Identification of shared populations of human immunodeficiency virus type 1 infecting microglia and tissue macrophages outside the central nervous system

Citation for published version:

Digital Object Identifier (DOI):
10.1128/JVI.75.23.11686-11699.2001

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Journal of Virology

Publisher Rights Statement:
Copyright © 2001, American Society for Microbiology

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Identification of Shared Populations of Human Immunodeficiency Virus Type 1 Infecting Microglia and Tissue Macrophages outside the Central Nervous System

T. H. WANG,1 Y. K. DONALDSON,1 R. P. BRETTLE,2 J. E. BELL,3 AND P. SIMMONDS1*

Laboratory for Clinical and Molecular Virology, University of Edinburgh, Summerhall, Edinburgh EH9 1QH,1 and Regional Infectious Diseases Unit2 and Department of Neuropathology, University of Edinburgh,3 Western General Hospital, Edinburgh EH4 2UX, United Kingdom

Received 10 May 2001/Accepted 3 August 2001

Infection of microglia and other cells of the macrophage/monocyte lineage in the central nervous system (CNS) by human immunodeficiency virus type 1 (HIV-1) underlies the development of giant cell encephalitis (GCE). It is currently unknown whether GCE depends on the emergence of virus populations specifically adapted to replicate in cells of the monocyte/macrophage lineage and whether this also leads to the specific targeting of macrophages in other nonlymphoid tissues. Autopsy samples from lymph node, brain (frontal region), lung, and full-thickness colon sections were obtained from nine study subjects with GCE and from nine without. The two groups showed no significant differences in CD4 counts, disease progression, or treatment history before death. Genetic relatedness between variants recovered from lymph node and nonlymphoid tissues was assessed by sequence comparison of V3 and p17ex regions using a newly developed method that scores the sample composition at successive nodes in a neighbor-joining tree. The association index enabled objective, numerical comparisons on the degree of tissue compartmentalization to be made. High proviral loads and p24 antigen expression in the brain were confined to the nine individuals with GCE. GCE was also associated with significantly higher proviral loads in colon samples (median of the GCE+ group: 1,010 copies/10⁶ cells; median of the GCE− group, 10/10⁶ cells; P = 0.006). In contrast, there were no significant differences in proviral load between the GCE+ and GCE− groups in lymph node or lung samples, where HIV infection was manifested predominantly by infiltrates of lymphoid cells. V3 sequences from brain samples of individuals with GCE showed the greatest compartmentalization from those of lymph node, although samples from other tissues, particularly the colon, frequently contained variants phylogenetically related to those found in brain. The existence of shared, distinct populations of HIV specifically distributed in cells of the monocyte/macrophage lineage was further indicated by immunocytochemical detection of CD68+, multinucleated giant cells expressing p24 antigen in samples of lung and colon in two individuals with GCE. This study provides the basis for future investigation of possible phenotypic similarities that underline the shared distributions of HIV variants infecting microglia and tissue macrophages outside the CNS.

The cellular tropism of human immunodeficiency virus type 1 (HIV-1) is governed at a variety of entry and postentry steps, including the attachment, fusion, and entry of HIV into the cell, reverse transcription, integration, and gene expression (16, 31, 39). Differences between CD4 lymphocytes, the principal targets of HIV-1 in vivo, and other potential cellular targets for HIV-1, such as macrophages and microglia in the brain, exist at many of these levels, particularly in the expression of CD4 and chemokine coreceptors required for virus entry. The ability of HIV-1 to target and productively infect these different cell types in vivo may therefore depend on strain-specific differences of HIV-1 or on the evolution of adaptive differences during the course of infection.

Primary and laboratory isolates show a wide range of cellular tropisms (including ability to grow in transformed T-cell lines, primary cultures of monocyte-derived macrophages), cytopathology (syncytium induction) and coreceptor usage (CXCR4, CCR5, CCR3). These differences have in the past been linked to variability in the rate of disease progression in HIV-infected individuals, in whom the emergence of CXCR4-using, non-macrophage tropic isolates of HIV-1 is accompanied by a more rapid decline in CD4 lymphocyte numbers and the onset of AIDS-related disease (3, 5, 10, 15, 20, 26, 29, 30, 47, 70, 78). Much less is understood about the existence of differential cellular tropism of HIV variants infecting different anatomical locations and tissue types in vivo, and it is not known whether adaptive changes are responsible for direct virus-mediated outcomes of infection, such as the invasion of the central nervous system (CNS) and the subsequent development of giant cell encephalitis (GCE). It is also unknown whether the ability of HIV-1 to productively infect nonlymphoid tissues, such as the brain, is dependent on the same adaptive changes that underlie the CCR5-using, macrophage-tropic phenotype characterized in in vitro studies.

In this study we have used a combination of immunocytochemical detection of p24 antigen (75), PCR for quantitation of proviral DNA sequences (69), and genetic characterization to examine the cell types and virological characteristics of HIV infecting samples of lung and colon collected at autopsy from a large number of HIV-seropositive individuals. We examined the genetic relationships between HIV variants infecting dif-
ferent cell types in these tissues and those recovered from lymph nodes, where lymphocytes are the predominant cell type infected, and also those present in the brains of individuals with GCE, where the principal target cells are cells of the monocyte/macrophage lineage (infiltrating macrophages and microglia). Our findings provide evidence for the existence of genetically distinct populations of HIV targeting cells of the monocyte/macrophage lineage with shared distributions in the CNS and tissue macrophages in the lung and gastrointestinal (GI) tract. Biological characterization of variants recovered from lymphoid and nonlymphoid cells identified in this study will allow the phenotypic differences underlying the in vivo differences in cellular tropism to be identified.

MATERIALS AND METHODS

Study subjects. All of the tissue samples used in this study were held in the Brain and Tissue Bank of Edinburgh (Western General Hospital, Edinburgh, United Kingdom). Tissue samples, including brain, lung, colon, and lymph node, were obtained from 18 autopsies (NA425, NA021, NA020, NA017, NA272, NA371, NA025, NA446, NA199, NA284, NA116, NA369, NA270, NA308, NA246, NA420, NA118) from the cohort of HIV-infected individuals in Edinburgh. All study subjects died of complications associated with HIV infection, including opportunistic infections or neoplasms. Their ages ranged from 32 to 49, and they had histories of HIV-1 infections lasting 4 to 13 years before death. Other clinical and background information for the study group is listed in Table 1, indicating that this variable was not a confounding factor in the differences observed between the two groups in proviral loads found in autopsy samples of brain and colon.

DNA extraction and provirus quantitation. DNA was extracted from frozen brain, lymph node, colon, and lung tissues as previously described (69). Total DNA concentration was determined by spectrophotometry at 260 nm. Proviral load was determined by previously described limiting-dilution nested PCR (69) using the previously described p17primers with 35S-dATP, the thermosequenase-radiolabeled terminator cycle sequencing reaction technique, as previously described (75). Twenty-four replicates at last positive dilution were used to indicate the minimum proviral load in the sample, assuming a Poisson distribution for each sample by the formula $I - p^d$, where $p$ is the proportion of positive samples and $d$ is the dilution (69). Viral load was expressed as copies per million cells, based on the DNA composition of human diploid cells of 6.6 pg of DNA.

Immunohistochemical examination. Approximately 2 weeks after formalin fixation, autopsy tissues were processed through a routine 41-h program in the Vacuum Infiltration Processor (Tissue Tek), followed by paraffin wax embedding using a Tissue Tek embedding console. Five-micrometer sections of formalin-fixed paraffin-embedded tissues from all of the study subjects were examined by immunohistological staining with anti-HIV-1 p24 antibody (DuPont) and double immunohistochemical detection of p24 antigen from three study subjects, NA425, NA021, NA020, NA017, NA272, NA371, NA025, NA446, NA199, NA284, NA116, NA369, NA270, NA308, NA246, NA420, and NA118 from the cohort of HIV-infected individuals in Edinburgh. All study subjects died of complications associated with HIV infection, including opportunistic infections or neoplasms. Their ages ranged from 32 to 49, and they had histories of HIV-1 infections lasting 4 to 13 years before death. Other clinical and background information for the study group is listed in Table 1, indicating that this variable was not a confounding factor in the differences observed between the two groups in proviral loads found in autopsy samples of brain and colon.

Nucleotide sequencing and analysis. Nucleotide sequences from p17 were amplified using previously described primers (68, 85). For the majority of study subjects, single molecules of HIV-1 provirus were isolated by limiting dilution and sequenced directly using the Thermal-cycle Sequenase kit (Amersham). For seven of the study subjects (NA425, NA021, NA020, NA017, NA272, NA371, and NA025), 1-μg aliquots of extracted DNA were amplified and cloned into pGEM, using T overhangs (pGEM-T easy vector system; Promega). Miniprepped DNA from clones were sequenced using the Sequenase version 2.0 kit (United States Biologicals) following the manufacturer's instructions. To avoid template resampling that may occur when the cloning method is used to obtain sequences from samples with low frequencies of amplifiable sequences, its application was restricted to samples with proviral loads greater than 60 copies per 10^6 cells. Deoxynucleotide sequencing of cloned DNA sequences was carried out using USB sequencing 2.0 kit (Amersham Life Science) with dATP, the thermosequenase-radiolabeled terminator cycle sequencing kit, according to the manufacturer's instructions.

### TABLE 1. Proviral loads in lymphoid and nonlymphoid autopsy samples from study subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>HIV risk factora</th>
<th>No. of lymphocytes (counts/μl) at time of death</th>
<th>Duration of infection</th>
<th>Interval between death and autopsy (h)</th>
<th>V3 predictionb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CD4</td>
<td>CD8</td>
<td>HIV (yr)</td>
<td>AIDS (mo)</td>
</tr>
<tr>
<td>NA371</td>
<td>M</td>
<td>0.5</td>
<td>56</td>
<td>N</td>
<td>4</td>
</tr>
<tr>
<td>NA425</td>
<td>I</td>
<td>1</td>
<td>314</td>
<td>N</td>
<td>13</td>
</tr>
<tr>
<td>NA116</td>
<td>I</td>
<td>4</td>
<td>60</td>
<td>N</td>
<td>13</td>
</tr>
<tr>
<td>NA20</td>
<td>M</td>
<td>8</td>
<td>388</td>
<td>N</td>
<td>13</td>
</tr>
<tr>
<td>NA17</td>
<td>I</td>
<td>17</td>
<td>109</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>NA308</td>
<td>I</td>
<td>23</td>
<td>212</td>
<td>N</td>
<td>4</td>
</tr>
<tr>
<td>NA199</td>
<td>I</td>
<td>28</td>
<td>164</td>
<td>N</td>
<td>12</td>
</tr>
<tr>
<td>NA272</td>
<td>I</td>
<td>105</td>
<td>411</td>
<td>N</td>
<td>13</td>
</tr>
<tr>
<td>NA270</td>
<td>I</td>
<td>246</td>
<td>510</td>
<td>N</td>
<td>13</td>
</tr>
<tr>
<td>Median GCE−</td>
<td></td>
<td>17</td>
<td>212</td>
<td>P</td>
<td>10</td>
</tr>
<tr>
<td>NA38</td>
<td>I</td>
<td>3</td>
<td>210</td>
<td>P</td>
<td>5</td>
</tr>
<tr>
<td>NA420</td>
<td>I</td>
<td>8</td>
<td>110</td>
<td>P</td>
<td>5</td>
</tr>
<tr>
<td>NA246</td>
<td>I</td>
<td>18</td>
<td>110</td>
<td>P</td>
<td>13</td>
</tr>
<tr>
<td>NA284</td>
<td>I</td>
<td>23</td>
<td>440</td>
<td>P</td>
<td>14</td>
</tr>
<tr>
<td>NA369</td>
<td>I</td>
<td>5</td>
<td>10</td>
<td>N</td>
<td>9</td>
</tr>
<tr>
<td>NA118</td>
<td>I</td>
<td>87</td>
<td>1,779</td>
<td>P</td>
<td>13</td>
</tr>
<tr>
<td>NA21</td>
<td>I</td>
<td>137</td>
<td>2,611</td>
<td>P</td>
<td>14</td>
</tr>
<tr>
<td>NA446</td>
<td>I</td>
<td>297</td>
<td>232</td>
<td>P</td>
<td>13</td>
</tr>
<tr>
<td>Median GCE+</td>
<td></td>
<td>23</td>
<td>336</td>
<td>P</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of lymphocytes</th>
<th>GCE diagnosis</th>
<th>Duration of infection</th>
<th>Interval between death and autopsy</th>
<th>V3 prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>(counts/μl) at time of death</td>
<td>HIV (yr)</td>
<td>AIDS (mo)</td>
<td>death and autopsy (h)</td>
<td></td>
</tr>
</tbody>
</table>

a M, Male homosexual; I, injecting drug user; H, heterosexual contact.
b Determined during autopsy examination of the brain. N, negative; P, positive.

c Presence of SI, CXCR4-dependent variants of HIV, predicted from V3 loop sequence (19, 32).
d NK, not known.
Sequences were aligned using the Simmonic 2000 Sequence Editor package. Phylogenetic trees were constructed by the neighbor-joining method using Jukes-Cantor corrected sequence distances in the MEGA package (48). The nucleotide sequences from p17\(^{gag}\) and V3 regions from each of the study subjects were compared with each other and with a range of standard HIV-1 variants. Each set of sequences from the 12 study subjects compared was monophyletic in both genomic regions and distinct from those of the published sequences of subtype B: HIV-1\(_{SF2}\) (K02007), HIV-1\(_{RF}\) (M17451), HIV-1\(_{YU2}\) (M74978), HIV-1\(_{YU2}\) (M93258), HIV-1\(_{CAM1}\) (D10112), HIV-1\(_{YU2}\) (M38431), HIV-1\(_{HAN}\) (U43131), HIV-1\(_{WM2G}\) (M12507), and HIV-1\(_{PFAA}\) (M65024). This comparison provided no evidence for coinfection with epidemiologically unrelated HIV strains, nor for intersample or exogenous laboratory contamination.

Analysis of phylogenetic groupings. The degree of genetic segregation between variants recovered from different samples was scored using a novel method for scoring phylogenetic trees. For each sequence comparison, a phylogeny was calculated with the programs DNADIST and NEIGHBOR in the PHYLIP package (28) using an epidemiologically unlinked sequence (HIV-1\(_{SF2}\)) as an outgroup. Starting from the root of the tree, the composition of sequences in each successive bifurcating node was calculated. An association value, \(d\), for the tree was calculated by summation of values individually calculated from each node, according to the formula 
\[
\frac{1}{H} \sum \frac{n}{H} \left(1 - \frac{f}{n}\right),
\]
where \(n\) is the number of sequences below the node and \(f\) is the frequency of most common sample type. Values of \(d\) expected from the null hypothesis (i.e., for samples showing no phylogenetic grouping) were calculated by random reassignment of the sequences to different samples. Finally, the influence of tree robustness on the association value was indicated by bootstrap resampling using the program SEQBOOT in the PHYLIP package. The association index (AI) represents the mean ratio of 100 bootstrap replicates of the association value calculated from the test sequences to those of 10 sample-reassigned controls.

Confidence intervals for AI values were difficult to calculate from first principles, as the variance depended on the number of sequences compared, the degree of sequence divergence, and the phylogeny of the sequences. For the purpose of this study, in which each comparison of tissue samples contained similar numbers of sequence with similar divergence, we estimated confidence intervals empirically for AI values using (sample-reassigned) control values from independent lymph node and brain comparisons from 10 of the study subjects. The distribution of AI values corresponded closely to a normal distribution around 1, and confidence intervals for 2 standard deviations ranged from 0.74 to 1.36.

Nucleotide sequence accession numbers. The nucleotide sequences obtained in the present study have been submitted to GenBank and have been assigned accession no. AF353734 through AF35394 and AF409200 through AF409685.

RESULTS

Detection of HIV-1 proviral sequences in nonlymphoid tissue. DNA was extracted from the left frontal region of the cerebral cortex in the brain, from lung, and from full-thickness autopsy samples of colon from 18 HIV-infected individuals. Proviral loads were measured by limiting-dilution PCR using highly conserved primers from the p17\(^{gag}\) region (Table 1). Proviral loads were measured by limiting-dilution PCR using highly conserved primers from the p17\(^{gag}\) region (Table 1). High frequencies of infected cells in the lymph node samples were detected in most individuals (range, 15 to 110,000 proviral copies/10\(^6\) cells).

Proviral loads in nonlymphoid tissue were more variable. In brain, high proviral loads were observed only in study subjects with a postmortem diagnosis of GCE (median proviral load, 540/10\(^6\) cells compared with 6/10\(^6\) cells in those without en-
HIV-1 p24 antigen was detected in lung samples of 6 of the 18 study subjects examined. p24 detection was more frequent in samples with high proviral loads (Table 1; \( P = 0.005 \); Spearman’s rank correlation test). p24 detection in lung was generally restricted to germinal centers in peribronchial lymphoid aggregates, and most of the p24-positive cells were morphologically mononucleated (Fig. 2E). However, a large number of infected MGCs which double stained with anti-p24 MAb and the PG-M1 marker, were observed in subject NA246 (Fig. 2F).

p24 immunopositivity was detected in colon samples from only two study subjects, in differing cell types. In the sample from NA25, p24 was detected in lymphoid infiltrate in the colon wall (Table 2). In contrast, p24 antigen expression was detected in both the lymphoid infiltrates (Fig. 2G) and CD68-positive cells with a tissue macrophage morphology in the colon sample from NA284 (Fig. 2H). In this sample, infected cells frequently formed pronounced multinucleated syncytia, reminiscent of the appearance of infected macrophages in the lung and microglia in the brain (Fig. 2D and F).

**Nucleotide sequence analysis.** The association between GCE and high proviral loads in nonlymphoid tissue and the histological evidence for infection of cells of the monocyte/macrophage lineage in lung and colon in NA284 and NA246 suggested that HIV populations in nonlymphoid tissue may be related to those infecting the brain. To investigate genetic relationships between variants in different tissues, we sequenced the V3 hypervariable and flanking regions from proviral sequences recovered from lymph node tissue of each study subject and, where appropriate, from brain, lung, and colon tissues. The proviral loads determined by limiting dilution were essential in interpreting the sequencing results obtained, in view of the possibility that residual blood contamination (particularly in highly vascular tissue such as the lung) may contribute partly or entirely to the sequences amplified by PCR.

Problems associated with analysis of sequences to define distributions of sequences from different samples are the arbitrariness of the phylogenetic groups assigned and the absence of any defensible statistical test to detect differences in the

---

**TABLE 2. Detection of p24 antigen in lymphoid and nonlymphoid autopsy samples from study subjects**

<table>
<thead>
<tr>
<th>Subject</th>
<th>No. of CD4 (counts/μl)</th>
<th>GCE diagnosis</th>
<th>LN&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Nonlymphoid samples&lt;sup&gt;b&lt;/sup&gt;</th>
<th>LF</th>
<th>CO</th>
<th>LU</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA371</td>
<td>0.5</td>
<td>N</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA425</td>
<td>1</td>
<td>N</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA116</td>
<td>4</td>
<td>N</td>
<td>+ +</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA20</td>
<td>8</td>
<td>N</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA17</td>
<td>17</td>
<td>N</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA308</td>
<td>23</td>
<td>N</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA199</td>
<td>28</td>
<td>N</td>
<td>+ + +</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA272</td>
<td>105</td>
<td>N</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA270</td>
<td>246</td>
<td>N</td>
<td>+ + +</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA38</td>
<td>3</td>
<td>P</td>
<td>+ + +</td>
<td>⊕⊕⊕ (sparse)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA420</td>
<td>7</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA246</td>
<td>8</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA284</td>
<td>8</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>+ ⊕ ⊕⊕</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA360</td>
<td>23</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA118</td>
<td>56</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>+ ⊕ ⊕⊕</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA21</td>
<td>87</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA446</td>
<td>137</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NA25</td>
<td>297</td>
<td>P</td>
<td>+</td>
<td>⊕⊕⊕ (frequent)</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<sup>a</sup> LN, lymph node; N, negative; P, positive.

<sup>b</sup> Detection of p24 antigen in autopsy samples by immunocytochemistry; positive samples were graded on a scale from + (sparse) to + + + (frequent), based on numbers of infected cells.

<sup>c</sup> p24 antigen expression detected in cells of monocyte/macrophage lineages is indicated on a scale of ⊕ (sparse), to ⊕⊕⊕ (frequent). LF, left frontal region of the brain; CO, colon; LU, lung.

<sup>d</sup> –, not detected.

cerebral and cerebellar inflammation and gliosis; correlation coefficient = 0.629 by Spearman’s rank correlation test). p24 detection in lung was generally restricted to germinal centers in peribronchial lymphoid aggregates, and most of the p24-positive cells were morphologically mononucleated (Fig. 2E). However, a large number of infected MGCs which double stained with anti-p24 MAb and the PG-M1 marker, were observed in subject NA246 (Fig. 2F).
composition of the clades. We have developed a method to analyze the large amount of sequence information generated in the genetic comparison of V3 sequences from different tissues from the 18 study subjects. The method is based on the detection of deviation from randomness in the position of sequences from different tissues in phylogenetic trees. A number of methods can be used to score the grouping of sequences in trees; in this study, we calculated a numerical index (grouping value) that is derived from the composition of descendants from each bifurcating node in a standard neighbor-joining tree. Trees showing a high degree of compartmentalization of sequences would contain relatively few nodes with descendants from different samples (in this case, from lymph node, brain, lung, or colon), and would be assigned a lower grouping score than for one where each descendant node contained mixed variants. The grouping value for a sequence dataset without compartmentalization (the null hypothesis) is also influenced by the number of sequences analyzed and the shape of phylogenetic tree. Therefore, compartmentalization has to be demonstrated by comparison of the grouping value with that calculated for a control dataset that retains the same phylogeny. The most convenient way to achieve this computationally is to randomly reassign the sequence labels, while retaining the relative numbers of sequences from each compartment and the tree structure. In the analysis presented here, a mean value from 10 label reassignments was used for each control value.

Finally, the grouping value is also dependent on the robustness of the tree analyzed. For this reason, grouping values were calculated on 100 sets of bootstrap-resampled replicates of the sequences, each with 10 sample label reassignments. Sequence datasets with grouping values consistently lower than label-reassigned controls therefore showed evidence for tissue-specific compartmentalization. Although the method can detect groupings in datasets containing sequences from any number of tissue compartments, grouping values are most relevant for comparing pairs of samples (such as brain and lymph node). To allow for different values from control sequences, results from the association test are expressed as an association index, AI, which is the ratio of the association value of sequences over that of controls). Examples of phylogenetic trees and their derived AI values for sequence comparisons between lymphoid and nonlymphoid tissues are shown in Fig. 3.

**Comparisons of HIV proviral sequences from different tissues.** Approximately 10 V3 and flanking region sequences were obtained from the frontal region tissues of the nine study subjects with GCE. These sequences were compared with those amplified from lymph node (Table 3). With one exception (NA38; AI value, 0.52), brain V3 sequences were phylogenetically distinct from the lymphoid HIV population (AIs from 0.03 to 0.0002).

For other tissues, phylogenetic relationships with lymphoid populations were more variable. Most samples from lung autopsy tissue contained proviral populations that were interspersed with lymph node sequences (Table 3), with only two study subjects showing clearly distinct lung populations. All lung-derived sequences from NA25 segregated separately from lymph node sequences (Table 3), while lung sequences from NA246 comprised two viral populations, one interspersed with lymph node sequences and another grouping separately (see below). There was no association between proviral loads in lung and their degree of sequence relatedness to lymphoid-cell populations (Tables 1 and 3).

Sequences from colon also showed variable relationships with the lymphoid populations; approximately half of the colon samples (from NA420, NA246, NA284, NA21, and NA25, all with evidence of GCE) contained sequences grouping separately from those in lymph node samples (Table 3). There was a tendency for colon samples with high viral loads to show greater genetic segregation, although this difference did not reach statistical significance.

To provide independent evidence for genetic segregation of proviral populations in lymphoid and nonlymphoid cells, samples from seven study subjects were sequenced in the p17resg region and association indices were calculated (Fig. 4). There was a close correlation between sequence relationships from the two genomic regions ($R = 0.71$; $P = 0.0002$ by Spearman’s rank correlation test), demonstrating substantial concordance about which populations were genetically distinct and which were interspersed. An outlier from this association was NA425, for which samples from colon were genetically distinct from lymph node samples in the p17resg region (AI, 0.0022) but interspersed in V3 (AI, 0.41).

**Genetic relationships between nonlymphoid populations.** The combination of the previously described findings suggests that virus populations in the brain (in which the principal target cells are microglia; Fig. 2C and D) may be genetically linked to those in nonlymphoid populations, particularly in samples with infection of tissue macrophages (NA284 and NA246). In all study subjects with GCE, except for NA38, variants infecting the brain were distinct from those in lymphoid tissue, potentially allowing the identification of macrophage-associated populations elsewhere. Analysis was concentrated on the four individuals with evidence for infection of the colon and/or lung with variants of HIV genetically distinct from those in lymph node (NA420, NA246, NA284, and NA25; Fig. 5). All four individuals showed evidence for GCE, and comparison was therefore also made with variants amplified from the left frontal region.

V3 sequences from lymph node, brain, lung, and colon of NA246 grouped into two main phylogenetic groups. One contained exclusively variants from nonlymphoid tissue, with sequences from brain, colon, and lung interspersed with each

---

**FIG. 2.** Identification of HIV-expressing cells from different tissues in vivo. (A and B) Serial sections of lymph node from NA25 immunostained with diaminobenzidine, using monoclonal antibodies to CD21 (A) and p24 antigen (B), localizing HIV to follicle center. (C) Low-power magnification of same section as for panel C, costained for p24 antigen with DAB and for CD68 with new fuchsin red. (D) Higher magnification of section of brain tissue from NA21 immunostained for p24 antigen, showing cluster of p24 antigen-expressing cells. (E) Lung tissue from NA21 showing p24 antigen expression in lymphoid follicle. (F) Same section as for panel E, identifying p24 expression (DAB) on MGCs expressing CD68 (new fuchsin red). (G) Colon tissue from NA284 showing p24 antigen expression in lymphoid follicle. (H) Same section as for panel G, identifying p24 expression (DAB) on MGCs expressing CD68 (new fuchsin red).
other. The other lineage contained predominantly lymph node-derived variant, with some from lung. The genetic division between lymphoid- and nonlymphoid-derived variants was also observed for NA25; this subject’s sequences from brain, colon, and lung were mixed in the two clades which contained exclusively nonlymphoid tissue variants. Sequence relationships for the remaining two individuals were more complex. Sequences from brain, colon, and lymph node from NA246

FIG. 3. Examples of phylogenetic comparisons between sequences from lymph node with brain (left), lung (center), and colon (right) from samples showing high (upper panel) and low (lower panel) AI values. Trees were constructed by the neighbor-joining method using the MEGA package; frequency of bootstrap replicates supporting individual clades is indicated on branches (only values ≥70% are shown). All trees were plotted to the scale indicated by the bar at the bottom of the figure.
each grouped separately from each other, with a colon-specific population distinct from both lymphoid variants and those from brain. Finally, most colon-derived variants from NA284 formed a separate clade which also contained lymph node sequences. These sequences shared a basic amino acid at position 320 in the V3 loop, associated with a CXCR4-dependent phenotype (Fig. 6). The remaining colon sequences grouped in a separate clade with a colon-specific amino acid at a particular site or combination of sites. From individuals with and without GCE in the distribution of residues differentiating lymphoid and nonlymphoid sample regions.

Sequences in the p17 gag regions were obtained from NA284 and NA25 (Fig. 5). Sequence relationships in the p17 gag region were generally congruent with those in V3, with genetic compartmentalization of lymph node sequences similar to those derived from nonlymphoid tissue. A noticeable difference was the grouping of colon-derived sequences from NA284. While the majority of V3 sequences grouped in a separate clade with lymph node sequences, there was no evidence for shared grouping of colon and lymph node sequences in the p17 gag region. Instead, all sequences from the colon were interspersed with those from the brain.

Summarizing the results of this comparison, most variants from colon and lung showed common genetic lineages with proviral sequences recovered from brain, while sequences resembling those from lymphoid tissue were generally in the minority in both V3 and p17 gag regions. However, sequences from nonlymphoid tissue were not always monophyletic, and in the case of NA246 there was a tripartite division between brain-, colon-, and lymph node-derived variants. Taken together, however, the comparisons reveal a closer genetic relationship between sequences within nonlymphoid tissue than with those derived from infected lymphocytes.

Comparison of V3 hypervariable region amino acid sequences. To investigate whether variants from nonlymphoid tissue shared sequence motifs in the V3 region which distinguished them from lymphoid populations, a consensus amino acid sequence for variants from each tissue from the four study subjects was calculated (Fig. 6A). Each consensus sequence in this comparison showed acidic or neutral residues at positions 306 and 320 in the V3 loop. By chi² analysis, there were no residues differentiating lymphoid and nonlymphoid sample types or individual nonlymphoid tissues that were shared between different study subjects in the V3 loop or flank regions.

To investigate whether the occurrence of GCE was associated with particular sequences in the V3 region, consensus sequences from lymph node were calculated from each study subject. This analysis was complicated by the occurrence of various proportions of V3 sequences associated with CXCR4 usage in samples from six of the study subjects (Table 1). As these are genetically distinct from viruses with predicted CCR5 usage, separate V3 consensus sequences were calculated for each (Fig. 6B and C). For either predicted phenotype, there was no evidence for any systematic difference between sequences from individuals with and without GCE in the distribution of amino acids at a particular site or combination of sites.

### DISCUSSION

Limiting-dilution PCR was used to quantify proviral sequences in a range of lymphoid and nonlymphoid autopsy tissues from HIV-infected individuals. Consistent with previous studies (1, 7–9, 25, 57), there was a strong association between the development of HIV-related GCE and proviral load measured in a representative sample of the brain. For the subjects analyzed in this study, ranges of proviral loads between GCE and GCE groups were almost nonoverlapping (Fig. 1). Identification of infected cells in the brain and in other tissues was carried out by immunocytochemical staining for p24 antigen expression and was therefore restricted to cells containing actively replicating virus. p24 antigen staining was
FIG. 5. Phylogenetic comparison of lymph node, lung, colon, and brain sequences from the four study subjects with evidence for separate proviral populations in colon and/or lung samples. p17\textsuperscript{gag} region sequences from NA284 and NA25 are placed in reversed orientation on the right. Trees were constructed by the neighbor-joining method using the MEGA package; the frequencies of bootstrap replicates supporting individual clades are indicated on branches (only values $\geq 70\%$ are shown). All trees were plotted to the scale indicated by the bar at the bottom of the figure.
Localized to CD68+ cells of the macrophage/monocyte lineage in the CNS. Morphologically, CD68+ cells corresponded both to perivascularg microglia and macrophages and to microglia in the brain parenchyma, although more precise identification awaits the development of markers to distinguish cells of this lineage. None of the findings presented in this paper rule out the possibility for additional restricted or latent infection by HIV of other cell types in the CNS. For example, there may be nonproductive infection of astrocytes (reviewed in reference 11), associated with the expression of nonstructural gene products such as Nef (21, 44, 65), although whether they contribute significantly to the proviral population detected by PCR awaits more precise quantitative studies based on methods such as in situ PCR (4, 76) which do not rely on gene expression for detection of infected cells.

CNS viral populations in brain autopsy samples from the study subjects were relatively uniform in the V3 region of env and genetically distinct from variants infecting lymph node tissue in eight of the nine study subjects (Fig. 3 and 6), consistent with a large number of previous studies (6, 22, 24, 27, 33, 34, 36, 43, 45, 52, 53, 56, 58, 62, 63, 74, 81). While this overwhelming body of evidence of compartmentalization suggests the existence of adaptive differences involved in infection of the CNS, it has been alternatively hypothesized that differences in the rates of virus turnover in different cell types may lead to the population differences observed between brain and lymph node, given the rapid temporal change in HIV populations over time in peripheral blood mononuclear cells and other lymphoid cell types (45). However, the strict association between proviral load and the detection of replicating virus in the CNS by immunocytochemistry and the likelihood of continuous trafficking of HIV-infected cells through the blood brain barrier and intermingling of HIV variants at this interface (59, 83) are inconsistent with the concept that CNS-derived sequences represent an inactive or slower-replicating archival HIV population. Further indirect evidence for a functional difference with lymphoid- and CNS-derived virus variants is provided by the comparison of sequence relationships in the env gene with regions elsewhere in the genome not involved in cellular tropism, such as gag (41, 56). To summarize a large amount of comparative sequence information, only sequences derived from the V1/V2 and V3 regions demonstrated consistent tissue specificity; brain-derived sequences in the p17gag region showed much greater sequence diversity and in some individuals or from certain brain regions may be genetically indistinct from those recovered from lymphoid tissue. Greater diversity and weaker partitioning are also evident from the comparison of V3 and p17gag sequences in the present...
study (Fig. 5). Given the propensity of HIV to frequently recombine in vivo, the restriction of tissue specificity to V1/V2 and V3 provides evidence for phenotypic selection in this region not exerted elsewhere in the genome.

Direct evidence for differences in cellular tropism between HIV variants recovered from CNS and lymphoid tissues has been obtained by comparison of the biological properties of virus isolated from the CNS and virus isolates passaged through microglia or through the characterization of infectious clones or pseudotypes constructed from env gene sequences amplified from brain and lymphoid tissue (13, 23, 35, 55, 67, 72, 84). All brain-derived or microglia-passaged variants analyzed to date require CD4 for virus attachment and use predominantly or exclusively the chemokine receptor CCR5 for entry (2, 13, 18, 23, 37, 38, 51, 67, 72, 84). This clearly is not the defining feature of CNS-derived viruses, as the majority of HIV variants derived from lymphoid cell types also use CCR5. Indeed, while there is a consensus that CCR5 is the principal coreceptor by which HIV enters cells of the macrophage/monocyte lineage, additional phenotypic differences are likely to underlie the specific macrophage tropism demonstrated by certain HIV-1 isolates.

Evidence supporting the hypothesis for additional factor(s) required for productive infection of macrophages includes the observations that many CCR5-using isolates of HIV fail to infect macrophages in vitro (14, 18, 23, 42, 71). Macrophages can be infected in vitro through the CXCR4 receptor by dual-tropic isolates of HIV (17, 38), while still retaining an apparent preferential tropism for macrophages (18). In vitro passaging of a blood-derived CCR5-using isolate of HIV-1 in microglia led to the emergence of a variant with an enhanced ability to replicate in macrophages, a change in phenotype dependent on amino acid changes in the V1/V2 hypervariable region (55, 67). These changes had no apparent effect on its usage of CCR5 or other coreceptors but enhanced its ability to use low cell surface concentration of CD4 for attachment. This may have direct adaptive advantages for infection of cells of the macrophage/monocyte lineage, where levels of CD4 expression are lower than on the surfaces of CD4 lymphocytes (55, 60, 67).

Together, these observations indicate the complexity of the HIV-macrophage interaction and support the original hypothesis for multiple entry or postentry restrictions to their productive infection (31), additional to requirements for CD4 and CCR5 cell surface expression. Adaptive changes required for replication in macrophages, and not coreceptor usage, may therefore underlie the consistent sequence differences between brain and lymph node populations in this study and previous genetic analyses of autopsy-derived tissues. This is particularly relevant for published comparisons of brain- and lymph node-derived variants of HIV-1 in which there are no differences in net V3 charge that might suggest use of different coreceptors.

A close association between macrophage tropism and the ability of isolates to infect microglia has been long recognized (35, 46, 66, 77, 82), although whether infection of microglia or neurovirulence requires additional adaptive changes remains unclear. In the simian immunodeficiency virus macaque model of AIDS, it has been shown that the gp120 gene determines macrophage tropism, while sequences in gp41 and/or nef are required for neurotropism (54). The lack of an animal model has precluded comparable investigations of neurovirulence of HIV-1, although there is evidence for similar or identical replication kinetics of various CCR5-using isolates in macrophages and microglia in vitro (35, 38, 77), indicating similarity in the phenotypic determinants underlying tropism for the two cell types. The development of GCE may therefore reflect the evolution of HIV variants, capable of infecting not only cells of the CNS but also macrophage populations in other nonlymphoid tissues. The relationship between the development of GCE and infection of macrophage populations elsewhere in the body was the focus of our quantitative and genetic comparison of HIV proviral sequences in different autopsy tissues.

Relationship between infections of the CNS and of other nonlymphoid tissues. The most striking link between the occurrence of GCE and infection of other nonlymphoid tissues was the observation of significantly higher proviral loads in colon samples of the study subjects with GCE (Fig. 1). This phenomenon was not the result of differences in disease progression, treatment history, or extent of destruction of lymphoid tissue between the GCE+ and GCE− groups on autopsy examination (data not shown). Similarly, there was no significant difference in CD4 count before death, nor were there differences in proviral loads in lymph node samples between the two groups (Fig. 2). These observations led us to the hypothesis that the development of GCE reflects a broader tendency of some virus infections to spread beyond lymphocytes to other cell types in nonlymphoid tissues.

The finding of increased proviral loads in colon samples in the GCE+ group was associated with the frequent detection of V3 sequences closely related to those found in the CNS and distinct from those in lymphoid tissues (NA246, NA284, NA25; Fig. 5). The finding of virus populations distinct from lymphoid tissue in some colon autopsy specimens is consistent with previous observations for partial genetic partitioning of HIV between GI- and blood-derived samples. For example, different frequencies of antiviral resistant variants in gut mucosal biopsy samples with circulating virus in plasma were obtained from a minority of individuals (61). Similarly, V3 sequence differences were frequently found between variants of HIV recovered from feces and those found in peripheral blood (79, 80).

In one case, it was possible to detect p24 antigen-positive CD68+ MGCs in the mucosal layer of the colon, representing productive infection of tissue macrophages. The associated syncytial cytopathology was strikingly reminiscent of giant cells associated with GCE in the brain. V3 sequences from the colon of this individual comprised two populations; one group contained positively charged residues at positions 322 and 324 and grouped with lymph node-derived variants (Fig. 5, 6). The others were closely similar to those recovered from the CNS (Fig. 5). As this sample showed both macrophage infection and lymphoid cell infiltration (Table 2), it is possible that the former cell types harbored the HIV variants similar to those in the CNS. We are currently using microdissection to allow the separate genetic analysis of macrophage- and lymphoid-associated infected cells in this autopsy sample to confirm this hypothesis. Despite this example, detection of productively infected macrophages in the GI tract by p24 antigen immunocytochemistry was infrequent in this series and has been previously described in only one case report (49). It is possible that restricted replication of HIV in macrophages prevented the detection of infected macrophages in other colon samples.
where there was evidence for the presence of CNS-related variants. Restricted or latent infection may result from the reduced expression of CCR5 found in gut-derived macrophages compared with those derived from blood monocytes or in microglia (50, 73). More extensive analysis of sections by immunocytochemistry or the application of in situ PCR to identify latently infected cells may help resolve this issue.

In contrast to infection of the colon, that of the lung was almost invariably the result of lymphoid infiltration. There was no association between proviral load in lung and the development of GCE (Fig. 1), and V3 and p17\textsuperscript{pol} sequences from variants infecting lung tissue were generally interspersed with those from lymph node (Table 3). The evidence for lack of genetic partitioning between lymph node- and lung-derived HIV variants is consistent with previously observed similarities between lymph node and lung env sequences (71) and suggests extensive trafficking and infection of similar cell types between the two tissue types. Supporting this, p24 antigen detection was, with one exception, confined to lymphoid follicle centers (Table 2).

However, the sample from NA246 showed a high frequency of p24 antigen-positive CD68\textsuperscript{+} giant cells, representing the productive infection of alveolar macrophages. This histological picture was remarkably similar to the observed infection of infected macrophages in the colon samples from NA284 (see above). Genetic analysis of V3 sequences from the lung sample of NA246 revealed two distinct virus populations; the majority form (6 of 10) grouped with brain-derived variants, while two were interspersed with lymph node-derived sequences. As with the colon samples, it is possible that these different populations of lung-derived variants have specific associations with the observed giant cells in the alveoli and the lymphoid cells infiltrating the lung.

**Mechanism of infection of nonlymphoid tissue.** The combined genetic and immunohistochemical findings in this study provide evidence of at least two distinct mechanisms for infection of nonlymphoid tissue. Infiltration of lymphoid cells, perhaps in response to inflammatory processes associated with opportunistic infection, was most evident in the lung and was observed to a lesser extent in the colon, but it was not a feature of the lymphoid infiltrates in the CNS. Infection of nonlymphoid tissue was alternatively or additionally manifested by the appearance of productively infected CD68\textsuperscript{+} cells of the macrophage/monocyte lineage, in which giant cell formation was a marked histological feature. This pattern of infection typified that of the CNS but was also found in colon and lung samples of two study subjects, both of which harbored virus populations closely related to those found in the brain.

This study forms the basis for future functional analysis of macrophage-associated populations in vivo. In particular, it will be important to determine whether the colon and lung-associated variants from NA284 and NA21 show evidence for greater phenotypic similarities to CNS-derived variants than the lymphoid-related variants recovered from corresponding samples of other study subjects. Secondly, it will be relevant to establish whether the particular ability of certain variants of HIV to infect cells of the macrophage/monocyte lineage reflects larger-scale epidemiological differences between HIV variants infecting injecting drug users (IDUs) in Edinburgh and other risk groups. HIV infection in the majority of Edinburgh IDUs, including those described in the present study, originated from a common source introduced into a specific suburb of Edinburgh around 1982-1983 (12, 40, 64). Despite careful analysis of confounding factors, such as treatment compliance and risk-taking behavior, it has remained difficult to account for the much higher incidence of GCE in this group than in other risk groups in Edinburgh or in IDU cohorts in other cities (8, 9). One possibility is that the virus introduced into this group has a greater propensity for infection of macrophages than other subtype B variants of HIV. Phenotypic analysis of the variants genetically characterized in the present study, with particular attention to the existence of specific adaptive changes, such as efficient use of lower cell surface CD4 concentration for attachment previously associated with macrophage tropism (55, 60, 67), will provide an important test of this theory and an insight into the disease manifestations outside the CNS in this group.

**ACKNOWLEDGMENTS**

We are grateful to Francis Brennan for preparation of the sections for immunocytochemistry and for the provision of frozen samples from the MRC Edinburgh HIV Brain and Tissue Bank (SG9708080).

The virological studies were supported by a Medical Research Council strategic project grant (G9632414).

**REFERENCES**


63. Reddy, R. T., C. L. Achim, D. A. Sirko, S. Tehranchi, F. G. Kraus, F. 

64. Robertson, J. R., A. B. V. Bucknall, P. D. Welsby, J. J. K. Roberts, J. M. 


77. Strizki, J. M., A. V. Albright, H. Sheng, M. O’Connor, L. Perrin, and F. Gonzalez-Scarano. 1996. Infection of primary human microglia and mono-


79. Vanderheide, L., C. A. Sol, J. Maas, V. V. Lukashov, C. L. Kuiken, and J. Goudsmit. 1998. Genetic differences between human immunodeficiency vi-


