Goudsmit, 2002). The significantly different estimates provoked a controversy between the authors (Lukashov and Goudsmit, 2003; Smith *et al.,* 2003). Even with the same data set, different analytical methods estimated different dates (Lukashov and Goudsmit, 2002). Using two genes (*nef* and *vif*) and two methods (BA and ML) for each gene, we were able to obtain very similar results, may be due to high similarity of the K_CB sequences compared to other subtype B sequences (Park *et al*., 2006, 2008). The Korean clade diversified 8 to 17 years after the emergence of the subtype B.

The diversification of the Korean clade from the subtype B may be viewed as a speciation event. And once separated, the Korean clade appeared to have undergone different evolutionary pathway from the other subtype B and the hallmark of the evolutionary pathway seems to be the K_CB-specific signature amino acid pattern (Park *et al*., 2006, 2008). HIV-1 *nef* gene encodes a transactivating factor that may reduce or increase viral replication depending on cell type (Welker *et al*., 1996; Levy, 1998). The vif protein protects HIV-1 from antiretroviral protein human APOBEC3G (hA3G)-mediated inhibition of HIV-1 replication by recruiting an E3-ubiquitin ligase to hA3G (Sheehy *et al*., 2002; Yu *et al.,* 2003; Farrow and Sheehy, 2008). However, the K_CB-specific signature amino acid patterns revealed that the amino acid residues critical for nef or vif protein action were not altered in K_CB. Thus, the evolution of K_CB does not involve the change in the critical functions of nef or vif protein.

Rather the evolution of the Korean clade might be due to escape from cytotoxic T lymphocyte (CTL) response (Price *et al*., 1997; McMichael and Rowland-Jones, 2001). CTL escape mutations could be positively selected within an individual host, and would be characteristic for specific HLA class I alleles across an HLA-diverse host population (Moore *et al*., 2002). Previously we analyzed substitution events in the K_CB and some of the nonsynonymous substitutions were identified to constitute signature amino acid residues whose frequencies were significantly higher among K_CB (Park *et al*., 2006, 2007, 2008). Some of the signature amino acids of the K_CB nef could be located to HLA-restricted CTL epitope motifs where the amino acids are enriched in the presence of a specific HLA allele thus presumably represent escape variants (Brumme *et al*., 2007). Koreans are relatively pure in ethnicity and exhibit certain HLA class I alleles more frequently than other alleles (Lee *et al*., 2005). Thus, the ancestor virus of the Korean clade might have evolved characteristically by adaptation to HLA-restricted CTL escape while circulating among the Korean population. Further and extensive analysis of HLA class I alleles of Korean HIV-1 infected individuals would give some idea on the origin and evolution of the Korean clade.

Positive selection could be inferred by a higher rate of nonsynonymous (dN) to synonymous (dS) substitution per site (Hughes and Nei, 1989; Mindell, 1996; Sharp, 1997). The rate of nonsynonymous substitutions might increase after the cross-species transmission of a pathogen, reflecting adaptation to the new hosts, while synonymous substitutions are expected to be largely neutral, and reflect the underlying rate of mutation and replication (Sharp *et al*., 2001). We estimated dN/dS ratio (ω) according to positive selection model (Nielsen and Yang, 1998; Zanotto *et al*., 1999), and it could be suggested that the *nef* and *vif* genes of the K_CB have undergone positive selection. On the other hand, there was little evidence suggesting adaptive evolution (either positive or negative) for the *nef* gene of the NK_CB, while the *vif* gene of the NK_CB seemed to have undergone negative selection.

Lastly a comment should be made about the limit about the data acquisition in this study. Although all Korean *nef* and *vif* sequences from NCBI GenBank were included available at the time this study was initiated, they did not include all Korean *nef* or *vif* sequences. However, it is not possible to acquire all Korean sequences since not all viruses were isolated from patients and sequenced, and, if sequenced, some of the sequenced data may not be reported to NCBI GenBank database. Despite this limitation in data acquisi-

tion, we believe that the conclusion drawn in this study may be the best we can get since we used all the sequences available to open public. Based on the sequence data and analytical methods we used in this study, it is concluded that the Korean clade of HIV-1 appeared around the year 1984 and to have undergone characteristic evolution by positive selection.

Acknowledgements

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (Ministry of Education, Science and Technology). (The Regional Research Universities Program/Chungbuk BIT Research-Oriented University Consortium)

References

Brumme, Z.L., C.J. Brumme, D. Heckerman, B.T. Korber, M. Daniels, J. Carlson, C. Kadie, T. Bhattacharya, C. Chui, J. Szinger, T. Mo, R.S. Hogg, J.S. Montaner, N. Frahm, C. Brander, B.D. Walker, and P.R. Harrigan. 2007. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. *PLoS Pathog.* 3, e94.

Farrow, M.A. and A.M. Sheehy. 2008. Vif and Apobec3G in the innate immune response to HIV: a tale of two proteins. *Future Microbiol.* 3, 145-154.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genetics, University of Washington, Seattle, USA.

Goldman, N. and Z. Yang. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* 11, 725-736.

Hughes, A.L. and M. Nei. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc. Natl. Acad. Sci. USA* 86, 958-962.

Kang, M.R., Y.K. Cho, J. Chun, Y.B. Kim, I. Lee, H.J. Lee, S.H. Kim, Y.K. Kim, K. Yoon, J.M. Yang, J.M. Kim, Y.O. Shin, C. Kang, J.S. Lee, K.W. Choi, D.G. Kim, W.M. Fitch, and S. Kim. 1998. Phylogenetic analysis of the *nef* gene reveals a distinctive monophyletic clade in Korean HIV-1 cases. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol*. 17, 58-68.

Kim, Y.B., Y.K. Cho, H.J. Lee, C.K. Kim, Y.K. Kim, and J.M. Yang. 1999a. Molecular phylogenetic analysis of human immunodeficiency virus type 1 strains obtained from Korean patients: *env* gene sequences. *AIDS Res. Hum. Retroviruses* 15, 303-307.

Kim, E.Y., Y.S. Cho, S.H. Maeng, C. Kang, J.G. Nam, and J.S. Lee. 1999b. Characterization of V3 loop sequences from HIV type 1 subtype B in South Korea: predominance of the GPGS motif. *AIDS Res. Hum. Retroviruses* 15, 681-686.

Korber, B., M. Muldoon, J. Theiler, F. Gao, R. Gupta, A. Lapedes, B.H. Hahn, S. Wolinsky, and T. Bhattacharya. 2000. Timing the ancestor of the HIV-1 pandemic strains. *Science* 288, 1789-1796.

Lee, KW., D.H. Oh, C. Lee, and S.Y. Yang. 2005. Allelic and haplotypic diversity of HLA-A, -B, -C, -DRB1, and -DQB1 genes in the Korean population. *Tissue Antigens* 65, 437-447.

Lee, D.H., Y. Yoon, and C.H. Lee. 2003. Phylogenetic analysis of the HIV-1 *nef* gene from Korean isolates. *J. Microbiol.* 41, 232-238.

Levy, J.A. 1998. HIV and the Pathogenesis of AIDS. 2nd ed. ASM Press, Washington, D.C., USA.

Lukashov, V.V. and J. Goudsmit. 2002. Recent evolutionary history of human immunodeficiency virus type 1 subtype B: reconstruction of epidemic onset based on sequence distances to the common ancestor. *J. Mol. Evol.* 54, 680-691.

Lukashov, V.V. and J.V. Goudsmit. 2003. Recent evolutionary history of HIV-1 subtype B-rebuttal. *J. Mol. Evol.* 56, 645-647.

McMichael, A.J. and S.L. Rowland-Jones. 2001. Cellular immune responses to HIV. *Nature* 410, 980-987.

Mindell, D.P. 1996. Positive selection and rates of evolution in immunodeficiency viruses from humans and chimpanzees. *Proc. Natl. Acad. Sci. USA* 93, 3284-3288.

Moore, C.B., M. John, I.R. James, F.T. Christiansen, C.S. Witt, and S.A. Mallal. 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* 296, 1439-1443.

Nielsen, R. and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148, 929-936.

Park, C.S., M.S. Kim, S.D. Lee, S.S. Kim, K.M. Lee, and C.H. Lee. 2006. Molecular phylogenetic analysis of HIV-1 *vif* gene from Korean isolates. *J. Microbiol*. 44, 655-659.

Park, C.S., M.S. Kim, H.A. Yi, D.H. Lee, K.M. Lee, and C.H. Lee. 2007. Analysis of substitution events in HIV-1 *vif* gene of the Korean clade. *J. Microbiol*. 45, 75-78.

Park, C.S., D.H. Lee, K.M. Lee, and C.H. Lee. 2008. Characterization and signature pattern analysis of Korean clade of HIV-1 using *nef* gene sequences. *J. Microbiol*. 46, 88-94.

Price, D.A., P.J. Goulder, P. Klenerman, A.K. Sewell, P.J. Easterbrook, M. Troop, C.R. Bangham, and R.E. Phillips. 1997. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proc. Natl. Acad. Sci. USA* 94, 1890-1895.

Salemi, M., K. Strimmer, W.W. Hall, M. Duffy, E. Delaporte, S. Mboup, M. Peeters, and A.M. Vandamme. 2001. Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. *FASEB J.* 15, 276-278.

Sharp, P.M. 1997. In search of molecular darwinism. *Nature* 385, 111-112.

Sharp, P.M., E. Bailes, R.R. Chaudhuri, C.M. Rodenburg, M.O. Santiago, and B.H. Hahn. 2001. The origins of acquired immune deficiency syndrome viruses: where and when? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 356, 867-876.

Sheehy, A.M., N.C. Gaddis, J.D. Choi, and M.H. Malim. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418, 646-650.

Smith, U.R., C. Kuiken, and B.T. Korber. 2003. Recent evolutionary history of human immunodeficiency virus type 1 subtype B-response. *J. Mol. Evol.* 56, 643-644.

Sung, H., B.T. Foley, I.G. Bae, H.S. Chi, and Y.K. Cho. 2001 Phylogenetic analysis of reverse transcriptase in antiretroviral drug-naive Korean HIV type 1 patients. *AIDS Res. Hum. Retroviruses* 17, 1549-1554.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876-4882.

Travers, S.A., J.P. Clewley, J.R. Glynn, P.E. Fine, A.C. Crampin, F. Sibande, D. Mulawa, J.O. McInerney, and G.P. McCormack. 2004. Timing and reconstruction of the most recent common ancestor of the subtype C clade of human immunodeficiency virus type 1. *J. Virol.* 78, 10501-10506.

Welker, R., H. Kottler, H.R. Kalbitzer, and H.G. Krausslich. 1996. Human immunodeficiency virus type 1 Nef protein is incorporated into virus particles and specifically cleaved by the viral proteinase. *Virology* 219, 228-236.

Yang, Z. 1997. Phylogenetic analysis by maximum likelihood (PAML). Version 1.4. Department of Biology, University of California,

Berkely, USA.
Yu, X., Y. Yu, B. Liu, K. Luo, W. Kong, P. Mao, and X.F. Yu. 2003. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302, 1056-1060.
Zanotto, P.M., E.G. Kallas, R.F. de Souza, and E.C. Holmes. 1999. Genealogical evidence for positive selection in the *nef* gene of HIV-1. *Genetics* 153, 1077-1089.