Phylogenetic relationship between human immunodeficiency virus type 1 (HIV-1) long terminal repeat natural variants present in the lymph node and peripheral blood of three HIV-1-infected individuals

Mounir Ait-Khaled and Vincent C. Emery*

Division of Communicable Diseases, Royal Free Hospital School of Medicine, Rowland Hill Street, Hampstead, London NW3 2QG, U.K.

PCR has been used to amplify a 540 base pair fragment of the human immunodeficiency virus type 1 (HIV-1) long terminal repeat region encompassing the U5/R/U3 regions from proviral HIV DNA that was present in the lymph nodes and peripheral blood mononuclear cells of three HIV-infected individuals. The resulting PCR products were cloned and the DNA sequence of multiple clones from each reaction was determined to enable the distribution and phylogenetic relationship between the variants in each body site to be assessed. The results of bootstrapped parsimony phylogenetic analyses showed that a distinct polarization was evident between peripheral blood and lymph node variants in the patient who possessed a histologically intact lymph node. In contrast, similar analyses conducted on samples from two patients with extensive lymph node disruption showed similar variants in lymph node and peripheral blood. In the patient with an intact lymph node, lymph node variants were significantly more heterogeneous than those present in blood samples taken either simultaneously or 11 months later. No differences in heterogeneity between lymph node and peripheral blood were observed in patients with disrupted lymph nodes. The significance of these results for our understanding of HIV pathogenesis is discussed.

Introduction

The work of Wain-Hobson, Leigh Brown and others has shown that human immunodeficiency virus (HIV) infection is caused by a collection of strains or a quasispecies (Goodenow et al., 1989; Meyerhans et al., 1989; Wain-Hobson, 1989; Simmonds et al., 1990; Balfe et al., 1990; Delassus et al., 1991, 1992). Although the temporal variation between proviral HIV strains in the blood and plasma virus strains has been extensively investigated (Simmonds et al., 1991; Wolfs et al., 1992), relatively little is known about the site(s) of strain evolution or the relationship of variants between various HIV-infected organs. Analysis of the V3/V4 regions of the env gene in brain, spleen and peripheral blood mononuclear cells (PBMCs) has indicated that strain evolution in each infected organ proceeds in parallel possibly reflecting cell tropism of particular variants (Epstein et al., 1991; Pang et al., 1991). Recent data have confirmed and extended earlier work from Janossy and colleagues (Tenner-Racz et al., 1986) highlighting the importance of the lymphoid tissues, particularly the lymph nodes, in HIV pathogenesis (Pantaleo et al., 1993; Embreston et al., 1993). Thus, proviral HIV DNA has been shown to be present in a high proportion of lymph node lymphocytes and macrophages at all stages of HIV infection (Pantaleo et al., 1993; Embreston et al., 1993). Such a situation contrasts with proviral HIV load in the peripheral blood, where a low proportion of CD4+ lymphocytes are infected at early and intermediate stages of HIV disease, although a higher load in the periphery is observed at late stages of disease. One explanation for this effect is the seeding of HIV strains from the lymph node to the periphery following the loss of an effective cellular and humoral immune system and/or extensive destruction of the lymph node structure.

Although previous studies on the tissue tropism of HIV-1 have concentrated on the variation of env sequences in different tissues, a major drawback for these analyses is that env regions are under selective pressure from both B and T cells and may exhibit convergent evolution (Holmes et al., 1992) thus complicating any phylogenetic analysis. In the light of this observation, we chose the long terminal repeat (LTR) region of the HIV genome, since it is not subject to extensive immune response effects. The LTR functions as the promoter--enhancer unit of HIV and comprises multiple functional elements controlling the differential and temporal ex-
pression of viral genes. These protein-binding sites are under the influence of cellular (NFkB, SPI, AP1, TCF-1) and viral (tat, nef) factors (reviewed by Gaynor, 1992). Several cellular transcription factors that bind to the HIV LTR region are induced by activation or differentiation of lymphocytes or monocytes, rendering the HIV genome subject to many of the same regulatory signals involved in the control of cellular gene expression. In addition to the promoter-enhancer region (-158 to -1), the upstream region of the LTR (-420 to -159) has been identified as a negative regulatory element whereas the proximal region contains the tat-responsive TAR sequence spanning positions +1 to +60. At present, HIV-1 LTR strain variation has been analysed only in the peripheral blood (Delassus et al., 1991).

In this study, we report phylogenetic analyses of HIV-1 LTR variants present simultaneously in the lymph node and PBMCs of three HIV-1-infected individuals and relate the results to proximal markers of disease, clinical stage of infection and lymph node histology.

**Methods**

**Patient samples**

**Patient 1.** A PMBC sample and mediastinal lymph node biopsy were taken from patient 1 in December 1990. The patient had unconfirmed lymphoma and infection with *Mycobacterium avium-intracellularare* had been detected. Histological analysis showed that the germinal centres were intact except for a small epithelial cell aggregate. The absolute CD4 lymphocyte count at biopsy was 18 x 10⁶/ℓ. A subsequent PBMC sample was taken from the same individual in November 1991. The absolute CD4 lymphocyte count at this time was 7 x 10⁶/ℓ.

**Patient 2.** PBMC and neck lymph node samples were obtained from patient 2 in December 1992. The patient had received zidovudine for 3 years and had a CD4 lymphocyte count of 220 x 10⁶/ℓ. *Pneumocystis*
lymph node.

DNA was extracted and purified from the lymph node from patient 3 in May 1992. The patient had received no antiretroviral therapy. Histological analysis showed follicular lysis and disruption of the germinal centres of the lymph node.

PCR analysis. PBMC and axillary lymph node samples were obtained from patient 3 in May 1992. The patient had received no antiretroviral therapy and had a CD4 lymphocyte count of 680 × 10^6/l. Toxoplasma myocarditis and extrapulmonary tuberculosis (lymph node) were diagnosed. Histological analysis showed extensive destruction and replacement of the lymph node by granulomas.

DNA extraction. PBMCs were purified by Ficoll-Hyphaque density gradients. DNA was extracted and purified from the lymph node biopsy sample and PBMCs by protease K digestion, phenol extraction and ethanol precipitation using established methodologies (Sambrook et al., 1989).

PCR analysis. PCR reactions were performed on extracted cell DNA (0.5 μg) using a Hybrid DNA thermal cycler as previously described (Fox et al., 1991). Primers consisted of a positive-sense oligonucleotide spanning positions −395 to −375 and a negative-sense oligonucleotide spanning positions +125 to +145 relative to the start of transcription, and were 5'-phosphorylated prior to use. The sequences were 5' GATCTCTAGTTACAGATCA and 5' CACACAAAAGCTTCCCTGA. The PCR product was 540 bp in length (based on the HXB2 prototype sequence). PCR-amplified products were purified using the Gene Clean II kit (Bio 101) according to the manufacturer's instructions. The amplified LTR quasispecies were cloned by blunt-end ligation into lambda phagemid pUC19. Colonies harbouring inserts were identified using standard methods (Sambrook et al., 1989). Plasmid DNA from recombinant colonies was extracted using a standard plasmid mini-preparation procedure (Sambrook et al., 1989) and 5 μg of DNA was used for sequencing by the methods of Chen and Seeburg (as modified by Murphy & Kavanagh, 1988). Ten to 20 clones from each PCR were sequenced. In order to ensure the fidelity of the LTR sequence variation, an LTR clone of each PCR was sequenced by the methods of Chen and Seeburg (as modified by Murphy & Kavanagh, 1988). Ten to 20 clones from each PCR were sequenced. In order to ensure the fidelity of the LTR sequence variation, an LTR clone of each PCR was sequenced.

Patient 3. PBMC and axillary lymph node samples were obtained from patient 3 in May 1992. The patient had received no antiretroviral therapy and had a CD4 lymphocyte count of 680 × 10^6/l. Toxoplasma myocarditis and extrapulmonary tuberculosis (lymph node) were diagnosed. Histological analysis showed extensive destruction and replacement of the lymph node by granulomas.

DNA extraction. PBMCs were purified by Ficoll-Hyphaque density gradients. DNA was extracted and purified from the lymph node biopsy sample and PBMCs by protease K digestion, phenol extraction and ethanol precipitation using established methodologies (Sambrook et al., 1989).

PCR analysis. PCR reactions were performed on extracted cell DNA (0.5 μg) using a Hybrid DNA thermal cycler as previously described (Fox et al., 1991). Primers consisted of a positive-sense oligonucleotide spanning positions −395 to −375 and a negative-sense oligonucleotide spanning positions +125 to +145 relative to the start of transcription, and were 5'-phosphorylated prior to use. The sequences were 5' GATCTCTAGTTACAGATCA and 5' CACACAAAAGCTTCCCTGA. The PCR product was 540 bp in length (based on the HXB2 prototype sequence). PCR-amplified products were purified using the Gene Clean II kit (Bio 101) according to the manufacturer's instructions. The amplified LTR quasispecies were cloned by blunt-end ligation into lambda phagemid pUC19. Colonies harbouring inserts were identified using standard methods (Sambrook et al., 1989). Plasmid DNA from recombinant colonies was extracted using a standard plasmid mini-preparation procedure (Sambrook et al., 1989) and 5 μg of DNA was used for sequencing by the methods of Chen and Seeburg (as modified by Murphy & Kavanagh, 1988). Ten to 20 clones from each PCR were sequenced. In order to ensure the fidelity of the LTR sequence variation, an LTR clone of each PCR was sequenced.
Table 1. Mean divergence between LTR variants present in the lymph node and blood samples of patients 1, 2 and 3

<table>
<thead>
<tr>
<th>Patient</th>
<th>Time (months)</th>
<th>Absolute CD4 count $\times 10^6/l$</th>
<th>Histology</th>
<th>Median divergence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lymph node</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>18</td>
<td>Intact germinal centres</td>
<td>0.03±0.01$^{*}$</td>
</tr>
<tr>
<td></td>
<td>+11</td>
<td>7</td>
<td>–</td>
<td>NA$^{†}$</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>220</td>
<td>Follicular lysis and disruption of germinal centres</td>
<td>0.0131±0.0053$§$</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>680</td>
<td>Extensive destruction and granulomas</td>
<td>0.038±0.0128$|$</td>
</tr>
</tbody>
</table>

*† $P < 0.0001.$

$§$ NA, Not applicable.

$\| P$ is not significant.

DNA sequence was determined for 15 clones. The sequence of each clone was identical to the sequence of the original PCR template.

Comparative analysis of HIV LTR sequences. The sequences were analysed using the Sequence Analysis Software Package (version 7) by Genetic Computers Inc. from the University of Wisconsin's Department of Genetics (Devereux et al., 1984). Multiple sequence alignments were produced using the Clustal software and alignments were produced as PHYLIP-compatible files. The PHYLIP package (version 3.5) was used in all the subsequent phylogenetic analyses (Felsenstein, 1989). The genetic distance between each variant was calculated using a Kimura-2 parameter method (available within the DNADIST program). Maximum parsimony trees were generated using the DNAPARS program. The robustness of the trees was tested by bootstrap resampling of the multiple alignment using SEQBOOT (100 datasets) and consensus trees were generated using CONSENSE.

Statistical analysis. An analysis of the significance of differences in genetic distances as calculated from the output of the PHYLIP DNADIST program was performed using the Mann-Whitney U-test (Bland, 1987).

Results

The HIV-1 LTR region from proviral DNA present in the lymph node and PBMCs of three patients has been amplified by PCR, cloned into the SmaI site of pUC19 and DNA sequences from the resultant clones were determined. In the case of patient 1, DNA samples from 20 clones derived from DNA extracts of the lymph node, simultaneous blood samples and an 11 month follow-up blood sample were sequenced. In patients 2 and 3, DNA from 11 and 10 clones respectively from both lymph node and a simultaneous peripheral blood sample were sequenced. The DNA sequences were aligned using multiple pairwise alignment. A representative multiple sequence alignment of lymph node and peripheral blood variants from patient 1 is shown in Fig. 1.

The multiple sequence alignments were then used as the basis for a calculation of the genetic distance between variants in each compartment for each patient (Table 1). Patient 1 harboured a lymph node HIV-1 quasispecies that was more heterogeneous than either the simultaneous blood quasispecies or the 11 months follow-up blood quasispecies. This patient had a relatively low CD4 cell count ($18 \times 10^6/l$) when the biopsy was taken and was classified as having Center for Disease Control (CDC) stage IV disease, but had an intact lymph node structure. In patient 2, there were no significant differences in the heterogeneity between the lymph node and simultaneous blood quasispecies. This patient had a relatively high CD4 cell count ($220 \times 10^6/l$) when the biopsy was performed and was diagnosed as having CDC stage III disease with extensive disruption of the lymph node structure. In patient 3, there was also no significant difference in the heterogeneity between the lymph node and simultaneous blood quasispecies. This patient had a relatively high CD4 cell count when the
relationships between individual variants within each sample taken from patients we have used DNA parsimony phylogenetic methods with bootstrap resampling. The resulting consensus phylogenetic trees are shown for patients 1, 2 and 3 in Fig. 2, 3 and 4, respectively. In patient 1 (Fig. 2) the consensus tree showed a clear dichotomy between the lymph node and peripheral blood variants. This resulted in a polarized tree with two major branches: 32 out of 40 blood variants were contained in one branch (peripheral blood variants clade), whereas 19 out of 20 lymph node variants formed the other branch (lymph node variants clade). This polarization was detected in over 80% of the bootstrap resamplings. In contrast, the consensus trees constructed for patients 2 and 3 (Fig. 3 and 4) were substantially different from that of patient 1 and demonstrated that there was no evidence of polarization between lymph node and peripheral blood variants. These phylogenetic trees are consistent with the data on genetic heterogeneity shown in Table 1. The phylogenetic trees have also been constructed using alternative phylogenetic approaches, e.g. Fitch-Margoliash, and the polarized topology observed between lymph node and peripheral blood variants in patient 1 remained a consistent feature (data not shown).

Discussion

The HIV-1 LTR variants that are present in the lymph node and peripheral blood of three HIV-1-infected individuals have been analysed. The genetic variation in the LTR quasispecies that are present in the lymph node was significantly higher than that present in the peripheral blood quasispecies in the one patient who possessed a lymph node with an intact follicular dendritic cell network and germinal centres. In the two patients in whom substantial disruption of the lymph node structure had occurred, including follicular lysis and absence of germinal centres, there was no difference in genetic diversity between the two compartments. Lymph nodes have been shown to contain a higher HIV burden than PBMCs (Pantaleo et al., 1991, 1993). Although we have not determined which cells within the lymph node gave rise to the HIV quasispecies analysed in this study, recent results indicate that the majority of cells harbouring HIV proviral DNA are CD4 lymphocytes and macrophages and not follicular dendritic cells (Embretson et al., 1993). Follicular dendritic cells appear to trap virions on their villus processes early in the course of HIV infection and may transmit infection to cells as they migrate through the lymph node. However, further data must be obtained before the dynamics of strain evolution in multiple tissue sites and the crossover of strains between these sites is fully understood.
Phylogenetic analysis indicated that blood and lymph variants independently clustered together in clearly separated clades in patient 1. Moreover, three blood variants that were obtained at the time the lymph node biopsy was performed, three variants from the 11 month follow-up blood sample, and one lymph node variant clustered together, forming a sub-branch in the major lymph node clade. These peripheral blood variants are therefore closely related to the lymph node variants. In contrast to the results obtained for patient 1, a distinct clustering of variants was not observed in patients 2 and 3 whose lymph nodes exhibited extensive histological deterioration. The short inter-node branches observed in the consensus trees for these patients indicate that the LTR variants are closely related, with lymph node and blood variants present in common branches (Fig. 3 and 4). The polarized topology of the consensus tree that was obtained for patient 1 with a significant branch-length separating the lymph node and peripheral blood clades provides data in support of the hypothesis that independent evolution of HIV-1 variants can occur in these two compartments when the lymph node is intact.

The results presented in this study, albeit obtained from a small sample size, suggest that the lymph nodes may be important in the sequestration/evolution of distinct HIV variants. Hence, disruption of the lymph node structure would lead to seeding of HIV variants in the periphery, a prediction which is consistent with the data obtained from patients 2 and 3. However, in order to provide sufficient data for statistically significant conclusions of this type to be drawn, similar studies on a greater number of patients are required. It should be noted that these analyses are complicated by the fact that lymph node biopsies are not routinely taken from all HIV-infected patients, especially in the early stages of disease. Such sampling problems could be overcome by performing a similar analysis in appropriate animal models of HIV e.g. simian immunodeficiency virus-infected macaques. Furthermore, the functional effects of sequence variations within the LTR sequence present in the patients that were analysed and the subsequent disease course should be investigated. We are currently extending research into these areas. In this context, it is interesting to note that the clinical status of patient 1 remains relatively stable despite having low CD4 cell numbers and CDC stage IV disease, whereas patient 2 has had a continuously declining CD4 cell count with a rapid disease course (data not shown).

Further data on more patients must be accrued before the precise role of the lymph nodes in HIV strain evolution and pathogenesis is defined. However, our data illustrate the importance of performing both cross-sectional and longitudinal quasispecies analyses on many target organs within HIV-infected individuals in order to understand the complex interplay between HIV and its host.

We wish to thank the staff and patients of the Royal Free Hospital AIDS clinic for their assistance in obtaining samples for this study, especially Dr Margaret Johnson, and to Dr J. McLaughlin of the Department of Histopathology, Royal Free Hospital School of Medicine. We would also like to thank Dr Peter Balfe for interesting discussions. We are grateful to the Royal Free Hospital School of Medicine Computing Unit and the Oxford Molecular Biology Computing Unit. M.A.-K. is supported by the Algerian Government. Work in the laboratory of V.C.E. is supported by the U.K. MRC.

References


(Received 2 November 1993; Accepted 13 January 1994)