

# The Explosive Human Immunodeficiency Virus Type 1 Epidemic among Injecting Drug Users of Kathmandu, Nepal, Is Caused by a Subtype C Virus of Restricted Genetic Diversity

ROBERT B. OELRICHS,<sup>1\*</sup> ISWAR L. SHRESTHA,<sup>2</sup> DAVID A. ANDERSON,<sup>3</sup>  
AND NICHOLAS J. DEACON<sup>1,4</sup>

*AIDS Molecular Biology Unit,<sup>1</sup> Hepatitis Research Unit,<sup>3</sup> and The National Centre for HIV Virology Research,<sup>4</sup>  
Macfarlane Burnet Centre for Medical Research, Fairfield, Victoria 3078, Australia,  
and Siddhi Polyclinic, Kathmandu, Nepal<sup>2</sup>*

Received 15 April 1999/Accepted 2 November 1999

**An explosive epidemic of human immunodeficiency virus type 1 (HIV-1) has been documented among the injecting drug user population of Kathmandu, Nepal, whose seropositivity rate has risen from 0 to 40% between 1995 and 1997. By using Catrimox to preserve whole-blood RNA at ambient temperature for transportation, HIV-1 envelope V3–V4 sequences were obtained from 36 patients in this group. Analysis of the sequences indicated a homogenous epidemic of subtype C virus, with at least two independent introductions of the virus into the population. Viral diversity was restricted within two transmission subclusters, with the majority of variation occurring in V4. Calculation of the synonymous-to-nonsynonymous mutation ratio (Ks:Ka) across this region showed that significant evolutionary pressure had been experienced during the rapid horizontal spread of the virus in this population, most strongly directed to the region between V3 and V4.**

Despite initial optimism that an incipient epidemic of human immunodeficiency virus type 1 (HIV-1) in the Himalayan kingdom of Nepal had been averted through the establishment of education programs (27), more recent data have shown a sharp rise in the incidence of seropositivity in this country. Between 1995 and 1997, HIV-1 seroprevalence among the injecting drug users (IDU) of Kathmandu rose from 0% to 40 to 50% (N. Crofts, personal communication; R. L. Gurubacharya, V. L. Gurubacharya, J. Shrestha, and B. Pulchowk, Conf. Record, 12th World AIDS Conf., abstr. 23246, p. 390, 1998), associated with the shared injection in this group of Tidigestic (buprenorphine, a semisynthetic opiate pharmaceutical) or “brown sugar” (an impure form of heroin). The extent of high-risk behavior among the IDU in Kathmandu is also reflected in the alarming incidence of hepatitis C virus seropositivity, which is 94% among IDU, against a background of 0.6% in the general population (30).

An explosive spread of HIV-1 has recently been documented in geographically defined IDU cohorts in Russia (18) and Belarus (21), demonstrating low genetic diversity in these epidemics. The number of IDU in greater Kathmandu was estimated at 2,000 in 1997 (12) (although the number is now probably higher), and the seroprevalence data indicated a rapid spread of HIV-1 within this relatively isolated population. Since no prior investigations have documented the molecular epidemiology of the Nepalese HIV-1 epidemic, this examination was undertaken to survey the prevailing subtype and mode of evolution of the virus from IDU in Kathmandu.

Previous studies that have examined the molecular structure of HIV-1 from regions distant from the laboratory have largely used dried blood spots, collected on filter paper and transported at ambient temperature, to analyze small regions of proviral DNA (3, 4) or viral RNA (5) by PCR amplification. In

this study, a technique was devised that enabled the safe transfer and recovery of viral RNA from whole blood collected in Kathmandu and transported and stored at ambient temperature for as long as 3 months.

## MATERIALS AND METHODS

**Patients and sample collection.** Whole-blood samples were collected after informed consent from randomly selected subjects in Kathmandu, Nepal. A history of intravenous drug use was the sole selection criterion. A total of 46 IDU were examined (42 males and 4 females; average age, 27 years), 35 of whom had previously been found seropositive by an HIV-1 enzyme-linked immunosorbent assay (ELISA) (United Biomedical Incorporated, Huppauge, N.Y.), 2 of whom had previously been found HIV-1 seronegative by the same method, and 9 of whom had not yet been tested. In addition, negative-control samples from four subjects at low risk for HIV-1 exposure were collected (Table 1).

**Sample transport and RNA extraction.** The method used for sample transport and RNA extraction is outlined in Fig. 1. Catrimox (Catrimox-14 surfactant; Qiagen, Hilden, Germany) was dispensed in 900- $\mu$ l aliquots into 1.5-ml screw-cap microcentrifuge tubes in Melbourne, Australia, and shipped at ambient temperature to Kathmandu. Samples (200  $\mu$ l) of whole blood were added directly to the tube, followed by vigorous shaking by hand. The samples were then shipped by mail to Melbourne and stored at ambient temperature before analysis. The total time from specimen collection to analysis was 2 to 3 months.

Samples were mixed by vortexing for 30 s and were then spun at 20,000  $\times$  g for 3 min in a microcentrifuge. The supernatant was discarded by inversion, 1 ml of RNase-free water was added, and the pellet was resuspended by vortexing for 30 s. The sample was then centrifuged as previously for 1 min, the supernatant was discarded by inversion, and the wash with water was repeated. The pellet was resuspended in 350  $\mu$ l of RNA Lysis Tissue (RLT) Buffer (RNeasy mini kit; Qiagen) and disrupted by 6 passages through a 27-gauge needle (Terumo, Elkton, Md.). The remainder of the RNeasy protocol was then followed by using the manufacturer's buffers and conditions. The RNA was eluted in 30  $\mu$ l of RNase-free water, with incubation of the column and eluate at 65°C for 10 min before centrifugation.

Reverse transcription of 9.2  $\mu$ l of the RNA template was performed with an avian myeloblastosis virus reverse transcription (RT)-PCR kit (Boehringer Mannheim, Indianapolis, Ind.) and the following oligonucleotide primers in equimolar concentrations: RONPRT1, GTT CTG CTG TTG CAC TAT ACC AGA CAA TAA (pNL43 nucleotide positions [nt] 7873 to 7844); RONPRT2, TCC CTA GTC CCA CTG CTC TTT TTT CTC (pNL43 nt 7756 to 7737); and RONPRT3, TGC GGA TAA CAA TTT CAC ACA GGC AGA TGA GTT TTC CAG AGC AGC CC (pNL43 nt 8023 to 7994, with the 5' addition of the M13 reverse universal primer sequence). Triple-nested PCR was then performed with *Taq* DNA polymerase (Boehringer GmbH, Mannheim, Germany) and the manufacturer's buffers and conditions. Thermal cycling and reaction conditions were as follows. For the first round, 5  $\mu$ l of the RT reaction product was used as the

\* Corresponding author. Present address: HIV-NAT, Program on AIDS, Thai Red Cross Society, 1871 Rama IV Rd., Pathumwan, Bangkok, Thailand 10330. Phone: 66-2256-4579. Fax: 66-2653-2488. E-mail: oelrichs@loxinfo.co.th.

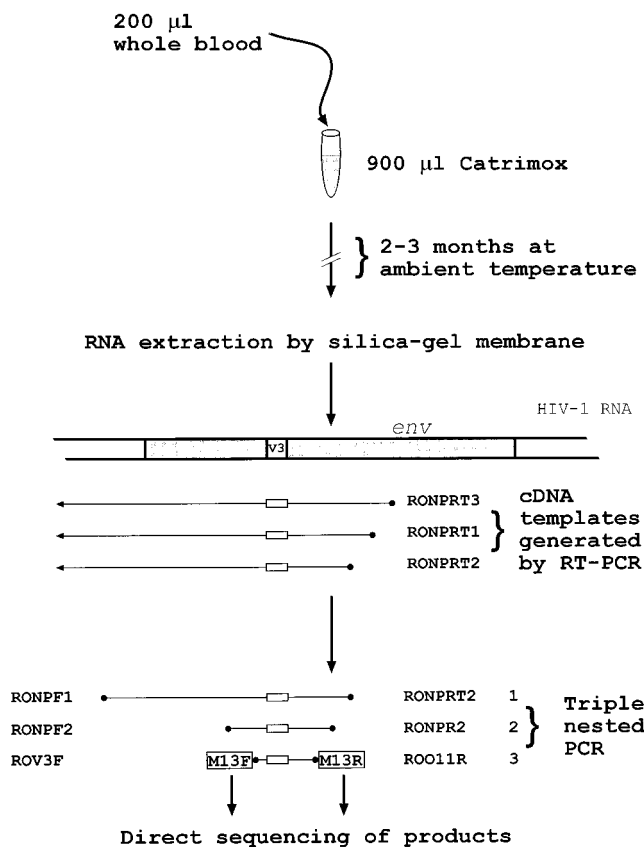


FIG. 1. Flow diagram outlining the method used to preserve, transport, and amplify HIV-1 RNA. Three oligonucleotide primers were included in the RT-PCR step to allow for the sequence variation in all HIV-1 subtypes. The final PCR step (with M13-tailed primers) generates products of sufficient quantity for direct automated DNA sequencing.

template, with 40 cycles of annealing at 50°C, extension at 72°C for 1 min, and oligonucleotide primers RONPF1 (GCT TAG GCA TCT CCT ATG GCA GGA AGA AA [pNL43 nt 5953 to 5981]) and RONPRT2. The second round was carried out with 5 µl of the first-round PCR product as the template, 40 cycles of annealing at 50°C, extension at 72°C for 1 min, and oligonucleotide primers RONPF2 (ATT ATT GTG CCC CGG CTG [pNL43 nt 6861 to 6878]) and RONPR2 (ATT TAT ATA ATT CAC TTC TCC AAT TGT C [pNL43 nt 7670 to 7643]). The third round was carried out with 2 µl of the second-round PCR product as the template and comprised 3 cycles of annealing at 55°C, extension at 72°C for 1 min, and finally 35 cycles in which denaturation was followed

directly by extension and annealing at 72°C for 1 min; the oligonucleotide primers used were ROV3F (TGC CAC GAC GTT GTA AAA CGA CGT CAG CAC AGT ACA ATG TAC ACA TGG [pNL43 nt 6938 to 6963, with the 5' addition of the M13 forward universal primer sequence]) and ROO11R (TGC GGA TAA CAA TTT CAC ACA GGT GGG AGG GGC ATA CAT TGC [pNL43 nt 7529 to 7511, with the 5' addition of the M13 reverse universal primer sequence]). The resulting 590-bp product was purified by using a spin column (Qiagen) and sequenced directly on both strands by using an automated DNA sequencer (LICOR, Lincoln, Nebr.). In the case of 99NP610 (an HIV-1 ELISA-positive IDU), proviral DNA sequence was obtained from a dried blood spot as described previously (4) under the PCR conditions described above. Samples from four individuals at low risk for HIV-1 exposure were collected, transported, and processed in tandem with positive specimens. None of these yielded detectable PCR products. Standard laboratory techniques were used to ensure that cross-contamination of samples did not occur, including the use of negative controls in all extractions and reactions and the physical separation of PCR products from templates in different laboratories. These data are summarized in Table 1.

**Sequence manipulation and phylogenetic analysis.** Primary sequence data were assembled and edited with Geneworks (Intelligenetics, Mountain View, Calif.) on a Power PC (Macintosh) platform. All subsequent analyses were performed by using the web interface of the Australian National Genomic Information Service (ANGIS). Sequences were aligned by using PILEUP, corrected, and codon aligned by hand. Square genetic distance matrices were generated by using EDNADIST; the Jukes-Cantor model of molecular evolution was used, correcting for multiple substitutions. The mean genetic distance within clusters, with range and standard deviation, was calculated from the matrix. Phylogenetic trees were constructed with a maximum-likelihood (ML) algorithm, EDNAML (a modified version of DNAML in PHYLIP, version 3.572c), by using empirical base frequencies, a transition/transversion ratio of 2.0, and a randomized input order of sequences. Trees were viewed and manipulated by using XTREE on a PC platform. The stability of tree topology was tested by the production of trees from the same alignment by using a maximum-parsimony algorithm (EDNAPARS) or a neighbor-joining (NJ) algorithm (ENEIGHBOR) and by using bootstrapping (ESEQBOOT with 50% resampling and 100 iterations) of the NJ trees. Synonymous (Ks) and nonsynonymous (Ka) substitutions per synonymous site were calculated by using DIVERGE (based on the algorithm of Li et al. [17], which ignores stop codons and codons that contain gap characters), and the mean Ks:Ka ratio was obtained from the matrices. Statistical significances of the difference of means were calculated by using an independent-sample *t* test in the Statistical Package for the Social Sciences (SPSS), version 7.0, with a two-tailed distribution and equal variances not assumed.

**HIV-1 sequences from published cohorts.** Published HIV-1 envelope DNA sequences from the following transmission cohorts were downloaded from GenBank for analysis: (i) the Florida dental cohort (index case and 5 recipients [6, 26]), (ii) the Kaliningrad IDU cohort (closely related sequences from 26 IDU [18]), and (iii) the Belarus IDU cohort (closely related sequences from 20 IDU [21]). Sequences from the Glenochil prison outbreak (an IDU transmission cluster of 13 [35]) were kindly provided by David Yirell. Sequences from the Sydney Blood Bank cohort (donor and three recipients) have been published previously (25). Where more than one viral sequence was available from an individual, one was selected at random.

**Nucleotide sequence accession numbers.** The 35 viral cDNA sequences obtained in this study have been deposited in GenBank under accession no. AF128961 through AF128995. The proviral sequence obtained from 99NP610 has been assigned accession no. AF128996.

TABLE 1. RT-PCR results for samples from each group of subjects<sup>a</sup>

Group	RT-PCR-negative samples	RT-PCR-positive samples	No. RT-PCR positive/total no. (%)
HIV-1 ELISA-positive IDU	99NP24, -29, -38, -39, -43 through -48	99NP13, -14, -15, -17, -18, -19, -21, -22, -23, -25, -27, -28, -31 through -37, -41, -50, -51, -57, -58, -59	25/35 (71)
HIV-1 ELISA-negative IDU		99NP30, -40	2/2 (100)
IDU of unknown serostatus	99NP49	99NP20, -26, -42, -52 through -56	8/9 (89)
Negative control (low risk)	99NP60 through -63		0/4 (0)

<sup>a</sup> Blood samples from Nepalese IDU and low-risk negative controls were transported and stored in Catrimox, viral RNA was recovered, and RT-PCR and direct sequencing of the products were performed. The percentage of samples that yielded products is shown for IDU whose HIV-1 serostatus was positive, negative, or unknown. Two patients who had tested negative by HIV-1 ELISA were found to be positive by RT-PCR. These patients were most probably sampled in the period after HIV-1 infection and before the development of anti-HIV-1 antibodies. Samples from four negative controls yielded no products by RT-PCR.

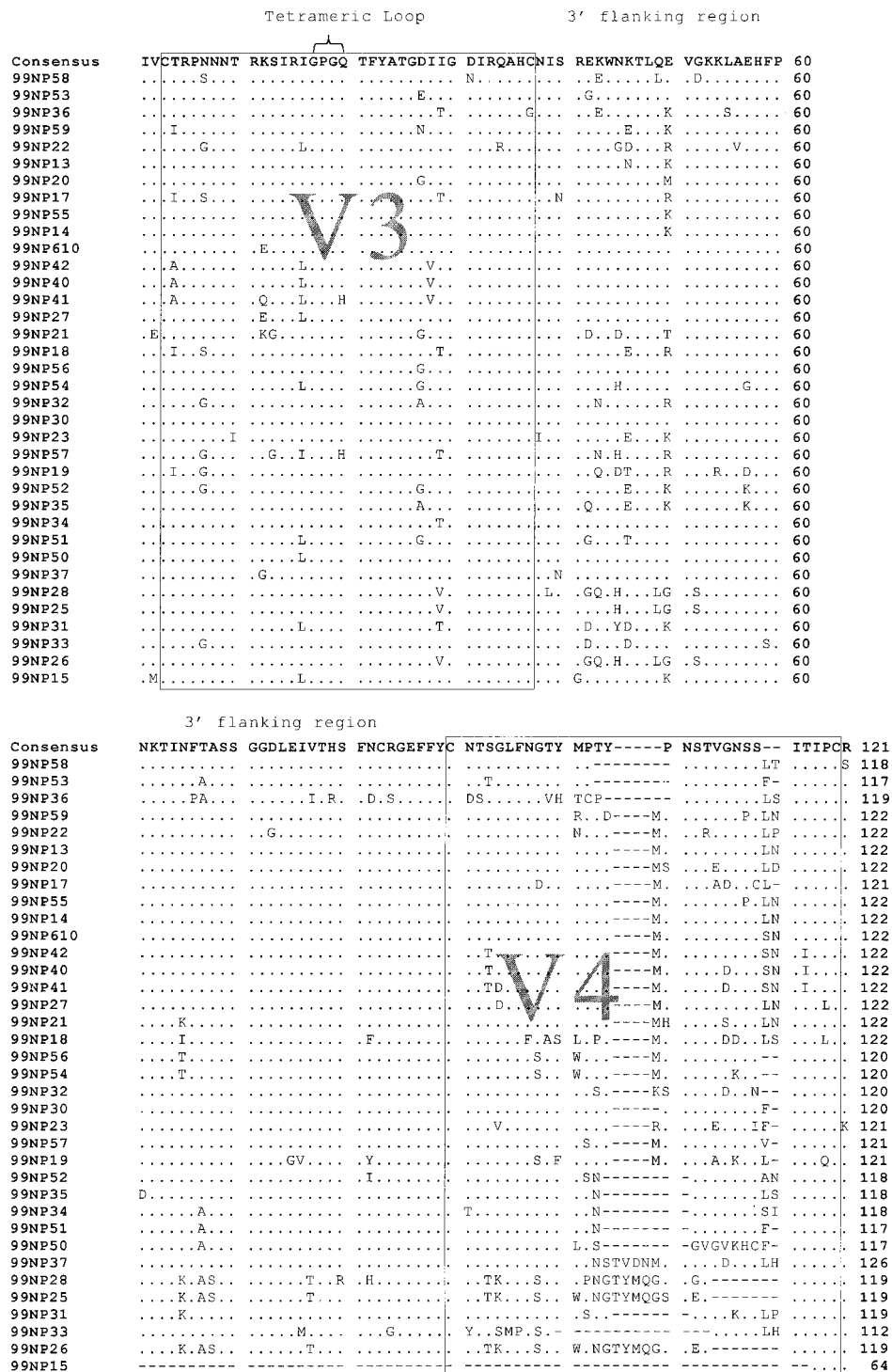


FIG. 2. Alignment of the predicted amino acid sequence of the HIV-1 envelope V3-V4 region from 36 Nepalese IDU. The two variable regions are boxed at the terminal cysteine residues. The consensus sequence is shown, and residues agreeing with the consensus are indicated by periods. Conservation of the GPGQ motif in the tetrameric V3 loop (indicated by a brace) is almost total. The most substantial variation occurs in V4, with numerous deletions (indicated by dashes), insertions, and mutations. Sample 99NP15 contains a 186-bp deletion encompassing almost the entire V4 region and was excluded from some of the following analyses.

RESULTS

**Amplification from viral RNA.** Of samples taken from 35 HIV-1-seropositive IDU, 25 (71%) yielded detectable products by RT-PCR (Table 1). Of the samples from nine IDU

whose HIV-1 serostatus was unknown, eight (89%) gave detectable products; 99NP49 alone was negative. Additionally, the two IDU who had previously been found HIV-1 seronegative yielded detectable HIV-1 PCR products. These were

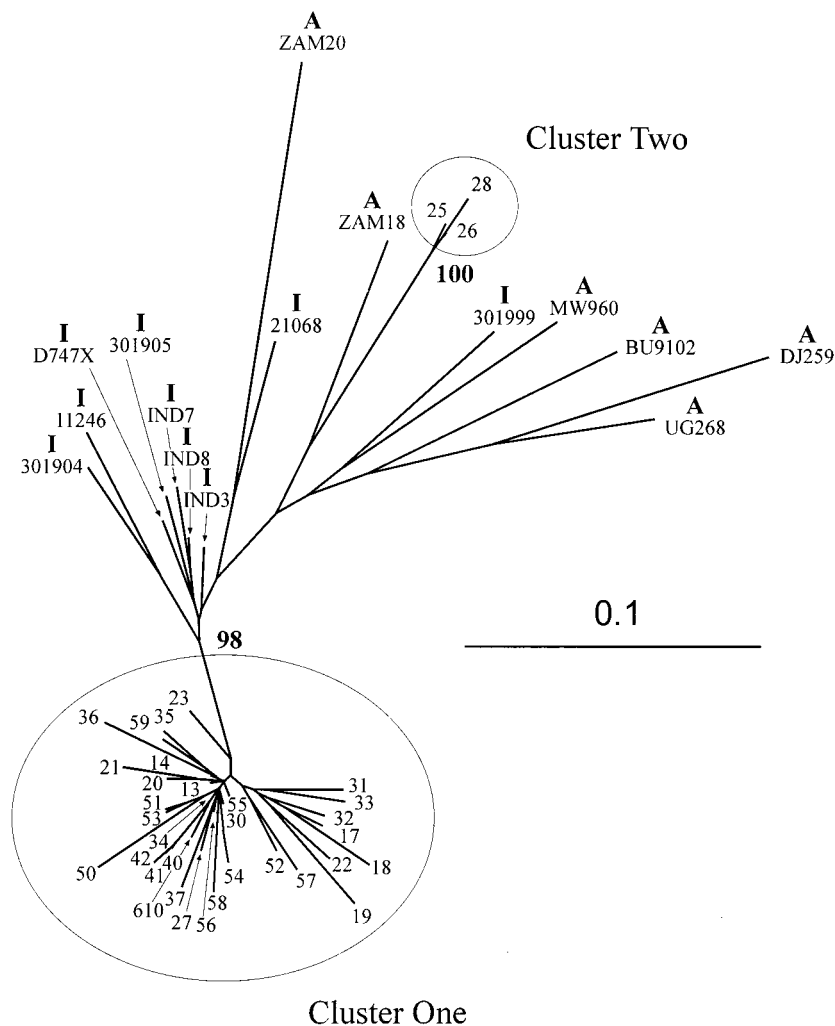


FIG. 3. Phylogenetic analysis of sequences from the HIV-1 envelope C3–V4 region (pNL43 nt 7023 to 7477) from 35 Nepalese IDU and 23 reference HIV-1 subtype C isolates from India and Africa. The country of isolate origin is indicated by I (Indian) or A (African). The phylogenetic tree was constructed by using an ML algorithm. The stability of the tree topology was verified by construction of a NJ tree by using the same alignment. This NJ tree was tested by using bootstrapping, and the percentage of bootstrap trees (from 100 data sets) for which the sequences at one end of the branch are a monophyletic group is indicated. The Nepalese sequences form two clusters that are distinct from other reported subtype C sequences (clusters 1 and 2). Cluster 1 is closest to a group of Indian isolates, whereas the nearest neighbor to cluster 2 is the African isolate ZAM18. The nucleotide alignment used to generate this analysis may be found at the R. B. Oelrichs website (<http://www.burnet.edu.au/oelrichs/align1.htm>). For sequences of Indian isolates, see reference 19 for 11246, 21068, 301904, 301905, and 301999; reference 9 for D747X; and reference 33 for IND3, IND7, and IND8. The sequences of African isolates BW01B03, BW0402, BW0504, BW1106, BW1210, BW15C05, BW16B01, and BW17B03 were obtained from GenBank (V. Novitsky unpublished data), and those of other African isolates may be found in the references cited, as follows: BU9102 (28); DJ259, UG268, ZAM18, and ZAM20 (20); and MW960 (11).

most probably sampled during the “window period” after HIV-1 infection but before the development of anti-HIV-1 antibodies. It has not yet been possible to repeat an HIV-1 ELISA on these patients. Assuming that the specificity of the ELISA is close to 100% and that subject 99NP49 was uninfected, the sensitivity of transport in Catrimox and amplification by this method is >76%. There was no evidence for cross-contamination of samples during transport or processing, and all derived nucleotide sequences were unique (<http://www.burnet.edu.au/oelrichs/align2.htm>).

**Analysis of V3–V4 envelope sequences.** Figure 2 shows an alignment of the predicted HIV-1 amino acid sequence derived from 35 samples by RT-PCR of viral RNA and from 1 additional sample (99NP610) by PCR of proviral DNA. The alignment has been truncated to show the V3–V4 region. All but 2 of the 36 viral sequences show the sequence GPGQ at the

tetrameric V3 loop tip, the conservation of which has been noted previously in HIV-1 subtype C (10). V3 is highly conserved in general, and much of the variation between sequences occurs in the V4 region. The sequence from 99NP15 contains a large deletion (186 bp) encompassing almost the entire V4 region. A total of three independent RT-PCRs yielded the same-sized fragment from this individual.

**Subtype of Nepalese sequences.** An NJ phylogenetic analysis generated from an alignment of the 36 Nepalese sequences and reference HIV-1 subtype sequences (15) over the region C3–V4 (pNL43 nt 7023 to 7477) showed the Nepalese sequences clustering with subtype C isolates with a high degree of certainty (bootstrap value from 100 iterations, 98% [data not shown]).

**Two distinct Nepalese clusters.** Figure 3 shows an ML tree estimating the phylogenetic relationship between the Nepalese

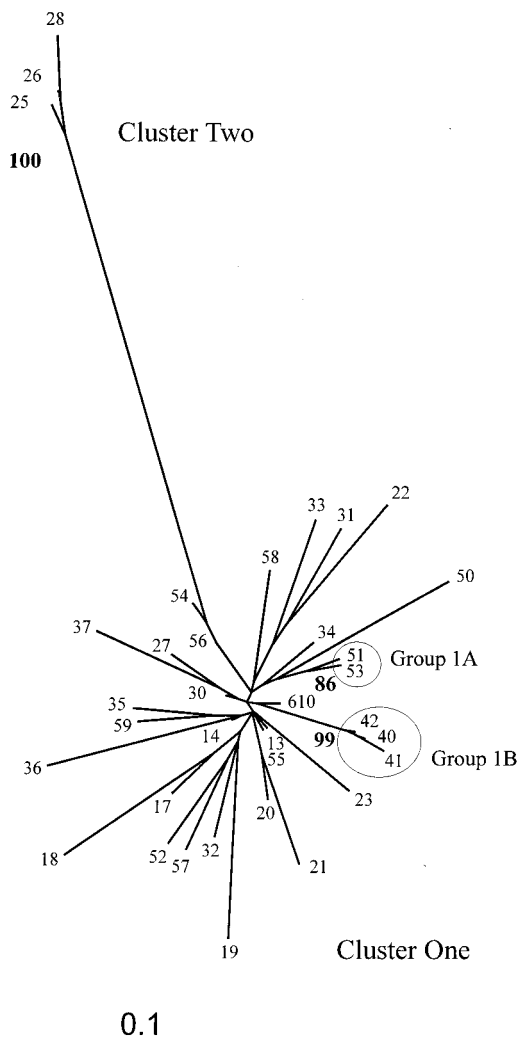


FIG. 4. Phylogenetic analysis of sequences from the HIV-1 envelope C3-V4 region (pNL43 nt 7023 to 7477) from the 35 Nepalese IDU HIV-1 subtype C isolates alone. An ML tree is shown with significant NJ bootstrap values (from 100 data sets) by the relevant nodes. The major and minor clusters (clusters 1 and 2, respectively) observed in Fig. 3 are preserved. Within cluster 1, two additional monophyletic groups (clusters 1A and 1B) are supported by highly significant bootstrap values. The nucleotide alignment used to generate this analysis may be found at the R. B. Oelrichs website (<http://www.burnet.edu.au/oelrichs/align2.htm>).

sequences and HIV-1 subtype C viruses from India and Africa over the region C3-V4. For this and subsequent estimates, the sequence from 99NP15 was removed, because the large deletion would have confused the analysis. Two monophyletic groups are evident within the Nepalese sequences, both supported by highly significant bootstrap values in NJ analysis and by maximum-parsimony analysis (data not shown). The larger of these (cluster 1), containing 33 of the sequences, is most closely related to a group of Indian isolates. The second, smaller cluster of three individuals (cluster 2) has a longer branch length between the ancestral node and other subtype C sequences—the nearest neighbor being ZAM18, from Africa. Although the three sequences in cluster 2 share a distinctive motif in V4 [QGPN(E/G)T (Fig. 2)], the distinction into two groups does not appear to be the result of a single “signature sequence” or discrete genomic region, as this separation is

maintained when NJ analysis is performed on smaller regions of the alignment encompassing V3, V4, or the intervening constant region (data not shown).

Figure 4 shows this phylogeny in more detail, presenting an ML tree estimating the relationships among the Nepalese sequences alone. The long branch length that separates the two major monophyletic clusters indicates the significant degree of divergence between the two. Within the major cluster of sequences, two subgroups are defined by significant NJ bootstrap values and supported by maximum-parsimony analysis (data not shown)—group 1A (99NP51 and 99NP53) and group 1B (99NP40, 99NP41, and 99NP42). It was not possible to assign a reason for the close association of the viral strains in these clusters from patient history or origin.

**Calculation of mean genetic distance.** Viral sequence diversity was calculated by mean pairwise genetic distance for the Nepalese IDU cohort (in total and separated into clusters 1 and 2) and a number of published HIV-1 parenteral transmission clusters and IDU epidemic cohorts (Table 2). The data were truncated to the largest region shared by all available sequences (V3 loop; pNL43 nt 7100 to 7207). The results are presented in order of diminishing mean genetic distance. The diversity of the Nepalese cohort is lower than that of the three nosocomial clusters but significantly higher than that for other published IDU outbreaks. The major Nepalese transmission group (cluster 1) alone did not differ in diversity from the entire IDU cohort analyzed together, although cluster 2 was of substantially lower diversity.

**Genetic distances and Ks:Ka ratios within the Nepalese cluster.** The distribution of genetic distance values in the Nepalese IDU C3-V4 region was examined in more detail. The entire alignment of the 35 viral cDNA sequences (pNL43 nt 7025 to 7477) was divided into four codon-aligned alignments: the 5' flanking region (pNL43 nt 7025 to 7099), the V3 region (pNL43 nt 7100 to 7207), the 3' flanking region (pNL43 nt 7208 to 7366), and the V4 region (pNL43 nt 7367 to 7468). The latter three regions are indicated in the translated amino acid sequences shown in Fig. 2. The mean genetic distance was calculated for each of these four alignments. Figure 5 shows the results of these calculations, with the range of pairwise genetic distances in each set presented to show the distribution of variation. As is evident from the predicted amino acid sequence alignment in Fig. 2, the majority of sequence variation is due to the V4 region, which has the highest mean genetic distance ( $12.9\% \pm 10.3\%$ ) as well as the broadest range of distances. By contrast, the distance values of V3 cluster more tightly around a lower mean of  $4.6\% \pm 2.3\%$ , and the distances of the 5' flanking region are confined below 6.0%. A significant amount of sequence variation is found within the 3' flanking region ( $5.5\% \pm 3.8\%$ ), and this region is more variable than V3, as has been described previously for HIV-1 subtype C isolates (10). To examine whether this sequence variability was due to increased neutral (random) mutation or was the product of evolutionary selection pressure, the Ks:Ka ratio was calculated for these alignments (Fig. 5). Ks:Ka values below 1.0 indicate a predominance of mutations driven by selective pressure as opposed to those occurring by random genetic drift (16, 24, 31, 36). Analysis of the entire alignment yielded a mean pairwise Ks:Ka ratio of 0.86 ( $P < 0.001$ ), indicating significant evolutionary pressure on the C3-V4 region as a whole. Interestingly, examination of the smaller regions alone revealed that the majority of this pressure is operating upon the 3' flanking region (Ks:Ka = 0.74;  $P < 0.001$ ), between V3 and V4. The Ks:Ka ratio within the V3 (0.93;  $P = 0.133$ ) and V4 (0.84;  $P < 0.001$ ) regions is also  $< 1.0$ , although, in concordance with the viral diversity, most evolution is occurring in the V4 rather

TABLE 2. Comparison of mean genetic distance in the V3 region among the Nepalese IDU cohort and six other, previously described parenteral-transmission cohorts<sup>a</sup>

Cohort	Length of epidemic (yr)	Subtype	Mean genetic distance (%)	SD (%)	Range (%)
Sydney Blood Bank ( <i>n</i> = 4)	10–16	B	13.9	4.6	7.2–22.7
Nosocomial pediatric ( <i>n</i> = 22)	2–4	G	6.1	3.6	0.0–22.0
Florida dental ( <i>n</i> = 6)	2–3	B	5.0	1.8	2.5–8.8
Kathmandu IDU ( <i>n</i> = 35)	<3–4	C	4.6	2.3	0.0–11.3
Cluster 1 ( <i>n</i> = 32)	<3–4	C	4.6	2.3	0.0–11.3
Cluster 2 ( <i>n</i> = 3)	<3–4	C	1.9	1.4	0.0–2.9
Belarussian IDU ( <i>n</i> = 20)	1–2	A	0.6	0.9	0.0–3.9
Kaliningrad IDU ( <i>n</i> = 26)	<0.5	A/B	0.6	0.7	0.0–2.9
Glenochil IDU ( <i>n</i> = 13)	<0.5	B	0.4	0.5	0.0–1.9

<sup>a</sup> Cohorts are arranged in order of declining mean genetic distance. The Nepalese IDU cohort shows a value consistent with previously published data from diverse HIV-1 subtypes. Mean genetic distance increases substantially in proportion to the time between the commencement of the epidemic (or transmission cluster) and sampling.

than the V3 region, and the Ks:Ka ratio in V3 does not achieve statistical significance. The 5' flanking region shows no evidence of selected mutations, with a Ks:Ka ratio of 2.3 ( $P < 0.001$ ).

## DISCUSSION

As the global pandemic of HIV-1 continues to expand, previously naive populations experience patterns of infection that are a product of the source of the virus and the predominant mode of transmission. The object of this study was to investigate the explosive spread of HIV-1 in the relatively confined population of IDU in Kathmandu, and the results have shed light on the possible source of the epidemic in that group and the particular mode of evolution of HIV-1 subtype C during rapid horizontal transmission.

Initially a less commonly reported subtype of HIV-1, subtype C has come to dominate the global pandemic and is responsible for rapidly spreading epidemics in South Africa and neighboring countries (2, 34) and in the Indian subcontinent (10, 19). It is not clear whether this predominance is a reflection of chance transmission of HIV-1 subtype C into fertile naive populations or whether it may reflect the relatively higher fitness of a virus that has adapted further to human transmission. Unlike other subtypes of HIV-1, strains of subtype C may contain an additional one or two copies of the binding site for the transcription factor NF- $\kappa$ B within the long terminal repeat promoter region (14), perhaps resulting in enhanced viral transcription and fitness.

HIV-1 subtype C was the only subtype detected in a total of 36 randomly selected IDU in Kathmandu, indicating that the epidemic is homogenous and possibly of Indian origin, rather than spreading from Southeast Asia, where the predominant subtypes are B' and A/E (10). Other neighboring countries have also reported subtype C infections, however, and the differences between reported strains are not sufficiently great to allow assignment of an isolate to a particular country of origin on the basis of sequence data alone.

The Kathmandu HIV-1 sequence information will be of relevance to any forthcoming vaccine strategies in Nepal. It will also be of interest to examine the HIV-1 epidemics in separate at-risk populations, for example, commercial sex workers or women receiving prenatal care. The detection of different subtypes of the virus in distinct at-risk groups, as was the case in the early HIV-1 epidemic in Thailand (22), would indicate a separate epidemiological origin of these epidemics.

A more detailed analysis of the 35 subtype C envelope sequences revealed clear segregation into minor and major clusters (Fig. 3 and 4). This segregation was not due to the presence of a discrete signature sequence but was demonstrable along the length of the aligned sequences. The larger of the two clusters has a star-shaped phylogeny, suggestive of horizontal transfer of the virus in a short time (10), in concordance with the available epidemiological information. It has not been possible to separate the three members of the minor cluster from the majority on the basis of geographical origin, ethnicity, or preferred mode of drug injection. It is probable, however, that these three are simply members of a more recent transmission cluster, sampled by chance, as the mean genetic distance in this group ( $1.9\% \pm 1.4\%$ ) is substantially lower than that in the larger cluster ( $4.6\% \pm 2.3\%$ ), consistent with a shorter time since transmission. These data indicate that the Nepalese IDU epidemic has been the product of a limited number (possibly as few as two) of introductions of HIV-1 subtype C into the population and a subsequent explosive transmission within it.

Comparison of these data with previously published sequences of parenteral transmission cohorts shows that the diversity of the Nepalese IDU sequences (as measured by mean pairwise genetic distance) is approximately centrally located within a broad range (Table 2). The highest diversity calculated is for the Sydney Blood Bank cohort (7), a function of the increased viral diversity that has been noted in such cohorts of long-term non-progressors to HIV disease (8, 23). By contrast, extremely low diversity is calculated for three large IDU transmission cohorts (from Kaliningrad, Belarus, and Glenochil) where sampling was performed early in the epidemic. A comparable diversity was calculated for the Nepalese IDU cohort and two other cohorts with epidemics of similar duration. These data indicate that, despite the explosive nature of the spread of HIV-1 in Kathmandu, it has already achieved substantial diversity, comparable to that in previously reported instances.

Examination of an alignment of the predicted amino acid sequences from the Nepalese IDU HIV-1 clearly shows that most variation has occurred in the V4 region (Fig. 2). Higher rates of evolution in the V3-flanking regions, compared to V3 itself, have been noted previously in HIV-1 subtype C (10, B. T. Foley, personal communication). In order to analyze the differential rates of evolution of these envelope regions, the data were divided into four separate alignments covering V3, V4, and the 5' and 3' sequences flanking V3. Viral diversity

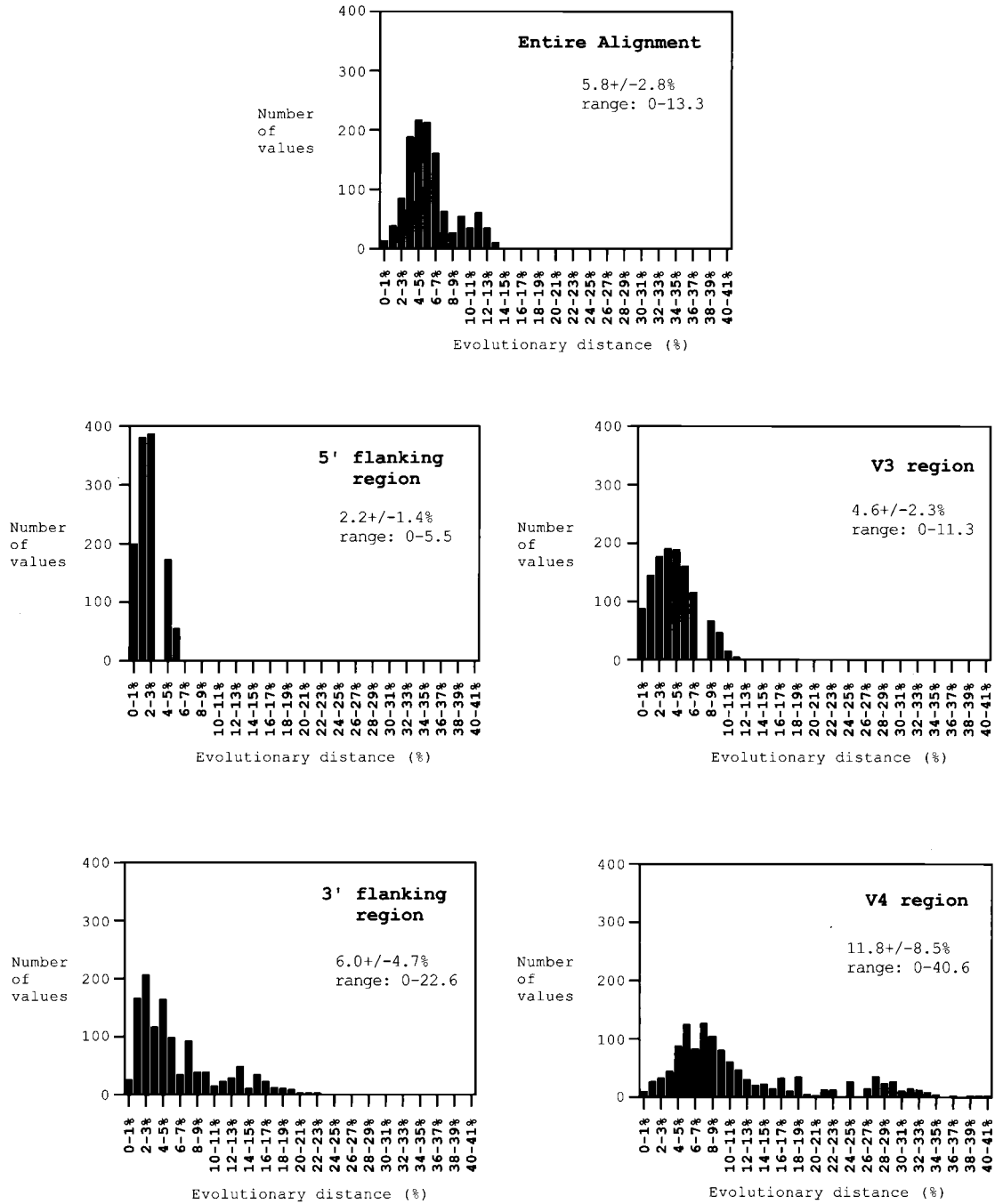


FIG. 5. Plots of the distribution of pairwise genetic distances in the entire alignment of 35 Nepalese IDU HIV-1 cDNA sequences (pNL43 nt 7025 to 7477) and in four subalignments covering the 5' flanking region (pNL43 nt 7025 to 7099), the V3 region (pNL43 nt 7100 to 7207), the 3' flanking region (pNL43 nt 7208 to 7366), and the V4 region (pNL43 nt 7367 to 7468). The number of distance values within a given percentage range (y axis) is plotted against an increasing percentage range, from 0 to 1% to 44 to 45% (x axis). The mean genetic distance, standard deviation, and range are given. The entire alignment shows moderate genetic diversity (average, 5.8%) and evidence of positive selection pressure (Ks:Ka = 0.86;  $P < 0.001$ ). The V4 region contributes most highly to genetic diversity, with the largest mean genetic distance and the broadest range of values. The 3' flanking region also contributes substantially to diversity and experiences the strongest evolutionary pressure (Ks:Ka = 0.74;  $P < 0.001$ ).

and Ks:Ka ratios were calculated from these (Fig. 5). All regions apart from the 5' flanking region show substantial diversity, with most variation occurring in V4 and the 3' flanking region rather than in V3. A Ks:Ka ratio of 0.86 ( $P < 0.001$ ) for the entire alignment indicates that the virus has experienced significant selective mutation during its rapid spread through

the Nepalese IDU. Interestingly, the highest evolutionary pressure has been directed towards the 3' flanking region (Ks:Ka = 0.73;  $P < 0.001$ ) rather than to the variable regions. These data confirm previous observations that the highest genetic diversity in the C3-V4 region of HIV-1 subtype C is found outside V3 and indicate that the flanking region immediately 3' to V3 is

the target of an unknown positive or purifying selective pressure during transmission.

The use of the cationic surfactant Catrimox-14 in the isolation of viral RNA has been described for the analysis of hepatitis C virus (1, 29, 32) and bovine viral diarrhoea virus (13). Addition of Catrimox leads to the disruption of cellular and viral particles and to the preservation of RNA from both these sources. This is the first report of its use in the transport and extended storage at ambient temperature of HIV-1 RNA in whole blood. This technique has a number of advantages over the use of dried-blood spots for the transport of HIV-1-positive specimens over extended distances. The collection procedure is simple, and the addition of the Catrimox reagent to the blood destroys infectious blood-borne viral particles, rendering the sample safer to transport through conventional mail. The extraction procedure is also simple and allows the analysis of viral RNA directly rather than by amplifying integrated proviral DNA copies of the virus that might be defective or archival. The technique is also efficient, achieving recovery of HIV-1 from >76% of ELISA-positive samples in this study.

This study reports the first genetic information from the explosive HIV-1 epidemic in Nepal, finding the subtype spreading amongst IDU to be exclusively subtype C and possibly of Indian origin. It reports a new application of the reagent Catrimox in the preservation and transport of HIV-1 RNA for analysis in laboratories distant from the site of collection, which offers advantages over currently used techniques. The study also presents the largest HIV-1 parenteral-transmission cohort examined to date and confirms previous observations concerning the unusual evolution of the V3 flanking regions of HIV-1 subtype C. These data have relevance for both existing and proposed public health interventions and vaccine strategies, designed to curb this potentially devastating epidemic.

#### ACKNOWLEDGMENTS

We thank Brian Thomas Foley and Nick Crofts for advice and helpful discussions and Mana Khongphatthanayothin and Scott Cowcher for help with statistical analysis. We gratefully acknowledge Jack Stapleton for suggesting the use of Catrimox in this study and M. R. Shambhu for sample collection.

This work was supported in part by an Australian government grant to the National Centre for HIV Virology Research and the Macfarlane Burnet Centre for Medical Research Fund.

#### REFERENCES

- Ali, S. A., B. Kubik, H. Gulle, M. M. Eibl, and A. Steinkasserer. 1998. Rapid isolation of HCV RNA from Catrimox-lysed whole blood using QIAamp spin columns. *BioTechniques* 25:975-978.
- Bredell, H., C. Williamson, P. Sonnenberg, D. J. Martin, and L. Morris. 1998. Genetic characterization of HIV type 1 from migrant workers in three South African gold mines. *AIDS Res. Hum. Retrovir.* 14:677-684.
- Cassol, S., T. Salas, M. Arella, P. Neumann, M. T. Schechter, and M. O'Shaughnessy. 1991. Use of dried blood spot specimens in the detection of human immunodeficiency virus type 1 by the polymerase chain reaction. *J. Clin. Microbiol.* 29:667-671.
- Cassol, S., B. G. Weniger, P. G. Babu, M. O. Salminen, X. Zheng, M. T. Htoon, A. Delaney, M. O'Shaughnessy, and C. Y. Ou. 1996. Detection of HIV type 1 env subtypes A, B, C, and E in Asia using dried blood spots: a new surveillance tool for molecular epidemiology. *AIDS Res. Hum. Retrovir.* 12:1435-1441.
- Cassol, S., M. J. Gill, R. Pilon, M. Cormier, R. F. Voigt, B. Willoughby, and J. Forbes. 1997. Quantification of human immunodeficiency virus type 1 RNA from dried plasma spots collected on filter paper. *J. Clin. Microbiol.* 35:2795-2801.
- Crandall, K. A. 1995. Intraspecific phylogenetics: support for dental transmission of human immunodeficiency virus. *J. Virol.* 69:2351-2356.
- Deacon, N., A. Tsykin, A. Solomon, K. Smith, M. Ludford-Menting, D. Hooker, D. McPhee, A. Greenway, A. Ellett, C. Chatfield, V. Lawson, S. Crowe, A. Maerz, S. Sonza, J. Learmont, J. Sullivan, A. Cunningham, D. Dwyer, D. Dowton, and J. Mills. 1995. Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. *Science* 270:988-991.
- Delwart, E. L., H. Pan, H. W. Sheppard, D. Wolpert, A. U. Neumann, B. Korber, and J. I. Mullins. 1997. Slower evolution of human immunodeficiency virus type 1 quasispecies during progression to AIDS. *J. Virol.* 71:7498-7508.
- Dietrich, U., M. Grez, H. von Briesen, B. Panhans, M. Geissendorfer, H. Kuhnel, J. Maniar, G. Mahambre, W. B. Becker, M. L. Becker, et al. 1993. HIV-1 strains from India are highly divergent from prototypic African and US/European strains, but are linked to a South African isolate. *AIDS* 7:23-27.
- Dighe, P. K., B. T. Korber, and Brian T. Foley. 1997. Global variation in the HIV-1 V3 region, p. III-74-III-206. *In* B. Korber, B. Hahn, B. Foley, J. W. Mellors, T. Leitner, G. Myers, F. McCutchan, and C. L. Kuiken (ed.), *Human retroviruses and AIDS 1997: a compilation and analysis of nucleic acid and amino acid sequences*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
- Gao, F., S. G. Morrison, D. L. Robertson, C. L. Thornton, S. Craig, G. Karlsson, J. Sodroski, M. Morgado, B. Galvao-Castro, H. von Briesen, S. Beddows, J. Weber, P. M. Sharp, G. M. Shaw, and B. H. Hahn. 1996. Molecular cloning and analysis of functional envelope genes from human immunodeficiency virus type 1 sequence subtypes A through G. The WHO and NIAID Networks for HIV Isolation and Characterization. *J. Virol.* 70:1651-1657.
- Hannum, J. 1997. AIDS in Nepal, communities confronting an emerging epidemic. AmFar publication. Seven Stories Press, New York, N.Y.
- Hyndman, L., S. Vilcek, J. Conner, and P. F. Nettleton. 1998. A novel nested reverse transcription PCR detects bovine viral diarrhoea virus in fluids from aborted bovine fetuses. *J. Virol. Methods* 71:69-76.
- Korber, B., B. Hahn, B. Foley, J. W. Mellors, T. Leitner, G. Myers, F. McCutchan, and C. L. Kuiken (ed.). 1997. *Human retroviruses and AIDS 1997: a compilation and analysis of nucleic acid and amino acid sequences*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
- Leitner, T., B. Korber, D. Robertson, F. Gao, and B. Hahn. 1997. Updated proposal of reference sequences of HIV-1 genetic subtypes. The human retroviruses and AIDS 1997 compendium, III-19-III-24. *In* B. Korber, B. Hahn, B. Foley, J. W. Mellors, T. Leitner, G. Myers, F. McCutchan, and C. L. Kuiken (ed.), *Human retroviruses and AIDS 1997: a compilation and analysis of nucleic acid and amino acid sequences*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
- Li, W. H., and D. Graur. 1991. *Fundamentals of molecular evolution*. Sinauer Associates, International, Sunderland, Mass.
- Li, W. H. 1993. Unbiased estimation of the rates of synonymous and non-synonymous substitution. *J. Mol. Evol.* 36:96-99.
- Liitsola, K., I. Tashkinova, T. Laukkanen, G. Korovina, T. Smolskaja, O. Momot, N. Mashkilleyson, S. Chaplinskias, H. Brummer-Korvenkontio, J. Vanhatalo, P. Leinikki, and M. O. Salminen. 1998. HIV-1 genetic subtype A/B recombinant strain causing an explosive epidemic in injecting drug users in Kaliningrad. *AIDS* 12:1907-1919.
- Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadhari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73:152-160.
- Louwagie, J., W. Janssens, J. Mascola, L. Heyndrickx, P. Hegerich, G. van der Groen, F. E. McCutchan, and D. S. Burke. 1995. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. *J. Virol.* 69:263-271.
- Lukashov, V. V., E. V. Karamov, V. F. Eremin, L. P. Titov, and J. Goudsmit. 1998. Extreme founder effect in an HIV type 1 subtype A epidemic among drug users in Svetlogorsk, Belarus. *AIDS Res. Hum. Retrovir.* 14:1299-1303.
- Mastro, T. D., C. Kumanusont, T. J. Dondero, and C. Wasi. 1997. Why do HIV-1 subtypes segregate among persons with different risk behaviors in South Africa and Thailand? *AIDS* 11:113-116.
- McDonald, R. A., D. L. Mayers, R. C. Chung, K. F. Wagner, S. Ratto-Kim, D. L. Birx, and N. L. Michael. 1997. Evolution of human immunodeficiency virus type 1 env sequence variation in patients with diverse rates of disease progression and T-cell function. *J. Virol.* 71:1871-1879.
- Myers, G., and G. N. Pavlakis. 1992. Evolutionary potential of complex retroviruses, p. 51-104. *In* J. A. Levy (ed.), *The retroviruses*. Plenum Press, New York, N.Y.
- Oelrichs, R., A. Tsykin, D. Rhodes, A. Solomon, A. Ellett, D. McPhee, and N. Deacon. 1998. Genomic sequence of HIV type 1 from four members of the Sydney Blood Bank Cohort of long-term nonprogressors. *AIDS Res. Hum. Retrovir.* 14:811-814.
- Ou, C. Y., C. A. Ciesielski, G. Myers, C. I. Bandea, C. C. Luo, B. T. M. Korber, J. I. Mullins, G. Schochetman, R. L. Berkelman, A. N. Economou, J. J. Witte, L. J. Furman, G. A. Satten, J. W. Curran, and H. W. Jaffe. 1992. Molecular epidemiology of HIV transmission in a dental practice. *Science* 256:1165-1171.
- Peak, A., S. Rana, S. H. Maharjan, D. Jolley, and N. Crofts. 1995. Declining



- risk for HIV among injecting drug users in Kathmandu, Nepal: the impact of a harm-reduction programme. *AIDS* **9**:1067–1070.
28. Penny, M. A., S. J. Thomas, N. W. Douglas, S. Ranjbar, H. Holmes, and R. S. Daniels. 1996. Env-gene sequences of primary HIV-1 isolates of subtypes B, C, D, E and F obtained from the WHO-Network for HIV Isolation and Characterization. *AIDS Res. Hum. Retrovir.* **12**:741–747.
  29. Schmidt, W. N., D. Klinzman, D. R. LaBrecque, D. E. Macfarlane, and J. T. Stapleton. 1995. Direct detection of hepatitis C virus (HCV) RNA from whole blood, and comparison with HCV RNA in plasma and peripheral blood mononuclear cells. *J. Med. Virol.* **47**:153–160.
  30. Shrestha, S. M., N. B. Subedi, S. Shrestha, K. G. Maharjan, F. Tsuda, and H. Okamoto. 1998. Epidemiology of hepatitis C virus infection in Nepal. *Trop. Gastroenterol.* **19**:102–104.
  31. Simmonds, P., P. Balfe, C. A. Ludlam, J. O. Bishop, and A. J. Brown. 1990. Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1. *J. Virol.* **64**:5840–5850.
  32. Stapleton, J. T., D. Klinzman, W. N. Schmidt, M. A. Pfaller, P. Wu, D. R. LaBrecque, J. Q. Han, M. J. Phillips, R. Woolson, and B. Alden. 1999. Prospective comparison of whole-blood- and plasma-based hepatitis C virus RNA detection systems: improved detection using whole blood as the source of viral RNA. *J. Clin. Microbiol.* **37**:484–489.
  33. Tripathy, S., B. Renjifo, W. K. Wang, M. F. McLane, R. Bollinger, J. Rodrigues, J. Osterman, S. Tripathy, and M. Essex. 1996. Envelope glycoprotein 120 sequences of primary HIV type 1 isolates from Pune and New Delhi, India. *AIDS Res. Hum. Retrovir.* **12**:1199–1202.
  34. Van Harmelen, J. H., E. Van der Ryst, A. S. Loubser, D. York, S. Madurai, S. Lyons, R. Wood, and C. Williamson. 1999. A predominantly HIV type 1 subtype C-restricted epidemic in South African urban populations. *AIDS Res. Hum. Retrovir.* **15**:395–398.
  35. Yirrell, D. L., P. Robertson, D. J. Goldberg, J. McMenamin, S. Cameron, and A. J. Leigh Brown. 1997. Molecular investigation into outbreak of HIV in a Scottish prison. *BMJ* **314**:1446–1450.
  36. Zhang, L., R. S. Diaz, D. D. Ho, J. W. Mosley, M. P. Busch, and A. Mayer. 1997. Host-specific driving force in human immunodeficiency virus type 1 evolution in vivo. *J. Virol.* **71**:2555–2561.