HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 SUBTYPES AMONG MALAYSIAN INTRAVENOUS DRUG USERS

TS Saraswathy[1], KP Ng[2], M Sinniah[1]

[1]Institute for Medical Research, Jalan Pahang, 50588 Kuala Lumpur, Malaysia; [2]Department of Medical Microbiology, Faculty of Medicine, University Malaya, 59100 Kuala Lumpur, Malaysia

Abstract. The HIV-1 genetic variation in 60 infected Malaysian intravenous drug users (IDU) was studied by comparison of the nucleotide sequences and their predicted amino acid sequences in the V3 loop of the external glycoprotein gp120. In this study, HIV-1 B, C and E subtypes were identified among Malaysian IDU, with HIV-1 B being the predominant subtype (91.7%). HIV-1 C and HIV-1 E were minority subtypes among Malaysian IDU. Analysis of the amino acid alignment of the C2-V3 region of the *env* gene suggests a genetic relationship between Thai and Malaysian B and E subtype strains. This study serves as a baseline for monitoring HIV-1 genetic diversity and spread in Malaysia.

INTRODUCTION

The extensive variability of the human immunodeficiency virus type 1 (HIV-1) has been well characterized and documented (Hu *et al*, 1996). Most of the sequence variability found in the extracellular envelope gene, *env,* is clustered in 5 hypervariable domains, V1-V5. The third hypervariable variable region V3 is an important target of research worldwide because it contains a principal neutralizing determinant (PND) as well as epitopes for T helper cell and cytotoxic T lymphocyte recognition. Genetic variation within the V3 loop have profound effects on viral infectivity, tropism and syncytia formation (Grimalia *et al*, 1992).

Analysis of HIV genetic variation involves comparison of multiple sequences covering the C2-V3 region of the *env* gene. Through the viral sequencing efforts of many groups of scientists, several clusters or subtypes of HIV-1 have now been identified. These subtypes are designated A through I, which constitute the major group M, and an additional divergent group outside M, provisionally categorized as Group O. Distinct geographic association between these subtypes have also been reported (UNAIDS, 1998).

The objective of this study was to investigate the HIV-1 genetic variability in 60 infected Malaysian intravenous drug users by analysis and comparison of the nucleotide and predicted amino acid sequences in the V3 loop of the *env* gene. An assessment of HIV subtypes in Malaysia would contribute to the present knowledge of HIV heterogeneity and molecular epidemiology of HIV infection in this region.

MATERIALS AND METHODS

Samples

EDTA blood samples were collected from 60 HIV-1 seropositive intravenous drug users during the period August 1996 and October 1997. The drug addiction history of these subjects were obtained from information either in questionnaire forms or from case histories in medical officers' reports.

DNA polymerase chain reaction

Peripheral blood mononuclear cells (PBMC) were isolated from whole blood by Ficoll Hypaque gradient centrifugation. Proviral DNA was extracted from PBMCs using lysis buffer containing detergents and proteinase K (Kellog and Kwok, 1990). The isolated DNA was subjected to a nested polymerase chain reaction (PCR) procedure to amplify the C2-V3 region of the HIV-1 *env* gene. The primer pairs for the primary PCR was MK603 (5′ CAGAAAAATTGTGGGTCACAGTCTATTATG GGGTACCT3′) and CO602 (5′ GCCCATAGTGC TTCCTGCTGCTCCCAAGAACC 3′). For the nested PCR, primers MK650 (5′ AATGTCAG CACAGTACAATGTACC 3′) and CO601 (5′ TT CTCCAATTGTCCCTCATATCTCCTCCTCCA 3′) were used (Ou *et al*, 1993). The PCR amplification cycle of 96ºC for 1 minute and 65ºC for 5 minutes was used to generate a 720 bp DNA fragment.

Nucleotide sequencing and analysis

The PCR amplified products were purified and directly sequenced according to the protocol described in the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction kit (Perkin Elmer

Corp, Applied Biosystem Division, Foster City, CA, USA). The internal primer used in the reaction was KH41 (5' TCAACTCAACTGCAGTTAAAT 3'), encompassing the V3 domain of the *env* gene and was obtained courtesy from Hideaki Tsuchie, Research Institute for Microbial Diseases, Osaka University, Japan. Sequencing reactions were performed in an automated DNA Sequencer (Applied Biosystems, Foster City, CA, USA). The nucleotide sequences for all the subjects were edited and translated to predicted amino acid sequences using DNASIS version 3.2 Software (Hitachi Software Engineering, Japan). Phylogenetic tree analysis was done with the PAUP (phylogenetic analysis using parsimony) program, version 3.0 (DL Swofford, Illinois Natural History Survey, Urbana, IL, USA).

RESULTS

Epidemiological data of subjects

Table 1 summarizes the HIV associated risk factor of the study group and HIV-1 subtype identified for each category. The IDU were all males in the productive age group ranging from 24 to 43 years (mean age = 33.0) 71.7% (43/60) of the subjects reported intravenous drug use without any other associated HIV risk factor. The remaining 28.3% (17/60) reported an additional heterosexual risk.

Predicted amino acid sequences of the V3 loop

From analysis of the amino acid sequence alignment in the C2-V3 region of the HIV-1 *env,* 91.7% (55/60) were identified as HIV-1 B subtype, 1.7% (1/60) as HIV-1 C subtype and 6.7% (4/60) as HIV-1 E subtype (Fig 1). The identification of the sub-

```
B-CON 95          CTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHC
                                                      N
SEQ TYPE 1        ------------PL-----W----Q----------  18
SEQ TYPE 2        ------------PL---Q-W----Q----------  12
SEQ TYPE 3        ------------PL---K-W----Q----------   3
SEQ TYPE 4        -------------L-----W----Q----------   3
SEQ TYPE 5        -------------L---Q-W----Q----------   3
SEQ TYPE 6        ------------P--------AI-N----------   1
SEQ TYPE 7        ------------NL-L---W----Q----------   1
SEQ TYPE 8        -------------L-Q---W----Q---N------   1
SEQ TYPE 9        -----------VPL-----W----Q----------   1
SEQ TYPE 10       ----------R-YL-----W----Q----------   1
SEQ TYPE 11       ------------PL-----W----R----------   1
SEQ TYPE 12       ------------PL-----W----Q-----GG---   1
SEQ TYPE 13       ------------PL---K-W----Q--------Y-   1
SEQ TYPE 14       ------------PL-----W----Q------R---   3
SEQ TYPE 15       ------------PL-----W----H------R---   1
SEQ TYPE 16       -------------------W-A--R--------Y-   1
SEQ TYPE 17       ------------PL---Q-W----R----------   1
SEQ TYPE 18       ------------PL-----W----Q-L--M-R---   1
SEQ TYPE 19       ------------PL-----W----D----------   1

C CON 95          CTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHC

S43352            -V----------------A----N-----------   1

E CON 95          CTRPSNNTRTSITIGPGQVFYRTGDIIGDIRQAHC

S44697            -----------L--------------T-N--K-Y-   1
UHR091            ------------PL-----C-T--Q----------   1
S29832            -------------L-----C-T---------G---   1
S43044            ---------K--PL-----C-T--H----------   1
```

Fig 1–Analysis of variation in the C2-V3 loop amino acid sequences of Malaysian HIV strains compared with B, C and E subtype consensus sequences from the Los Alamos HIV database. Residues identical to consensus sequence are shown as dashes ("-"), substitutions are indicated by standard amino acid code letter. N= number of subjects.

Table 1
Epidemiological data of HIV infected IDU.

Total number of IDU: 60		HIV-1 subtype		
	No. (%)	B	E	C
Age group (years)				
20-30	21 (35.0%)	17	3	1
31-40	34 (56.6%)	33	1	0
41-43	5 (8.0%)	5	0	0
HIV associated risk factor				
Intravenous drug use (IDU)	43 (71.7%)	39	3	1
IDU/heterosexual	17 (2.8%)	16	1	0

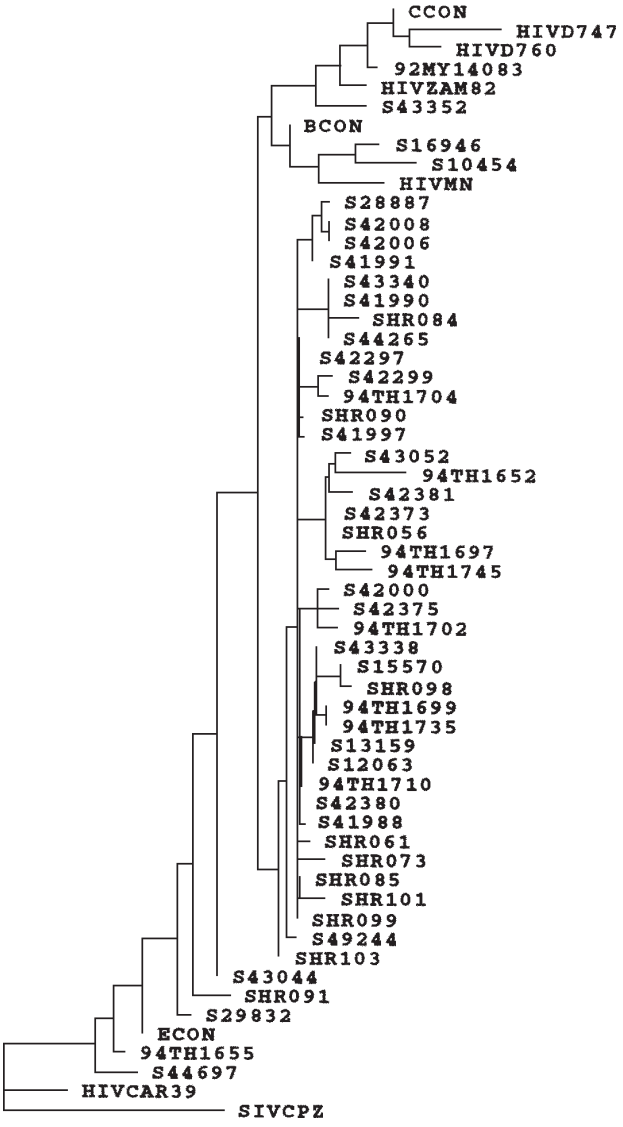

Fig 2–Phylogenetic tree comparing HIV-1 *env* C2-V3 region of Malaysian strains with subtypes B, C and E sequences from the Los Alamos HIV Database. The tree was constructed using PAUP version 3.0, with Simian Immunodeficiency Virus SIV_{CPZ} as the outgroup.

types were based on distinguishing features of various subtypes as well as close nucleotide and amino acid similarity with other sequences in the Los Alamos HIV Database (Myers *et al*, 1994). The genetic relationship among the Malaysian HIV-1 strains are depicted in the phylogenetic tree in Fig 2. Subtypes B, C and E sequences from the Los Alamos HIV Database were included for comparison. The strains used for comparison were North American/ European reference strain HIV_{MN}, Central African Republic HIV-1 E and C subtypes HIVCAR39 and HIVZAM82, Thailand genotype B and E strains, TH1697, TH1704, TH1652, TH1699, TH1702, TH1710, TH1735, TH1745, TH1655, Indian C subtype HIVD747 and Malaysian C subtype MY141083.

DISCUSSION

In this study the HIV-1 genetic variation in 60 HIV-1 infected Malaysian IDU was analysed by comparison of the nucleotide sequences and predicted amino acid sequences in the V3 loop of the *env* gene gp120.

The C2-V3 domain of the *env* was analysed because this domain contain nucleotide sequences with sufficient variability to distinguish between strains and identify subtypes. About 183 nucleotide base sequences (nucleotides 6625 to 6807) in the C2-V3 region of the *env* gene were sequenced. Analysis of the sequence data showed that B, C and E subtypes were prevalent among the Malaysian IDU population. This is consistent with findings of the Centers for the Disease Control, Atlanta, USA who had reported B and C subtypes among 13 Malaysian IDU and a E subtype in an infant born of Thai parentage (Brown *et al*, 1996). Our findings show no correlation between the type of amino acid sequence in the C2-V3 region and the ethnic origin or age of the subjects.

The predominant HIV-1 subtype among the Malaysian IDU was subtype B (91.7%). It is interesting to note that the studies of HIV-1 subtype prevalence in other parts of Asia (Thailand, Myanmar and China) showed subtype B to be the predominant subtype among IDU (Htoon *et al*, 1994; Weniger *et al*, 1994). The majority of the Malaysian HIV-1 subtype B V3 loop sequences had 5 amino acid substitutions when compared with the North American/European HIV-1 B consensus subtype and are closely related to the Thai genotype B reported among IDU in Thailand (Ou *et al*, 1993). Only one sample belonged to the HIV-1 C subtype. C subtype strains are characterized by an extra potential site for glycosylation close to the CD4 binding domain and a consistent absence of the highly conserved amino-linked glycosylation site proximal to the first cysteine (C) in the V3 loop. The source of HIV-1 infection for the subtype C case in this study could not be determined since no further epidemiological

data could be obtained from this case. No HIV-1 subtype has been reported from Thailand. Four samples in this study identified with HIV-1 subtype E. This was a minority subtype among the IDU population with 6.7% compared to 91.7% of HIV-1 subtype B. Although the sequences of each of the 4 subtype E were distinct, with 4 to 6 amino acid substitutions from the consensus E subtype sequence, they grouped closely with the Thai subtype E strains and not with the Central African Republic.

Sequence analysis observations in this study suggest a genetic relationship between Thai and Malaysian HIV-1 B and E subtypes. Malaysia shares an extensive border with Thailand and there is much sea borne traffic between the two countries. In Thailand, although both subtype B and E are present, subtype E is predominant with an estimated ratio of total E to B infections of about 10:1 (Weniger *et al*, 1994). The ratio was estimated bearing in mind that the number of people infected sexually greatly exceeds the number infected by injecting drug use. In Malaysia, the rates of HIV infection among prostitutes have so far been low, with the majority of those found to be HIV infected being Thai, however, the number tested were relatively small (Singh *et al*, 1994). In this study the predominant subtype among IDU was subtype B. Since HIV screening is routinely carried out for IDU in Drug Rehabilitation Centers and prisons, the reported cases would be higher for this category.

HIV-1 B, C and E subtypes were prevalent among IDU in Malaysia. The genetic relationship between the Malaysian and Thai B and E has implications for the understanding of molecular epidemiology of HIV infection in the two neighboring countries and in the design of HIV vaccines for use in this region.

ACKNOWLEDGEMENTS

We thank the Director of Institute for Medical Research for permission to publish this paper. This work was funded by a Research and Development Grant from the Ministry of Science and Technology.

REFERENCES

Brown TM, Robbins KE, Sinniah M, *et al.* HIV type 1 subtypes in Malaysia include B, C, and E. *AIDS Res Hum Retrovir* 1996;12:1655-7.

Grimalia RJ, Fuller BA, Rennert PD, *et al.* Mutations in the principal neutralization domain of the human immunodeficiency virus type 1 affect syncytium formation, virus infectivity, growth kinetics and neutralization. *J Virol* 1992;66: 1875-83.

Htoon MT, Lwin HT, San KO, Zan E, Thwe, M. HIV/AIDS in Myanmar. *AIDS* 1994;8 (suppl 2): S105-9.

Hu DJ, Dondero TJ, Rayfield MA, *et al.* The emerging genetic diversity of HIV. *JAMA* 1996; 275:210-6.

Kellog DE, Kwok S. Detection of human immunodeficiency virus. In: Innis Am, Gelfand DG, Sninsky JJ, White TJ, eds. PCR protocols: A guide to methods and applications. Academic Press, Harcout Brace Jovanovich Publishers, 1990;146-52.

Myers G, Korber B, Wain-Hobson S, Jeang KT, Henderson LE, Pavlakis GN. Human retroviruses and AIDS. A compilation and analysis of nucleic acid and amino acid sequences. Los Alamos, NM: Los Alamos National Library. 1994.

Ou CY, Takebe Y, Weniger BG, Luo, *et al.* Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. *Lancet* 1993;341: 1171-4.

Singh J, Sulaiman CR, Chong S, Yee KC, Crofts N. AIDS in Malaysia. *AIDS* 1994; 8 (suppl 2): S99-3.

UNAIDS. Joint United Nations Programme on HIV/AIDS. Report on the Global HIV/AIDS Epidemic, http://ww.unaids.org/unaids/fact/index.html, 1998.

Weniger BG, Takebe Y, Ou CY, Yamazaki S. The molecular epidemiology of HIV in Asia. *AIDS* 1994;8 (suppl 2): 513-28.