Viruses encounter changing selective pressures during transmission between hosts, including host-specific immune responses and potentially varying functional demands on specific proteins. The human immunodeficiency virus type 1 Nef protein performs several functions potentially important for successful infection, including immune escape via down-regulation of class I major histocompatibility complex (MHC-I) and direct enhancement of viral infectivity and replication. Nef is also a major target of the host cytotoxic T-lymphocyte (CTL) response.

To examine the impact of changing selective pressures on Nef functions following sexual transmission, we analyzed genetic and functional changes in nef clones from six transmission events. Phylogenetic analyses indicated that the diversity of nef was similar in both sources and acutely infected recipients, the patterns of selection across transmission were variable, and regions of Nef associated with distinct functions evolved similarly in sources and recipients. These results weighed against the selection of specific Nef functions by transmission or during acute infection. Measurement of Nef function provided no evidence that the down-regulation of either CD4 or MHC-I was optimized by transmission or during acute infection, although rare nef clones from sources that were impaired in these activities were not detected in recipients.

Nef-specific CTL activity was detected as early as 3 weeks after infection and appeared to be an evolutionary force driving the diversification of nef. Despite the change in selective pressure between the source and recipient immune systems and concomitant genetic diversity, the majority of Nef proteins maintained robust abilities to down-regulate MHC-I and CD4. These data suggest that both functions are important for the successful establishment of infection in a new host.

Sexual transmission is the primary route of new human immunodeficiency virus (HIV) infections worldwide. Although the source partner often has a chronic infection characterized by a polyclonal viral population, acute infection in the recipient partner is characterized by a mono- or oligoclonal population. This reduction in viral diversity in acutely infected recipients compared to the chronically infected sources appears to result from a transmission bottleneck. Possible factors contributing to this bottleneck include anatomical compartmentalization of virus in the genital tract of the source partner, a founder effect, or a release of diversifying immune pressures, resulting in reversion to a “wild-type,” more fit, virus.

In contrast to the reduction in the diversity of env (6, 10), the diversity of nef is maintained, at least during vertical transmission, indicating either the transmission of multiple strains or the rapid diversification of the quasispecies in the new host (40). This observation may indicate that transmission and acute infection do not exert new functional constraints upon Nef relative to chronic infection. Alternatively, Nef may maintain function despite diversification driven by selective pressures in the new host more easily than other viral proteins such as Env.

The functions of Nef include direct enhancement of viral replication and immune evasion. The direct effect of Nef on viral replication appears to be multifaceted, involving optimization of signaling pathways in T cells and an increase in the infectivity of virions, mediated in part by the down-regulation of CD4 (reviewed in reference 7). The interaction of Nef with the host immune system is also multifaceted. First, Nef-mediated down-modulation of class I major histocompatibility complex (MHC-I) from the surface of infected cells mitigates the cytotoxic T-lymphocyte (CTL) response (7a, 40a). Second, Nef itself is highly immunogenic, and MHC-I-restricted immune responses detected early in infection predominantly target the Nef protein (27). Escape mutations within these epitopes allow immune evasion and drive evolution of the nef sequence (3, 11, 16, 19, 26, 30).

Interestingly, these mechanisms of immune evasion may be at odds with each other; mutations that confer escape from CTL surveillance might interfere with the down-regulation of MHC-I. Such a cost of CTL escape has been observed in vitro by propagating HIV type 1 (HIV-1) in the presence of CTL clones specific for Nef (2). Similarly, mutations associated with
CTL escape could diminish other functions of Nef that directly affect viral replication and infectivity, such as the down-regulation of CD4. As a small protein with multiple functions, Nef might be especially sensitive to the fitness costs of mutations associated with CTL escape (1).

While Nef-mediated evasion of the CTL response likely contributes to the establishment and maintenance of chronic infection, the direct effects of Nef on viral infectivity and replication could increase the efficiency of transmission. These direct virologic effects of Nef could be optimized by the transmission event via at least two mechanisms. First, the absence of CTL-mediated selection pressure in the recipient (before the onset of adaptive immunity) could allow reversion of escape mutations generated in the source, yielding a more functional nef sequence. Second, the initial lack of CTL activity could allow selection of Nef proteins that are optimized for enhancement of viral infectivity and replication at the expense of down-regulation of MHC-I. Such a trade-off in Nef activities has been observed in the case of advanced chronic infection, when CTL responses have presumably faltered (5), but whether this occurs during acute infection is unknown. Notably, nef-defective HIV-1 can be transmitted, both parenterally and sexually (9, 39).

In this study, we examined the effect of sexual transmission on genetic adaptation in nef and on the activities of Nef proteins. We hypothesized that CTL activity would drive diversifying selection but that most polymorphisms maintained in vivo would minimally affect Nef function. We also hypothesized that in the acutely infected recipients, Nef would be optimized to down-regulate CD4, a phenotype that correlates with enhancement of viral replication in primary CD4-positive T lymphocytes (18, 28), whereas down-regulation of MHC-I would be optimized in chronically infected source hosts whose CTL responses were more robust. Strikingly, we observed broad CTL responses not only in the sources but also in the acutely infected recipients. Under these conditions, CTL-mediated selection pressure appeared to be a major evolutionary force for molecular diversification and adaptation of nef in all individuals. Despite such changes, the Nef proteins from both sources and recipients maintained remarkably robust functional activities for the down-regulation of both MHC-I and CD4. Neither activity was optimized by the transmission event, and in no case could gene-wide evolution be attributed to a single Nef function. These data indicate that Nef is sufficiently adaptable to maintain two independent functions despite the changing immune pressure associated with sexual transmission.

**TABLE 1. Demographics of the six transmission pairs**

<table>
<thead>
<tr>
<th>Pair no.</th>
<th>Subject&lt;sup&gt;a&lt;/sup&gt;</th>
<th>AEH class (length of infection [days])&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Age (yr)</th>
<th>CD4 count&lt;sup&gt;c&lt;/sup&gt;</th>
<th>BP vRNA (log)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Estimated days between sample/transmission event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S1</td>
<td>E1 (85)</td>
<td>49</td>
<td>476</td>
<td>6.22</td>
<td>−26</td>
</tr>
<tr>
<td>1</td>
<td>R1</td>
<td>A3 (45)</td>
<td>37</td>
<td>383</td>
<td>4.88</td>
<td>+45</td>
</tr>
<tr>
<td>2</td>
<td>S2</td>
<td>E2 (166)</td>
<td>41</td>
<td>599</td>
<td>5.04</td>
<td>+101</td>
</tr>
<tr>
<td>2</td>
<td>R2</td>
<td>A3 (46)</td>
<td>38</td>
<td>562</td>
<td>4.85</td>
<td>+45</td>
</tr>
<tr>
<td>3</td>
<td>S3</td>
<td>C (NA)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>36</td>
<td>324</td>
<td>4.43</td>
<td>−627</td>
</tr>
<tr>
<td>3</td>
<td>R3</td>
<td>A3 (44)</td>
<td>44</td>
<td>634</td>
<td>5.12</td>
<td>+45</td>
</tr>
<tr>
<td>4</td>
<td>S4</td>
<td>C (NA)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>36</td>
<td>469</td>
<td>4.68</td>
<td>+45</td>
</tr>
<tr>
<td>4</td>
<td>R4</td>
<td>A1 (21)</td>
<td>35</td>
<td>466</td>
<td>6.42</td>
<td>+21</td>
</tr>
<tr>
<td>5/6</td>
<td>S5/6</td>
<td>A1 (21)</td>
<td>36</td>
<td>536</td>
<td>5.75</td>
<td>−73 (to R5); −49 (to R6)</td>
</tr>
<tr>
<td>5/6</td>
<td>R5</td>
<td>A3 (44)</td>
<td>48</td>
<td>930</td>
<td>7.18</td>
<td>+45</td>
</tr>
<tr>
<td>5/6</td>
<td>R6</td>
<td>A1 (21)</td>
<td>44</td>
<td>647</td>
<td>6.70</td>
<td>+21</td>
</tr>
</tbody>
</table>

<sup>a</sup> Individuals are men who reported sex with men as their HIV risk factor and are enrolled in the San Diego AIEDRP cohort.

<sup>b</sup> AEH, acute and early HIV infection. Duration of infection was estimated using a previously described algorithm (43).

<sup>c</sup> CD4<sup>+</sup> cells per μl of blood.

<sup>d</sup> BP vRNA, blood plasma viral RNA (copies per ml of blood).

<sup>e</sup> NA, not available.

**MATERIALS AND METHODS**

**Study participants, specimen collection, and processing.** Eleven subjects were identified through the NIH-funded Acute Infection and Early Disease Research Program (AIEDRP), representing six transmission events (pair 6 included a different recipient matched to the same donor as pair 5). Three of the six pairs were identified by epidemiological linkage. The other three pairs were identified through screening of the AIEDRP cohort by phylogenetic analysis of baseline population-based pol sequences. Pairs that were linked through phylogenetic analysis demonstrated pol sequence homology such that transmission between the individuals was highly likely (<1% genetic divergence), as described previously (15) and below. Additionally, population based nef sequencing was performed from HIV RNA extracted from blood plasma from all suspected source and recipient partners to confirm transmissions by phylogenetic analysis (15). Study participants were classified as in either the acute, early, or chronic stage of HIV infection (reference 43; see also Table S1 in the supplemental material).

Samples were obtained from source partners from 26 days before to 45 days after the presumed date of transmission, with the exception of the source in pair 3, from whom blood was only available 627 days after transmission. Recipients were sampled 21 to 45 (mean, 37) days after transmission. The time of infection was estimated using established AIEDRP algorithms (26).

Blood from documented source partners and the recipients was collected and centrifuged to separate peripheral blood mononuclear cells (PBMCs) and blood plasma. Plasma samples were stored at −80°C until analysis.

**HIV-1 plasma RNA extraction and sequencing.** HIV RNA was extracted from blood plasma using the QiaAMP viral RNA mini kit (QIAGEN, Valencia, CA). RNA was amplified via reverse transcription-PCR using the Finnzyme system (Espoo, Finland) with primers situated outside of nef in the viral genome (see Table S2 in the supplemental material [OutNeF5′ and OutNeF3′]) followed by nested PCR with internal primers (see Table S2 [nef_INS5′ and nef_INS3′]) using Hi-Fidelity Platinum Taq (Invitrogen, Carlsbad, CA). To avoid contamination, source and recipient RNA was extracted and amplified on separate days and in separate locations. After purification with the QIAGEN PCR cleanup kit (QIAGEN), products were cloned into the pDNA 3.1 TOPO-V5/His vector (Invitrogen) following the manufacturer's instructions. Individual clones were sequenced using the T7 Forward and BGH Reverse primers provided with the TOPO-V5/His kit (Invitrogen) and an ABI 3100 genetic analyzer. A total of 138 clones were sequenced, 8 to 28 clones for each sample of plasma RNA. Abbreviations for sequences obtained in this study use the nomenclature Sx or Rx, where S is for the source partner, R is for the recipient partner, and x is the pair number, as delineated in Table 1. For the phyotypic analyses, the sequences are referred to as Sx-y or Rx-y, where y refers to the clone number. Sequences used in this study will be made available on GenBank.
Sequence analysis. Translated sequences were aligned against HXB2 nef (K03455), mapped back to corresponding codons, and manually adjusted using BioEdit. Aligned sequences were subsequently subjected to an array of maximum likelihood-based comparative analyses (described below) in an attempt to quantify the evolutionary process governing viral evolution during transmission and adaptation to the immune response. Unless otherwise stated, all analyses were performed using the HyPhy (www.hyphy.org) (37) software package. HyPhy scripts implementing the analyses are either a part of the package distribution or can be obtained from the authors upon request.

Clonal sequences from each transmission pair were screened for evidence of recombination. Phylogenetic trees were built using a heuristic maximum likelihood search procedure under the REV (44) substitution model with site-to-site rate variation corrected for with the beta-gamma distribution (24), performing randomized sequential addition with nearest neighbor interchange branch swapping after every 10 sequences were added.

To confirm that transmission events were genetically linked, six phylogenetic trees were reconstructed: one separately for each transmission group and a master tree for all transmission pairs. To assess diversity within and between transmission pairs, we estimated the mean pairwise sequence divergence within the source and the recipient of each transmission pair based on the maximum likelihood phylogeny and branch lengths derived by fitting the MG94kREV model of codon evolution (22) with the appropriate distribution of site-to-site substitution rates. Additionally, we computed approximate 95% confidence intervals on the mean divergence using profile likelihood.

Analysis of selection. For each transmission pair, selection within individuals was quantified using a random effects likelihood model, which assumes that all branches within the individual’s nef sequence tree share the same ratio of synonymous to nonsynonymous substitution rates (dn/dS) (35). For the transmission branches between individual in a pair, selection was quantified using a newly developed genetic algorithm to account for possible heterogeneity of site-to-site variation in both synonymous and nonsynonymous substitution rates with the rest of the sequence. First, we fitted a codon model which allowed site-to-site variation in both synonymous and nonsynonymous substitution rates (34). Second, we compared whether the estimates of p and dn/dS were significantly different between the regions using the likelihood ratio test with appropriately constrained models.

Selection on discordant CTL epitopes. The region-wide test for differential strength of selection (36) was extended to investigate whether adaptive evolution, measured both by the strength of selection (dn/dS) and the proportion of codons undergoing adaptive change (p, with dn > ds), was different within CTL epitope present only in the source or the recipient (but not in both) compared with the rest of the sequence. First, we fitted a codon model which allowed site-to-site variation in both synonymous and nonsynonymous substitution rates (dn/dS) (35). For the transmission branches between individual in a pair, selection was quantified using a newly developed genetic algorithm to account for possible heterogeneity of site-to-site variation in both synonymous and nonsynonymous substitution rates with the rest of the sequence.

Conservation of functional regions. Cumulative interpatient amino acid divergence was estimated using only the internal branches in the phylogenetic tree that encompassed between-patient evolution, after partitioning the nef sequence into several nonoverlapping regions of interest and considering the rest of the sequence as background. To measure divergence as expected amino acid substitutions per site per unit time, we fitted the Jones et al. (20) model of amino acid evolution to each sequence region individually using pooled source and recipient sequences. This model was chosen among 12 popular empirical evolutionary models using small sample Akaike’s information criterion scores. Approximate 95% confidence intervals for cumulative interpatient divergence were determined using profile likelihood for each region of interest and the background.

Lastly, we conducted a likelihood ratio test to determine whether interpatient divergence was significantly different between the regions.

HLA genotyping and mapping of Nef-specific CTL. HLA genotyping was performed as described previously using PCR-based sequencing (29). The CTL responses were detected as described previously (27). In brief, the presence of specific CTL was detected using expanded PBMCs and a gamma interferon (IFN-γ) enzyme-linked immunosorbent assay (ELISPOT). CD8+ cells were screened for reactivity as measured by secretion of IFN-γ using a library of 15-mer peptides representing the entire clade B consensus sequence of Nef (NIH AIDS Research and Reference Reagent Program). Spot-forming cells (SFC) were counted using an automated ELISPOT reader (AID). Counts were all normalized to SFC/10^6 cells. A response was considered significant if it was both greater than two times the average SFC/10^6 cells of the negative control wells and greater than 100 SFC/10^6 cells over the background of the negative control wells.

Phylogenetic construction for functional analysis. Selected clones used for functional analyses were amplified by PCR as follows. For pairs 1, 3, 4, and 5, the sense primer was used was EcoNef5′ and the antisense primers were pci_S1RI_Sal3′ for pairs 1 and 3, CF194_SalNeF3′ for pair 4, and SSR5_SalI3 for pair 5′. For pair 2, sense primer pci_S2R2_ecoRI_Sal5′ and antisense primer pci_S2R2_SalI3′ were used. Primer sequences are provided in Table S2 of the supplemental material. Following PCR with the Hi-Fidelity Platinum Tag system (Invitrogen), the products were inserted into the pCI-neo vector (Promega) by digestion with the EcoRI and SalI sites (see Table S2, in bold). All plasmids were confirmed by sequencing to exclude mutations induced by the PCR cloning process.

Transfection and flow cytometry. SupT1 cells (3 × 10^5) were transfected during exponential growth with 20 μg of the pCI-neo vector (empty or containing Nef clones) and 2 μg of the pGCG-FEP vector (a gift from Jacek Skowronska) as a transfection marker using the Amaxa cell kit V, protocol O-17 (Amaxa Systems, Gaithersburg, MD). Cells were incubated for 24 h after transfection and then stained with anti-CD4-allophycocyanin (Becton Dickinson) and anti-HLA-A2-phycocerythrin (a generous gift from David Camerini, University of California, Irvine). Cells were then fixed and analyzed on a Coulter Elite flow cytometer. The average phycocerythrin and allophycocyanin fluorescence intensities of the green fluorescent protein (GFP)-positive cells were plotted. Assays were performed in duplicate and are representative of two to six independent experiments for each transmission pair.

Western blotting. Samples for analysis by Western blotting were taken from the same transfected cells that were analyzed by flow cytometry. Cells were suspended in loading buffer containing sodium dodecyl sulfate and boiled for 10 min. After resolution on a 12% denaturing polyacrylamide gel (Bio-Rad), the proteins were transferred to a nitrocellulose membrane and blotted with the following antibodies: mouse antitubulin (1:6,000; Sigma); sheep anti-Nef (1:1,500; a gift from Celsa Spina, University of California San Diego). Detection was performed using a goat anti-mouse antibody linked to horseradish peroxidase (Bio-Rad, Hercules, CA) and a rabbit anti-sheep antibody linked to peroxidase (DAKO, Glostrup, Denmark), followed by development with enhanced chemiluminescence (Amersham-Pharmacia, Piscataway, NJ).

RESULTS

Characteristics of transmission cohort. To examine the evolution of nef after sexual transmission, we identified six transmission pairs from the San Diego site of the AIEDRP based on epidemiologic evidence and phylogenetic analysis of the pol gene (15) (Table 1). These data suggested that two recipients were infected from the same source (Fig. 1, pair 5/6). The average age of the subjects was 40.6 years (range, 36 to 49), and the average CD4+ T-lymphocytes/μL of blood was 547.8 (range, 383 to 930) (Table 1). A single sample of plasma RNA was analyzed from each individual; the concentrations ranged from 10^4.4 to 10^7.2 copies per ml (Table 1). The estimated intervals between the transmission events and the dates of sample collection are indicated in Table 1. Notably, two recipients, R4 and R6, were sampled within 21 days of the estimated date of transmission. At the time of sampling, none of these individuals had received antiretroviral therapy.

Phylogenetic linkage of transmission pairs by sequence analysis of nef. We analyzed nucleotide sequences of 138 nef clones from the six source and recipient pairs, with an average of 12 clones/individual, by cloning reverse transcription-PCR products derived from plasma RNA. Of all 138 clones, only 2 (both found in recipients) encoded Nef proteins predicted to be defective based on the presence of premature termination codons. The predicted amino acid sequences of the 136 intact open reading frames are shown aligned in Fig. S1A to E in the supplemental material.

Phylogenetic analysis (Fig. 1) indicated that the sequences from each individual and transmission pair formed distinct monophyletic groups, with strong bootstrap support for separating branches, with the exception of recipients R5 and R6, who grouped together and were infected by the same source (S5/6). This tree morphology was consistent with the putative transmission events. Recombination, which can severely bias phylogeny-based comparative methods (41), was not detected.
In any of the six transmission pairs, based on a robust maximum likelihood GARD test (23) (data not shown). Importantly, in each of the six transmission events only a single branch connected the source with the recipient sequences (Fig. 1). These data suggest that a monoclonal population of nef sequences was transmitted. In one pair (pair 4), a single source clone grouped with the recipient clones, suggesting the possibility of gene-wide purifying selection during the transmission event.

Genetic diversity of nef in sources and recipients. Previous data suggested that the genetic diversity of nef is maintained across vertical transmission (40). To determine whether this finding was true of sexual transmission, the diversity of nef sequences within each host (defined as the average length of the path connecting two sequences from a given individual in the phylogenetic tree) was assessed for each transmission pair. As shown in Fig. 1, the diversity of nef sequences within each individual was relatively low compared to the diversity between transmission pairs (0.16 to 1.80% versus 16.22%; 95% confidence interval [CI], 14.66 to 17.87%) (Fig. 1). With the exception of pair 3, in which diversity was lower in the recipient, the diversities of the nef sequences were statistically indistinguishable between the sources and the recipients of each pair.

Diversifying and purifying selective pressures on nef. To examine the selective pressures in the transmission pairs and within individuals, the dN/dS ratio was estimated from the entire nef gene (Fig. 1) (14). Distinct patterns of selection were detected in most pairs. Purifying selection was detected across the transmission branch of pair 1 (dN/dS, 0.26; 95% CI, 0.1 to 0.32), but within either individual evolution was essentially neutral. In contrast, in pair 2, diversifying selection was found both within the source (dN/dS, 2.6; 95% CI, 0.9 to 5.6) and the recipient (dN/dS, 5.1; 95% CI, 2.6 to 9.0), while selection across the transmission branch was essentially neutral. Pair 3 exhibited trends toward purifying selection within the source and the recipient but diversifying selection across transmission. However, this branch incorporates 18 months of sequence evolution in the source following transmission, rendering interpretation problematic. In pair 4, sequences in the recipient were under diversifying selection (dN/dS, 3; CI, 1.2 to 6.0), while

FIG. 1. Maximum likelihood phylogeny of nef sequences from six putative transmission pairs. A maximum likelihood phylogenetic tree was constructed with all nef sequences obtained from each individual in the study. Duplicate sequences were not included. All transmission events grouped together with 100% bootstrap support. Diversity of all clones within individuals is indicated. Transmission branch length represents the genetic distance between the most recent common ancestors of the clonal populations in two individuals of a transmission pair.
Functional regions within Nef evolve at similar rates in sources and recipients. We hypothesized that in acutely infected recipients, Nef would be optimized to enhance viral infectivity and replication, while in chronically infected sources, the down-regulation of MHC-I would be optimized. To test this hypothesis, we focused the analysis of protein diversification in specific functional regions were compared to the background rate for the rest of the protein. Three regions of Nef were assessed: (i) the N-terminal α-helix, spanning residues 61 to 80; and (ii) the acidic cluster/polyproline region, evolving at a rate similar to the background rate for the rest of the protein (excluding the other functional regions). Purifying selection across the transmission events was not consistently found, weighing against a scenario in which the release of CTL pressure in the nonimmune recipient results in convergent evolution towards a more fit nef sequence.

Importantly, key residues in these regions, such as the central prolines in the polyproline helix and the two leucine residues in the C-terminal loop, were universally conserved. Overall, the rates of evolution in regions of Nef related to specific functions did not appear to change measurably with transmission (Table 2), suggesting that differential selection of a specific function across transmission or during acute infection may not occur.

## CTL responses to Nef peptides were divergent between individuals and may drive diversification after transmission.

Many of the amino acid differences between the source and recipient clones, and between these clones and the HXB2 reference sequence (see Fig. S1 in the supplemental material), were previously described as escape adaptations to CTL activity (2, 5, 40). To test the hypothesis that these polymorphisms may be attributed to the selective pressure of CTL, both the HLA genotypes of the individuals and the breadth of their CTL activity were determined (Table 3; see also Fig. S1 in the supplemental material). For the detection of CTL activity, peripheral blood mononuclear cells obtained at the same time as the plasma RNA samples were expanded and responses were measured using an array of overlapping peptides that covered the entire Nef protein in an ELISPOT assay for IFN-γ (Table 3).

### CTL responses to clade B consensus Nef peptides were detected in 8 of 11 individuals. Two of the individuals were unable to be tested due to a lack of viable cells, and in individual R2, no responses were detected despite viable cell cultures. Notably, two individuals in the earliest stage of infection exhibited broad CTL responses (R4 and S5/6).

To assess the possibility that CTL activity was a driving force in the diversification of nef sequences across transmission and during acute infection, we took advantage of the observation that pair 3 exhibited fairly divergent and broad CTL responses (Table 3; see also Fig. S1 in the supplemental material). This allowed the comparison of selection within regions that were only targeted by the CTL of one individual (discordant epitopes) with selection elsewhere in the Nef protein. Accordingly, the 15% of the codons that were in discordant epitopes were under strong diversifying selection (dN/dS, 8.4). Of the codons outside discordant epitopes, 54.8% were under weak diversifying selection (dN/dS, 1.5), while the remaining codons were not under detectable selection. This difference was sig-

---

*Numbers represent the relative rate of evolution of specific regions of Nef (setting the rest of the sequence at 1) among interpatient branches in the trees, fitted using the JTT model of protein evolution (95% confidence intervals are in parentheses). P values indicate whether those rates are different within the functional region versus the rest of Nef, not including the other functional regions.

### Table 2. Relative rate of evolution of Nef sequences

<table>
<thead>
<tr>
<th>Region</th>
<th>Sources</th>
<th>Recipients</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-terminal α-helix</td>
<td>3.25 (1.68–5.66); P = 0.002</td>
<td>2.35 (1–4.63); P = 0.05</td>
</tr>
<tr>
<td>Acidic cluster/PxxP</td>
<td>0.66 (0.26–1.28); P = 0.27</td>
<td>0.66 (0.26–1.34); P = 0.86</td>
</tr>
<tr>
<td>C-terminal loop</td>
<td>1.66 (1.15–2.31); P = 0.02</td>
<td>1.64 (1.1–2.4); P = 0.025</td>
</tr>
<tr>
<td>Rest of Nef protein</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

* Numbers represent the relative rate of evolution of specific regions of Nef (setting the rest of the sequence at 1) along interpatient branches in the trees, fitted using the JTT model of protein evolution (95% confidence intervals are in parentheses). P values indicate whether those rates are different within the functional region versus the rest of Nef, not including the other functional regions.

### Table 3. HLA genotypes and CTL responses of transmission pairs

<table>
<thead>
<tr>
<th>Subject</th>
<th>Genotype</th>
<th>Region(s) in Nef triggering CTL response</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>HLA-A</td>
<td>01_03 40_08 03_07</td>
</tr>
<tr>
<td>R1</td>
<td>HLA-B</td>
<td>01_02 15_08 06_07</td>
</tr>
<tr>
<td>S2</td>
<td>HLA-C</td>
<td>30_11 70_14 07_07</td>
</tr>
<tr>
<td>R2</td>
<td></td>
<td>02_11 15_44 30_05</td>
</tr>
<tr>
<td>S3</td>
<td></td>
<td>02_24 39_48 08_12</td>
</tr>
<tr>
<td>R3</td>
<td></td>
<td>30_33 44_70 07_07</td>
</tr>
<tr>
<td>S4</td>
<td></td>
<td>01_02 53_57 04_06</td>
</tr>
<tr>
<td>R4</td>
<td></td>
<td>01_11 09_38 07_12</td>
</tr>
<tr>
<td>S5/6</td>
<td></td>
<td>32_68 40_44 03_07</td>
</tr>
<tr>
<td>R5</td>
<td></td>
<td>25_24 15_44 03_05</td>
</tr>
<tr>
<td>R6</td>
<td></td>
<td>11_29 44_44 16_16</td>
</tr>
</tbody>
</table>

* Numbering refers to the HIV-1 clade B Nef consensus sequence.

---

The interpatient branches of the tree were assessed between all sources and recipients.

We hypothesized that in acutely infected recipients, Nef would be optimized to enhance viral infectivity and replication, while in chronically infected sources, the down-regulation of MHC-I would be optimized. To test this hypothesis, we focused the analysis of protein diversification in specific functional regions. Notably, two individuals in the earliest stage of infection exhibited broad CTL responses (R4 and S5/6).
significant (P < 0.01), indicating that changes in CTL recognition between hosts can be a major driving force for the diversification of nef sequences during transmission and subsequent acute infection.

Functional effects of Nef variations in transmission pairs.
To examine the impact of observed amino acid polymorphisms on function, clones from all individuals were assayed for functions related to immune evasion (down-regulation of MHC-I) as well as down-regulation of CD4, a surrogate marker for enhancement of replication rate (28). The clones selected for analysis met at least one of the following criteria: (i) the clone encoded the amino acid sequence closest to the consensus sequence for that individual; (ii) the clone contained one or more mutations found in at least two other clones; or (iii) the clone contained mutations in or near a known functional region of Nef. An alignment of the tested clones is shown in Fig. 2.

Overall, 21 Nef-encoding clones from the six transmission pairs were assessed for the down-regulation of MHC-I and of CD4 using transiently transfected SupT1 T cells and a flow cytometric assay (Fig. 3). The expression of the various Nef proteins in the transfected cells was tested by Western blotting using a polyclonal antiserum. Of the 21 clones, four (S1-3, S1-4, S3-2, and S3-4) exhibited functional defects. In three of these cases, the defects could be attributed to poor expression (Fig. 3A and C). None of these four functionally impaired source clones represented the majority variant in that host.

In general, both the down-regulation of MHC-I and of CD4 were well maintained in both sources and recipients. No overall enhancement or diminution of either Nef-function was detected between source and recipient sequences. As noted above, certain minority clones encoded defective or partly defective proteins. Clone S1-3 was selectively impaired in the down-regulation of CD4 and contained a unique mutation, V166G (Fig. 2), located near the N-terminal end of the C-terminal flexible loop. Although apparently poorly expressed, this clone maintained robust down-regulation of MHC-I (Fig. 3A). Clone S1-4 was completely nonfunctional for CD4 or MHC-I down-regulation, and the expression of Nef was not detected by Western blotting, possibly due to V159A and/or W196R mutations. The W196R mutation was not found in other clones of the cohort, while the V159A mutation was found in single clones from R4 and S5/6 not analyzed in Fig. 3 (see Fig. S1D and E in the supplemental material). Minority clones derived from S3 also exhibited defects in Nef functions (S3-2 and S3-4) (Fig. 3C). The S3-2 clone was minimally active in the down-regulation of MHC-I, and it had an intermediate ability to down-regulate CD4. The S3-4 clone was virtually inactive for either function. These defects were not clearly attributable to specific, unique polymorphisms. However, both S3-2 and S3-4 encoded a substitution of serine 8 with leucine, and S3-4 encoded a substitution of serine 116 with proline; these mutations were not found in other clones of the cohort. The S3-3 and R3-1 clones were phenotypically wild type, despite relatively low levels of protein detected by Western blotting, suggesting antigenic differences that hindered the detection of these Nef proteins.

The remaining 17 of 21 clones analyzed were fully functional in both the down-regulation of MHC-I and of CD4. These data suggest that both of these Nef functions are important during HIV transmission and acute infection. The data also indicate that Nef is able to tolerate extensive and diverse polymorphisms, some of which are likely responses to the selective pressure of CTL, without compromising its immune evasion and virologic functions.

DISCUSSION
In this study, we explored the genetic and phenotypic changes in the HIV-1 nef gene in epidemiologically and phylogenetically linked transmission pairs. The transmission cohort represented a range of stages of HIV-1 infection. In principle, this allowed a test of the hypotheses that specific Nef sequences and functions were associated with transmission or were optimized during acute infection. Sequence and phylogenetic analyses indicated that although the diversity of the nef sequences in acutely infected recipients was usually similar to that in their matched sources, tree morphology suggested that a mono- or oligoclonal population of nef sequences was transmitted. The patterns of gene-wide selection across transmission were variable, and regions of Nef associated with distinct functions evolved similarly in sources and recipients. These results weighed against the selection of specific Nef functions by transmission or during acute infection. This conclusion was supported by direct measurement: the relative activities of down-regulation of CD4 and MHC-I, properties with distinct genetic requirements, were virtually unchanged by the transmission of HIV-1. Furthermore, in the majority of clones in both sources and recipients, these two functions were remarkably well preserved despite extensive sequence diversity.

In this cohort, the nef coding region varied by as much as 18% between transmission pairs but exhibited substantially lower diversity within individuals and within pairs. The diversity within the samples obtained from the individuals of each pair did not differ significantly, with one exception (pair 3), in which diversity decreased after transmission. The lack of a consistent reduction in the diversity between the individuals of each pair might suggest the absence of a genetic bottleneck during sexual transmission with respect to nef. As noted previously in the case of vertical transmission (40), this observation may reflect either the transmission of multiple viral strains or the rapid diversification of a mono- or oligoclonal population in the new host. In support of the latter, the diversity of nef sequences correlated to some extent with the duration of infection in the recipient (Fig. 1 and Table 1). For example, the recipients in pairs 1, 2, and 3 were in the later stages of acute infection (A3), and their Nef sequences were characterized by relatively high diversities compared to the recipients of pairs 4 and 5/6, who were in the earliest stages (A1). Pairs 1, 2, and 3 also have longer transmission branches than pairs 4 and 5/6, consistent with the hypothesis that diversity increases with time during acute infection.

The different patterns of selection (purifying versus diversifying) observed among the transmission pairs may reflect various pairings of concordant and discordant immune responses. The variability in selective pressures after transmission observed here in the case of nef is reminiscent of a characterization of evolution within env as driven by neutralizing antibodies (14). Under these variable conditions, the CTL response could drive either purifying or diversifying selection. For example,
transmission could drive purifying selection if CTL escape mutations generated in the source confer a selective advantage in the recipient due to recognition of the same epitopes. A possible example of this is found in pair 1; the individuals of this pair share three HLA alleles, and purifying selection was inferred across the transmission branch. A more common scenario may be one in which diversifying, positive selection results from new immune responses in the recipient relative to those in the source. While pairs 5 and 6 may represent such cases (diversifying selection occurred across transmission) the CTL response was unable to be evaluated in the recipients due to the unavailability of viable cells. Notably, positive selection was detected in sequences obtained from individuals S3 and R4, both of whom had broad CTL responses, a result consistent with the hypothesis that CTL pressure drives diversifying selection.

FIG. 2. Protein sequences of clones used in phenotypic analyses. Clones are listed by pair; S refers to sources, and R refers to recipients. The number after S or R indicates the number of the pair, as shown in Table 1, and the number after the dash designates a specific clone from that individual. The alternate shading separates the transmission groups. All amino acid sequences were aligned to HXB2 Nef as a reference. Dots indicate identity with the HXB2 sequence, and dashes indicate gaps. Bold boxes indicate positions at which differences were detected between the source and recipient of each pair. Annotations denote the significance of the clone chosen, as follows: 1, the clone was most representative of the consensus sequence for that individual; 2, the clone contained amino acid changes not representative of the population in the individual but was found in at least two other clones; 3, the clone encoded mutations in an amino acid sequence near a known functional region of Nef.
In addition to the selective pressures induced by host-to-host variation in CTL activity, the process of transmission might select for a specific Nef-function, for example, the enhancement of viral infectivity or replication. Genetic evidence of such selection was sought by examining the relative rates of amino acid evolution in regions of Nef associated with specific functions: either the down-regulation of MHC-I (the N-terminal α-helix and the acidic cluster/polyproline region) or the down-regulation of CD4/enhancement of infectivity and replication (the C-terminal flexible loop). The estimated rates of evolution of these regions were essentially the same in sources and recipients, providing genetic evidence against the association of a specific Nef-function with the efficiency of transmission or with viral replication in the acutely infected host. Consistent with previous data, the increased rate of evolution within the N-terminal α-helical region and the C-terminal flexible loop suggested that these areas are relatively more tolerant of polymorphisms than the acidic cluster/polyproline region (21, 42).

The diversity of Nef sequences and specific polymorphisms within this cohort were extensive, yet the majority of the Nef proteins tested were fully functional and previously defined key residues were conserved. For example, an N-terminal duplication (Q33-A38, pair 2; analogous to previously noted duplications found in 36% of available nef sequences) did not affect function (17). A key position in this region associated with the down-regulation of MHC-I (M20) was conserved or contained a conservative change (I20). The acidic cluster region associated with the down-regulation of MHC-I (E92-65 in HXB2; positions 73 to 78 in Fig. 2) exhibited insertions and changes that did not affect function; this observation is consistent with the reported insertion of a glycine or glutamic acid in the acidic cluster motif in ∼5% of nef sequences (17). An additional acidic residue in this motif did not confer improved activity in MHC-I down-regulation (compare S2-1 and S2-2 to S2-3).

Notably, all the prolines of the SH3-binding region were conserved throughout this cohort. The variability of positions within the C-terminal flexible loop was extensive. The di-acidic sequence associated with binding to β-COP (4, 32) had a high proportion of lysine substitutions at the second glutamic acid (position 168 in Fig. 2; also seen in references 21 and 42); the majority of these clones were fully functional (all tested clones of pairs 2, 4, and 5/6, as well as clone S3-3). Nonconservative substitutions within the acidic di-leucine motif (E169xxxLL175 in HXB2; positions 173 to 178 in Fig. 2) demonstrated unexpected flexibility at the +2 and +3 positions relative to the acidic residue. These polymorphisms included the replacement of polar, uncharged residues with hydrophobic and basic residues, substitutions that were surprisingly functional in view of the sequence preferences for the binding of such motifs to the adaptor protein complexes involved in endosomal trafficking (8). Nevertheless, the ExxLL adaptor protein-binding motif within the C-terminal flexible loop was universally conserved in this cohort. Finally, the F191 residue recently associated with binding of Nef to Pak2 (position 204 in Fig. 2) was also universally conserved (31).

As noted above, we initially hypothesized that the ability of Nef to down-regulate CD4 would be optimized in acutely infected patients at the expense of its ability to down-regulate MHC-I. This hypothesis was based on the assumption that subjects identified very early during acute infection would essentially be nonimmune. It also followed from the reported relationship between these two Nef-activities in late-stage, chronically infected patients whose CTL responses had presumably waned: a relative optimization of the down-regulation of CD4 at the expense of down-regulation of MHC-I (5). This hypothesis was not supported; both functions were robust in the majority of clones, and no change was detected in either activity across the transmission event. Why would the modulation of MHC-I be preserved in these acutely infected individuals? We suspect that at the time of sampling, both sources and recipients had developed CTL responses, which would provide a selective advantage to Nef proteins able to down-regulate MHC-I. Indeed, broad CTL activity was detected as early as 21 days after transmission (individual R4), suggesting that the temporal window during which Nef can evolve in the nonimmune, acutely infected host is very short.

Because we have not measured all the activities of Nef, it remains possible that a property of Nef so far untested is optimized during transmission or acute infection. Such properties include the ability of Nef to enhance viral infectivity or replication by a mechanism independent of the down-regulation of CD4 or the ability of Nef to facilitate T-cell activation. In this regard, we are aware of no residue that has been described as important for either of these Nef phenotypes that does not also contribute to either the down-regulation of CD4 or MHC-I. Taken together, the two genetically distinct Nef functions measured here may serve as an effective general screen for defects in Nef activity.

In conclusion, two distinct Nef functions, the down-regulation of CD4 and of MHC-I, were well conserved both before and after sexual transmission of HIV-1. Genetic and phenotypic data suggested that Nef tolerates multiple mutations, probably driven by CTL pressure, without a fitness cost. This conclusion is consistent with a recent report indicating that CTL escape variants of internal viral proteins are transmitted sexually in proportion to their frequency in the source; no diminished capacity for transmission was detected (12). The data also suggest that neither sexual transmission nor the subsequent acute infection subjects Nef to functional constraints distinct from those present during chronic infection. We speculate that the down-regulation of CD4 and MHC-I by Nef are each crucial to the successful establishment and maintenance of a new infection. Alternatively, these two activities may together reflect an as-yet-unrecognized overarching function of Nef that is preserved throughout various stages of disease.

ACKNOWLEDGMENTS

This work was supported by grants K23AI055276, AI38201, AI57167, AI27670, AI38858, UCSD Center for AIDS Research (AI36214, AI29164, AI47745, and AI051970 from the National Institutes of Health; the Flow Cytometry Research Core Facility of the VA San Diego Healthcare System; and the Research Center for AIDS and HIV Infection of the San Diego Veterans Healthcare System. C.M.N. is the recipient of a Universitywide AIDS Research Program Dissertation Fellowship, D05-VMRF-404#.

We thank Christopher Woelk for helpful comments and advice, David Camerini for the A2-specific antibody, and Nanette Van Damme, Jorge Jimenez, Theresa Russell, Parris Jordan, Caroline Ignacio, Judy Nordberg, and Nancy Keating for technical support.


