JOURNAL OF VIROLOGY, June 1995, p. 3778–3788 Vol. 69, No. 6
0022-538X/95/$04.00+0
Copyright © 1995, American Society for Microbiology

Analysis of Envelope Sequence Variants Suggests Multiple Mechanisms of Mother-to-Child Transmission of Human Immunodeficiency Virus Type 1

LAURENCE BRIANT,[1]† CHRISTOPHER M. WADE,[2] JACQUELINE PUEL,[1]
ANDREW J. LEIGH BROWN,[2] AND MIREILLE GUYADER[1]*

Laboratoire de Virologie, Centre Hospitalo-Universitaire Purpan, 31059 Toulouse, France,[1] and Centre for HIV Research, Institute of Cell, Animal and Population Biology, Division of Biological Sciences, The University of Edinburgh, Edinburgh EH9 3JN, United Kingdom[2]

Received 5 July 1994/Accepted 6 March 1995

In order to elucidate the molecular mechanisms involved in human immunodeficiency virus type 1 (HIV-1) mother-to-child transmission, we have analyzed the genetic variation within the V3 hypervariable domain and flanking regions of the HIV-1 envelope gene in four mother-child transmission pairs. Phylogenetic analysis and amino acid sequence comparison were performed on cell-associated viral sequences derived from maternal samples collected at different time points during pregnancy, after delivery, and from child samples collected from the time of birth until the child was approximately 1 year of age. Heterogeneous sequence populations were observed to be present in all maternal samples collected during pregnancy and postdelivery. In three newborns, viral sequence populations obtained within 2 weeks after birth revealed a high level of V3 sequence variability. In contrast, V3 sequences obtained from the fourth child (diagnosed at the age of 1 month) displayed a more restricted heterogeneity. The phylogenetic analysis performed for each mother-child sequence set suggested that several mechanisms may potentially be involved in HIV-1 vertical transmission. For one pair, child sequences were homogeneous and clustered in a single branch within the phylogenetic tree, consistent with selective transmission of a single maternal variant. For the other three pairs, the child sequences were more heterogeneous and clustered in several separate branches within the tree. In these cases, it appeared likely that more than one maternal variant was responsible for infection of the child. In conclusion, no single mechanism can account for mother-to-child HIV-1 transmission; both the selective transmission of a single maternal variant and multiple transmission events may occur.

Extensive genetic diversity is a characteristic feature of human immunodeficiency virus type 1 (HIV-1) (2, 19, 33, 39). Considerable genomic diversity has been observed among independent isolates from epidemiologically unlinked infections and has also been reported to occur to a lesser degree in vivo, among viral species from a single patient (4, 5, 22, 29, 44, 45). Studies of sequence variation within the V3 hypervariable region of the envelope gene have described the heterogeneous viral population in long-term-infected individuals and have shown that sequences collected from an individual shortly after infection appear to be less variable than those found in the donor at the time of transmission (45, 51). This observation supports the hypothesis that sexual or parenteral HIV-1 infection may be initiated by a limited number of molecular variants. Recent reports based on V3 sequence analyses of mother-child transmission pairs have also suggested that a very limited number of variants or even one particular variant could initiate infection within the child (31, 42, 50). This genotype may represent a minor maternal form, perhaps escaping a critical immune surveillance mechanism or having particular phenotypic properties. However, the molecular and biological properties of viruses transmitted perinatally to children have not yet been determined; whether cell-free or cell-associated virus is transmitted and at what time transmission occurs remain unclear. The detection of HIV nucleic acids in fetal tissues (8, 46) has indicated that transmission to the child may occur at an early stage of gestation. On the other hand, clinical studies have demonstrated that a high proportion of perinatally infected children show no sign of infection at delivery (12, 28). This would suggest transmission to the child either at a late stage of pregnancy or at delivery. In contrast, the early diagnosis of HIV infection in newborns (within the first few days following birth) has been interpreted to reflect infection early in pregnancy. Striking differences have also been reported regarding the evolution of disease in young perinatally infected children. AIDS has been shown to develop more rapidly in 20% of perinatally infected children than in adults (3, 38). Such rapid progression may be associated with infection of the child at an early stage of pregnancy.

In order to elucidate the molecular characteristics of HIV-1 variants involved in mother-to-child transmission, we have assessed the genetic diversity of proviral DNA sequences spanning the V3 loop and flanking regions (313 bp) in four perinatally infected children and their respective mothers. Since transmission may have occurred at any time during pregnancy or at delivery, we have analyzed maternal variants detected at different time points throughout pregnancy and following delivery and compared these with variants detected in the child over a period of 16 months. The analysis of longitudinal samples collected from mothers during pregnancy is essential to determine whether positive selection of maternal variants occurs in HIV-1 vertical transmission. In one mother-child pair we have observed transmission of a single maternal variant, which may indicate a selective process. However, we also re-

* Corresponding author. Present address: Centre d'Immunologie, Parc Scientifique et Technologique de Luminy, Case 906, 13288 Marseille cedex 09, France. Phone: (33) 91 26 94 94. Fax: (33) 91 26 94 30.
† Present address: Centre de Tri des Molécules anti-HIV, CRBM-CNRS, 34060 Montpellier, France.

TABLE 1. Times of sampling and clinical data from mother-child pairs studied

Sample	Time point[a]	CDC stage	$CD4^+$ count ($cells/mm^3$)[b]	Cell viremia ($TCID_{50}$[c]/10^6 cells)	Plasma viremia ($TCID_{50}/10^6$ cells)	Symptoms
Pair A						
Mother						
MA1	3.5 mo	III	618	1,548	0	None
MA2	4.5 mo	III	ND[d]	154	0	None
MA3	6 mo	III	276	25	0	None
MA4	7 mo	III	470	616	0	None
MA5	8 mo	III	ND	123	0	None
MA6	2 mo p-del	III	480	11	0	None
Child						
CA1	1 mo	P2A	2,204	ND	ND	None
CA2	2.5 mo	P2A	ND	ND	ND	None
CA3	16 mo	P2D	1,626	ND	ND	Cryptosporidiosis, herpes
Pair B						
Mother						
MB1	3.5 mo	III	576	55	0	None
MB2	4 mo	III	612	125	0	None
MB3	6.5 mo	III	ND	25	0	None
MB4	Delivery	IVc2	672	56	0	Oral leukoplakia
MB5	3 mo p-del	IVc2	858	126	0	Oral leukoplakia
Child						
CB1	5 days	P2A	ND	ND	ND	None
CB2	1.5 mo	P2A	ND	ND	ND	None
CB3	3.5 mo	P2A	1,912	ND	ND	None
Pair C						
Mother						
MC1	5 mo	II	392	16	0	None
MC2	6.5 mo	II	468	16	0	None
MC3	7.5 mo	II	240	275	3	None
Child						
CC1	1.5 mo	P2A	ND	ND	ND	None
CC2	2.5 mo	P2A	3,159	ND	ND	None
CC3	13 mo	P2A	1,912	ND	ND	None
Pair D						
Mother						
MD1	2 mo	II	528	1,548	0	None
MD2	3.5 mo	II	314	16	0	None
MD3	8.5 mo	IVc2	512	55	0	Zoster
MD4	4.5 mo p-del	IVc2	364	56	0	Zoster
Child						
CD1	1.5 mo	P2A	3,500	ND	ND	None
CD2	11.5 mo	P2A	1,550	ND	ND	None

[a] Time points are expressed as months of pregnancy and age for mothers and children, respectively, unless indicated otherwise. p-del, postdelivery.
[b] The average normal $CD4^+$ cell count for an adult is 1,000 cells per mm^3. For children under the age of 11 months, the normal $CD4^+$ cell count is between 1,700 and 2,880 cells per mm^3 (mean, 2,200) (49). For children between 1 and 6 years of age, the average $CD4^+$ cell count is between 1,000 and 1,800 cells per mm^3.
[c] $TCID_{50}$, 50% tissue culture infective dose.
[d] ND, not determined.

port the apparent infection of children by multiple maternal subtypes, with evidence provided for both early and late transmission events. Vertical transmission of HIV-1 is clearly complex, and further characterization of the transmitted viral species may be required to provide further insight and a better understanding of this transmission route.

MATERIALS AND METHODS

Patients. Sequence variation within the V3 domain and flanking regions in four HIV-1-infected mother-newborn pairs was assessed. All four mothers gave birth at La Grave Hospital in Toulouse, France, and were monitored over a 1-year period from the beginning of pregnancy. The four mother-child pairs were selected because they provided the best series of samples from the mother and child. The clinical status of the patients studied, the time of sampling, levels of cell and plasma viremia, and immunological data are presented in Table 1. Evidence of infection in newborns was provided by coculture and/or PCR using specific primers for the *gag* and *pol* genes (18).

Cell and plasma viremia. Maternal cell viremia and plasma viremia were assessed according to the standard methods of the French Agence Nationale de Recherches sur le SIDA collaborative group. Briefly, fresh peripheral blood mononuclear cells (PBMCs) and plasma were collected from blood samples after separation on Ficoll-Hypaque gradient (Pharmacia). Cells were washed twice in RPMI-1640 (Whittaker). Cells and plasma were then diluted separately from 5^0 to 5×10^6 and cocultivated with 2×10^6 fresh phytohemagglutinin-stimulated PBMCs from an HIV-negative donor in RPMI-1640 medium containing 15% fetal calf serum, 0.3 mg of glutamine per ml, 20 IU of human interleukin 2 per ml, and antibiotics. Four replicates of each sample were analyzed. Twice a week, the supernatant was collected and reverse transcriptase activity was monitored. Cell and plasma viremia titers were calculated according to the Karber method and expressed as 50% tissue culture infective doses per 10^6 cells (21).

Nucleic acid extraction. Total cellular DNA was obtained from patients' uncultured PBMCs. To avoid any risk of contamination, nucleic acid extraction was carried out in laboratories free of PCR products, cultured HIV isolates, or

cloned HIV sequences. DNA purification was performed separately for each of the pairs, and as a control, DNA was also extracted from uninfected PBMCs at the same time. After separation on Ficoll-Hypaque (Pharmacia), cells were washed twice in RPMI-1640 (Whittaker) and lysed for 2 h at 56°C in 10 mM Tris (pH 8.3)–50 mM KCl–2.5 mM $MgCl_2$–0.45% Nonidet P-40–0.45% Tween 20–80 mg of proteinase K per ml. After lysis, proteinase K was inactivated by heating the mixture for 10 min at 95°C.

PCR amplification, cloning, and sequencing of the V3 region. The region encoding the V3 loop and flanking sequences (positions 6615 to 6928 in the HIV-LAI genome [33]) was amplified in a nested PCR (32). One microgram of total cellular DNA was amplified in a 100-μl reaction mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM $MgCl_2$, 0.2 mM each deoxynucleoside triphosphate, 0.4 μg of each primer, and 1.5 U of *Taq* DNA polymerase (Perkin-Elmer Cetus) and overlaid with 25 μl of mineral oil. The first amplification step was performed for 25 cycles with the outer primers E1, 5′-TACAATGTACA CATGGAATT-3′, and E2, 5′-TTACAGTAGAAAAATTCCCC-3′ (positions 6551 to 6570 and 6955 to 6974, respectively, in the HIV-LAI genome [33]). Ten microliters of the first amplification product was then used as the template in a second amplification step performed for 30 cycles with primers E3, 5′-GTATCG GAATTCCTGCTGTTGAATGGC-3′ (positions 6592 to 6618), and E4, 5′-TT AGCAAGCTTCTGGGTCCCCTCCGAGGA-3′ (positions 6907 to 6935), containing *Eco*RI and *Hin*dIII restriction sites, respectively (underlined nucleotides). Primers E3 and E4 have been previously described (35), and all four primers show a high degree of identity with published reference sequences. Thermal cycling was performed by using a Perkin-Elmer Cetus 9600 thermal cycler with denaturing at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 1 min. This was followed by a final extension step for 2 min at 72°C. Ten percent of the secondary PCR product (313 bp) was analyzed by electrophoresis on a 2% agarose gel, with the expected band being visualized following ethidium bromide staining. Amplified products were digested with *Eco*RI and *Hin*dIII restriction enzymes, purified by using the Glassmax purification system (Life Technologies, Bethesda Research Laboratories), and cloned in *Eco*RI-*Hin*dIII-digested M13 vector. The ligated vector was used to transform DH5αF′ competent cells. One to 17 individual positive M13 clones were sequenced for each sample by the dideoxy chain termination method (41) using [α-^{35}S]dATP (T7 sequencing kit; Pharmacia). PCR amplification, cloning, and sequencing were performed separately for each pair in the following order: pair A, pair B, pair D, and pair C. Appropriate negative controls and DNA from uninfected cells were included in each reaction mixture. In addition, all negative controls from the first round of amplification were included in the second amplification step.

Sequence analysis. The nucleotide sequences from each mother-child pair were aligned by using the CLUSTAL V algorithm (20) as implemented in version 2.2 of the Genetic Data Environment package (kindly provided by the Harvard Genome Laboratory). The final alignment was improved manually by preferring gaps to transition differences and transition differences to transversion differences and by the insertion of gaps to maintain the reading frame. Translation of nucleotide sequences to amino acids was also undertaken with this package. Distance-based phylogenetic analyses were performed with programs taken from version 3.52c of the Phylogeny Inference Package (PHYLIP [15]). Nucleotide sequence distances for all pairwise sequence comparisons were estimated by using the generalized two-parameter (maximum-likelihood) model which uses the transition probability formulas of Kishino and Hasegawa (25), incorporating unequal rates of transition and transversion and allowing for different frequencies of the four nucleotides (program DNADIST). Phylogenies were reconstructed by both the neighbor-joining method (40) (program NEIGHBOR) and the Fitch-Margoliash (16) (program FITCH) distance method. Phylogenies were also reconstructed by the maximum likelihood method (13) using the modified PHYLIP program FASTDNAML (kindly provided by Gary Olsen of the University of Illinois at Urbana-Champain and the Ribosomal Database Project) (data not shown). Settings for the transition/transversion ratio were estimated from each data set. Bootstrap resampling (14) was employed on the neighbor-joining trees (programs SEQBOOT and CONSENSE) to assign approximate confidence limits to the branches. Two thousand bootstrap replications were performed. Alternative phylogenetic hypotheses were evaluated statistically by the Kishino-Hasegawa-Templeton likelihood ratio test (25), following the assignment of log likelihoods (program DNAML) to artificially generated topologies (program RETREE) (unpublished data).

Nucleotide sequence accession numbers. Nucleotide sequences reported in this study have been assigned the GenBank accession numbers U24717 to U24999 and U25001 to U25025.

RESULTS

Clinical statuses of patients. (i) Mothers. Mothers A, B, and D were of European origin and were formerly intravenous drug users. Mother C was of African origin (Angolan), had lived in Zaire, and is most likely to have been infected heterosexually. Nevertheless, since the HIV status of her partner was unknown, a risk factor could not reliably be identified for this woman. The clinical status of the mothers and children, the times of sampling, levels of cell and plasma viremia, and immunological data are presented in Table 1. At the commencement of pregnancy, mothers A, B, and D had been infected for at least 25 months, 5 years, and 7 months, respectively. Diagnosis of infection was performed at 5 months of pregnancy for mother C. All mothers were asymptomatic at the beginning of pregnancy (Centers for Disease Control and Prevention [CDC] stages II and III) (6), although mothers B and D became symptomatic (oral hairy leukoplakia and multidermal herpes zoster, respectively) during late pregnancy, at which point they were reclassified as CDC stage IVc2. The $CD4^+$ cell counts among the mothers varied between 276/mm^3 and 858/mm^3 with no significant variation in the mothers during pregnancy (Table 1). However, a significant decrease in the $CD4^+$ cell count was observed to occur in mother D 1 year after delivery (32 $CD4^+$ cells per mm^3), although this was not accompanied by further modification in clinical status (data not shown). Mothers A, B, and D were negative for plasma p24 antigen throughout pregnancy, but mother C was positive at all time points (128, 8, and 6 pg/ml for time points MC1 (time point 1 for mother C), MC2, and MC3, respectively). Cell viremia was observed throughout pregnancy in all four mothers studied (Table 1). In contrast, plasma viremia was not present in mothers B and D and was only transiently observed at low levels in mothers A and C (Table 1). In all cases, pregnancies were free of complications and all children were delivered vaginally. The mothers did not receive any antiretroviral therapy before or during pregnancy.

(ii) Newborns. Children A, B, and C were diagnosed as HIV positive (PCR and coculture positive) during the first 2 weeks of life (at 6, 5, and 12 days of age, respectively). Child D was negative by PCR and coculture at birth but became HIV positive at 1 month of age. All children were originally classified as CDC stage P2A (7). Children B, C, and D remained asymptomatic for the duration of the study (24 months) (Table 1). In contrast, child A displayed neurological signs at 1 month of age and was reclassified as CDC stage P2B at 3 months (Table 1). This child developed AIDS within the first year of life, with evidence of cryptosporidiosis and multidermal herpes zoster and was reclassified as CDC stage P2D at 16 months of age. All children received zidovudine treatment (50 to 70 mg three times a day). Children A, B, and C were treated with zidovudine from the age of 7 weeks, whereas child D received antiretroviral therapy from the age of 14 weeks.

Sequence diversity. We have analyzed genetic variation in 226 complete sequences spanning the V3 region and flanked 5′ and 3′ by 102 and 114 bp, respectively. Four incomplete sequences lacking only a few bases at the 5′ and/or 3′ end which did not lack any variable region were also included.

The within-sample genetic diversity and between-sample genetic distance were calculated for each pairwise comparison between sequences from the four mother-child pairs (Fig. 1 and 2). To compare the diversity between the mothers' samples and those of the infants, we have plotted each value as a histogram (Fig. 1). The overall means for all sequences obtained from each patient (Fig. 1e) lay between 3.5 and 5.2% for the mothers' samples, with the between-mother distances (5.3 to 8.2%; mean, 7.2%) (Fig. 2e) being greater, as expected. The within-child diversity had a greater range (0.74 to 8.8%) (Fig. 1e), and the between-child distances were mostly greater (7.6 to 10.8%; mean, 8.8%) than for the mothers (Fig. 2f). Comparing the overall within-patient diversity for the mothers with that of their infants (Fig. 1) revealed two pairs (pairs B and D) where the infants' samples were substantially less diverse than their mothers' (Fig. 1b and d), one where they were similar (Fig. 1a, pair A) and one (Fig. 1c, pair C) where substantially greater diversity was found in the child. The within-sample

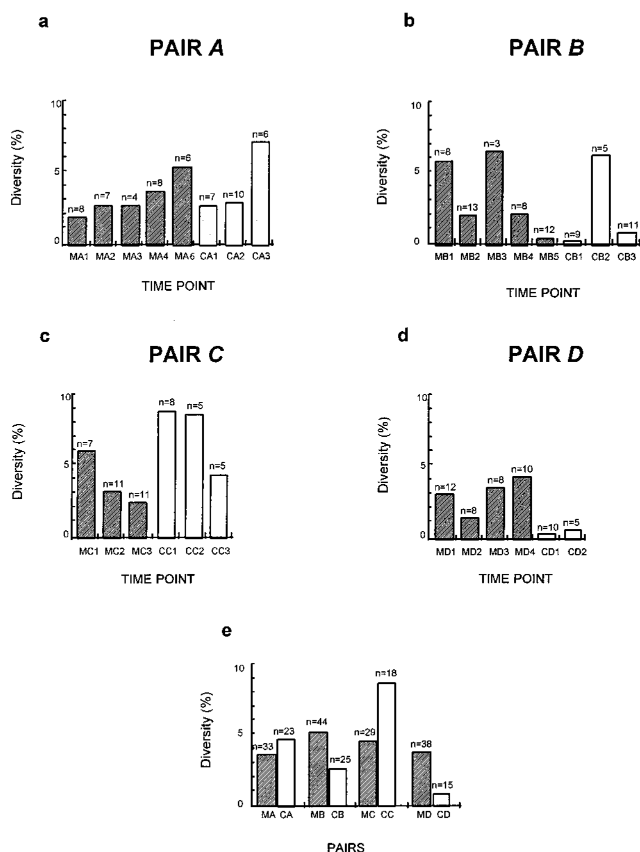

FIG. 1. Within-sample and between-sample nucleotide sequence diversity. (a to d) Within-sample diversity in mothers (dark grey columns) and children (light grey columns). (e) Mean within-patient diversity. The number of sequences for each time point is indicated at the top of each column.

genetic distances showed considerable variation within each patient. For three of the four pairs, the first child sample showed less diversity than the mother's sample closest to delivery (Fig. 1, pairs A, B, and D). However, exactly the opposite was true for pair C (Fig. 1c). In the mothers' samples, a progressive increase of genetic distance with time was found to occur in pair A (Fig. 2a; entries off-diagonal), with the distance from sample 1 to sample 6 (time interval, 7.5 months) almost twice that from sample 1 to sample 2 (interval, 1 month). Again, no such trend was observed with the other samples (Fig. 2b through d). Examination of individual values showed that divergence between samples from the same patient can reach levels similar to those found when comparisons between different patients are made. Clearly, a more detailed form of analysis is required to identify any evolutionary pattern that may exist in these data.

Phylogenetic analyses. The sequences of the four mother-child pairs were classified according to HIV-1 global subtype classification by phylogenetic comparison with reference sequences from the five designated *env* subtypes (33). The sequences of all four mother-child pairs (including sequences from mother C, who originated from Africa) were of subtype B (European or North American) (data not shown). The rarity of a B clade virus in the African-derived isolates makes it more likely that the infection of mother C was acquired in Europe.

(i) Mother-child pair A. Analysis of V3 sequences from the first transmission pair (pair A) led to the generation of a "star" phylogeny in which the different lineages radiated from a single point. The neighbor-joining tree reconstructed from the data set is presented in Fig. 3 (pair A). Similar topologies were obtained by each of the three methods of tree construction employed, as well as with the tree reconstructed on the basis of the amino acid sequence data (data not shown). Internal branches within the phylogeny were short, and bootstrap values across the tree were very low. As a consequence, there was little statistical support for any specific grouping within the tree, although three groups of child sequences were apparent. The first group (Fig. 3, pair A, child group 1) appeared to consist predominantly of early-time-point sequences (1 and 2.5 months of age), although sequences from the child's 16-month sample were also associated with the group. There appeared to be some association of these sequences with a number of second- and third-trimester sequences from the mother (4.5, 6, and 7 months). The second child group (Fig. 3, pair A, child group 2) consisted of three 2.5-month sequences which clustered within the main group of maternal sequences detected during pregnancy (3.5 and 4.5 months) but appeared to be most closely related to a small number of maternal sequences from the third trimester. The third group (Fig. 3, pair A, child group 3) consisted predominantly of late-time-point sequences (16 months) which appeared to be associated with sequences from the 2-month-postdelivery sample of the mother.

(ii) Mother-child pair B. For the second transmission pair (pair B), a clear division of all sequences into two distinct groups was apparent from the neighbor-joining tree (Fig. 3, pair B). All three tree-building methods gave similar trees from the nucleotide sequences, and the division into two distinct groups was also clearly apparent from the amino acid neighbor-joining tree (data not shown). The branch separating the two groups was resolved in 100% of bootstrap replicates, and child sequences were associated with both groups. The first sequence group (Fig. 3, pair B, group 1) formed the predominant lineage present within the maternal population throughout pregnancy and included sequences from all time points from 3.5 months into pregnancy until delivery. Only three child sequences, isolated from the child at the age of 5 weeks, were associated with this group. The majority of the child sequences were associated with the second group (Fig. 3, pair B, group 2), with sequence isolates from all child time points (5 days, 5 weeks, and 3.5 months). The maternal sequences associated with this group were mainly from the 3-month-postdelivery sample, whose sequences formed a tight cluster, although the variant was present as a minor form throughout pregnancy. It was interesting that the group 1 maternal variant which was predominant at delivery had been replaced with the group 2 sequence type by 3 months postdelivery.

(iii) Mother-child pair C. Two child groups were also clearly apparent in the neighbor-joining tree for the third transmission pair (Fig. 3, pair C). The main child group (Fig. 3, pair C, child group 1), consisting solely of sequences found at 1.5 and 2.5 months after birth, appeared to be descended from a lineage present within the mother during the second trimester (5- and 6.5-month maternal samples) (mother group 1). The lineage was represented by 4 of 7 sequences from the 5-month maternal sample and in only 1 of 11 sequences from the 6.5-month maternal sample. It was not represented in the 7.5-month sample. A single child sequence from the 13-month sample was also weakly clustered with this group. The level of bootstrap support for the cluster incorporating the main group of child and ancestral mother sequences was reasonably high, with the cluster supported in 87.4% of replicates. Except for one sequence (designated by an arrow in Fig. 3, pair C), the cluster was found in 96.1% of bootstrap replicates. The second minor child group (Fig. 3, pair C, child group 2) consisted of four (of

a

PAIR *A*

	MA1	MA2	MA3	MA4	MA5	MA6	CA1	CA2
MA2	0.025							
MA3	0.026	0.031						
MA4	0.032	0.035	0.034					
MA5	0.033	0.037	0.032	0.040				
MA6	0.046	0.049	0.047	0.055	0.051			
CA1	0.044	0.045	0.046	0.042	0.051	0.064		
CA2	0.039	0.041	0.041	0.037	0.048	0.060	0.031	
CA3	0.061	0.065	0.065	0.069	0.073	0.070	0.067	0.068

b

PAIR *B*

	MB1	MB2	MB3	MB4	MB5	CB1	CB2
MB2	0.055						
MB3	0.057	0.041					
MB4	0.053	0.022	0.042				
MB5	0.055	0.088	0.069	0.085			
CB1	0.055	0.089	0.069	0.086	0.004		
CB2	0.058	0.054	0.058	0.051	0.057	0.057	
CB3	0.057	0.089	0.071	0.085	0.010	0.009	0.059

c

PAIR *C*

	MC1	MC2	MC3	CC1	CC2
MC2	0.063				
MC3	0.064	0.033			
CC1	0.115	0.110	0.111		
CC2	0.108	0.103	0.103	0.080	
CC3	0.072	0.048	0.043	0.106	0.099

d

PAIR D

	MD1	MD2	MD3	MD4	CD1
MD2	0.033				
MD3	0.037	0.036			
MD4	0.042	0.038	0.042		
CD1	0.054	0.057	0.046	0.045	
CD2	0.050	0.052	0.042	0.042	0.009

e

MOTHERS

	MA	MB	MC
MB	0.053		
MC	0.082	0.062	
MD	0.082	0.079	0.076

f

CHILDREN

	CA	CB	CC
CB	0.083		
CC	0.108	0.090	
CD	0.084	0.076	0.088

FIG. 2. Nucleotide distances in the V3 loop and flanking regions of HIV-1 among mother-child transmission pairs. Nucleotide distances were estimated for each pairwise sequence comparison by using the generalized two-parameter model (15). (a to d) Nucleotide distances between mother and child time point sequences. (e) Mean nucleotide distances between all maternal sequences. (f) Mean nucleotide distances between all child sequences.

five) late-time-point sequences (13 months of age). The group was located within the main cluster of the maternal sequences (Fig. 3, pair C, mother group 2), closest to those found late in pregnancy (7.5 months), although only a low level of bootstrap support for this was indicated. The overall level of bootstrap support for the main mother group itself was high (96.1%, excluding intermediate sequences). The phylogenies reconstructed by alternative methods for mother-child pair C were again consistent (data not shown).

(iv) Mother-child pair D. For pair D, the three methods of phylogeny reconstruction employed all identified a single child group (Fig. 3, pair D). The sequences from both the 1.5- and 11.5-month infant samples were clustered fairly tightly within the phylogeny and showed an association with a number of maternal sequences from the 8.5- and 4.5-month-postdelivery samples. These sequences represented a minor form in both maternal samples (2 of 9 sequences from the 8.5-month sample and 2 of 10 sequences from the 4.5-month-postdelivery sample). Four maternal sequences, three from the 4.5-month-postdelivery sample and one from the 2-month sample, were intermediate between the mother-child group and the main maternal cluster. The presence of these intermediate sequences reduced the bootstrap support, but if they were excluded the branch separating the main maternal group and the mother-child group was found in 99.2% of bootstrap replicates. Once again the phylogenies inferred by the three methods of phylogeny reconstruction employed were highly congruent, and again the amino acid neighbor-joining tree was very similar to that reconstructed on the basis of nucleotide sequence data (data not shown). Testing the relative likelihoods of these phylogenetic hypotheses (23) confirmed the results of the bootstrap analyses.

Amino acid sequence heterogeneity between mothers and children. The amino acid sequence alignments for the four mother-child pairs are presented in Fig. 4. For each pair, a consensus sequence was constructed for each time point by assigning the amino acid most frequently observed in the clones to each position. When the phylogenetic analysis revealed more than one group within a time point, sequences from that time point were divided into clusters of related

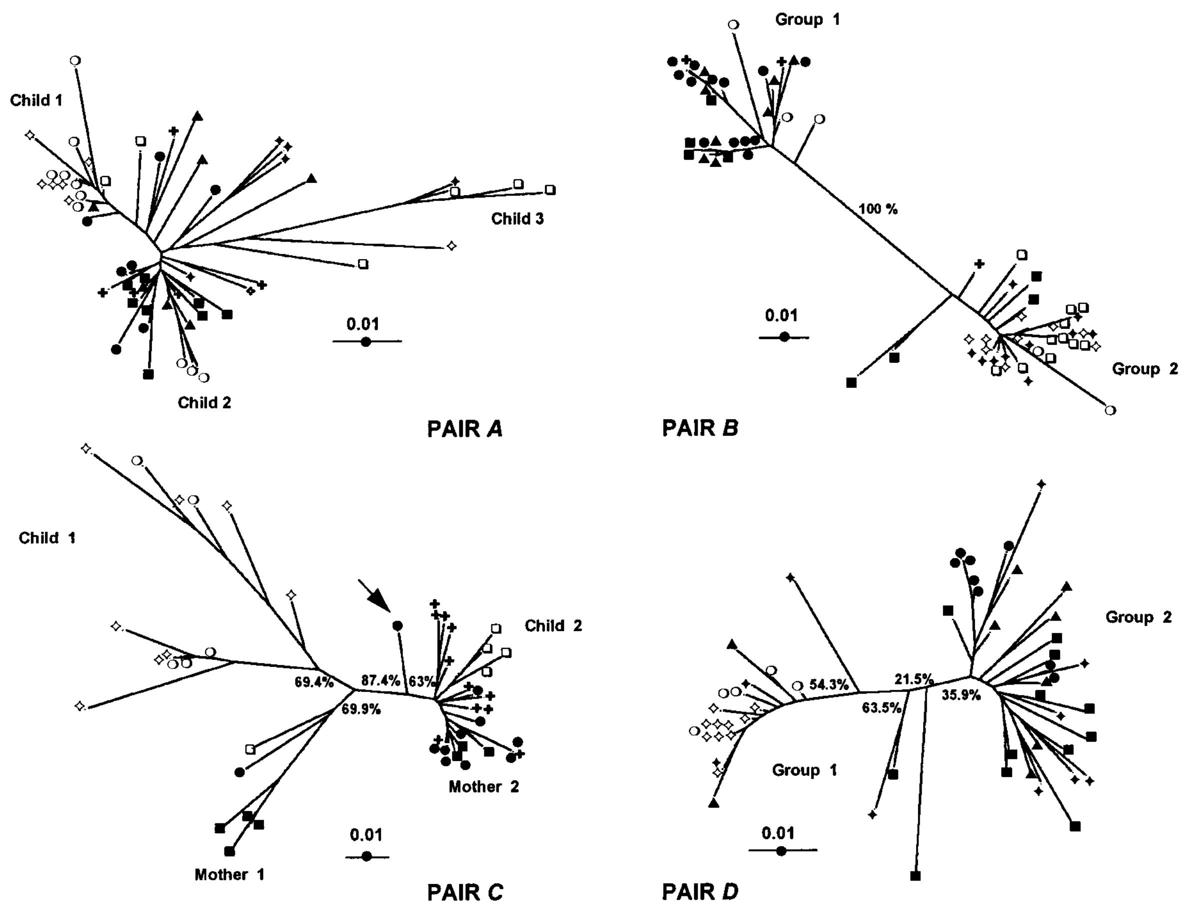

FIG. 3. Unrooted neighbor-joining trees for the four mother-child transmission pairs. Symbols at the tip of each branch denote the time point to which the sequence belongs. Open symbols represent one individual child sequence, and shaded symbols represent one individual maternal sequence. All branch lengths are drawn to scale. Bootstrap values are expressed as percentages for each branch and represent the percent occurrence of that branch per 2,000 bootstrap replicates. Symbols for pair A: ■, MA1; ●, MA2; ✚, MA3; ▲, MA4; ❖, MA5; ✦, MA6; ✧, CA1; ❍, CA2; ❑, CA3. Symbols for pair B: ■, MB1; ●, MB2; ✚, MB3; ▲, MB4; ✦, MB5; ✧, CB1; ❍, CB2; ❑, CB3. Symbols for pair C: ■, MC1; ●, MC2; ✚, MC3; ✧, CC1; ❍, CC2; ❑, CC3. Symbols for pair D: ■, MD1; ●, MD2; ▲, MD3; ✦, MD4; ✧, CD1; ❍, CD2.

clones. All consensus sequences were aligned in relation to the first time point consensus of the mother. The phylogenetic group to which each consensus sequence corresponds is reported at the right edge of Fig. 4.

For pair A, the three groups of child sequences found in the phylogenetic tree (Fig. 3, pair A) were associated with amino acid substitutions within both V3 and the flanking regions (positions 308, 320, 335, 339, 343, 346, and 360). Comparison of amino acid sequences of the mother with those of the child revealed a greater similarity between child group 1 sequences and maternal sequences detected late in pregnancy (Fig. 4a, MA4.C). Child group 2 sequences (Fig. 4a, CA2.C2) shared amino acid variants with maternal sequences from various time points (positions 320, 339, 346, and 360); however, no common amino acid sequence pattern was apparent between child group 3 sequences and the maternal sequence set.

For pair B, two distinct groups of sequences which correspond to the phylogenetic groupings observed in the neighbor-joining tree (Fig. 3, pair B) were apparent, on the basis of their amino acid sequences (Fig. 4b). The two groups were characterized by the following amino acid sequence patterns: group 1, T-283, S-291, S-300, R-305, S-306, T-308, T-317, K-342, V-345, and T-360; group 2, S-283, T-291, N-300, K-305, G-306, H-308, A-317, R-342, A-345, and N-360.

For pair C, the two subgroups of child sequences (Fig. 3, pair C) were characterized by distinct amino acid sequence patterns (Fig. 4c). The major amino acid differences between the two subgroups occurred at positions 289 (N or K) and 308 (D or S). Child group 2 sequences were mostly characterized by the presence of a basic lysine residue (K) at position 313. This modified the GPGR motif at the crown of the V3 loop to GPGK. A high degree of similarity was observed between mother group 2 and child group 2 sequences.

For pair D, the child sequences, which clustered in a single group within the phylogenetic tree (Fig. 3, pair D), were characterized by the following amino acid sequence pattern: K-293, S-306, K-313, A-317, D-320, D-324, K-341, N-346, G-349, K-360 (Fig. 4d). A small number of maternal sequences present as a minor form in the late-pregnancy–postdelivery maternal samples shared this amino acid sequence pattern (Fig. 4d, MD3.C2 and MD4.C2). The lysine residue observed at position 313, which gave rise to the GPGK motif at the crown of the V3 loop, was particularly characteristic of the group. This was in contrast to the GPGR motif observed in the majority of maternal sequences, which were also more heterogeneous than those of the child. The amino acid sequences of the main maternal group (Fig. 4d, group 2) diverged from those of the child by many amino acid substitutions (positions 306, 313, 317, 320, 349, and 360).

(i) Pattern of potential N-linked glycosylation between

a. PAIR A

```
               270        280      290        300      310       320       330       340       350       360         n  Phyl. group
MA1.C    GSLAEEEVVIRSENFTNNAKTIIVQLKESVEINCTRPSNNTRRSITIGPGRAFYTTGDIIGDIRQAHCNISRAKWNETLKQIVKKLGEQFKNKTIVFKQSSGGDPE 8

MA2.C    ---------------------------------------------------------------------------------------------------------- 7

MA3.C    ----------------------------------------------------------------------------D----------------------------- 4

                                                                                            K
MA4.C    ------------------------------------L---------------------------------------N------E---D------------------ 8

                                                                    T                    KV        .
MA6.C    ----------------D----------------------------N-------------I----------------D---QI-V------K--------------- 6

CA1.C    --------------I---------------------L--------N-----------A--------------V---N------E---D---------T-------- 7 1

CA2.C1   --------------I---------------------L--------N-----------A--------------V---N------E---D---------T-------- 7 1
CA2.C2   --------------I---------------------L--------------------A--------------E---D---H---------R--------------- 3 2

                       I
CA3.C    --------------F-----S-----N-T--------N----KG-H--------A---------------------D--R--AI-------------N-------- 11 3
```

b. PAIR B

```
               270        280      290        300      310       320       330       340       350       360         n  Phyl. group
MB1.C1   GSLAEEEVVIRSENFTNNAKTIIVQLKESVEINCTRPSNNTRRSITIGPGRAFYTTGDIIGDIRQAHCNISRAKWNDTLKQIVIKLGEQFKNKTIVFKQSSGGDPE 4 1
MB1.C2   --------------------S-----N-T--------N----KG-H--------A--D---D-----------------R--A--------------K-------- 4
                                                                  R   N                                   N

MB2.C    ---------------------------K-----------------------------A-T----------------E------K---------------------- 13 1

MB3.C    --------------------------NK------------------------------------------------E------?---------------------- 2 1
MB3.1    --------------------s-----n-t--------n----kg-h--------a------------------------r----------e----a---------- 1 2

MB4.C    ----------------------------------------------------------------------------E------K---------------------- 8 1

MB5.C    --------------------S-----N-T--------N----KG-H--------A------N----------S------R--A--------------N-------- 10 2

CB1.C    --------------------S-----N-T--------N----KG-H--------A------N----------S------R--A--------------N-------- 9 2

CB2.C1   ------------------------------------L----------------------------------------------K-------------T-------- 3 1
CB2.C2   --------------------S-----N-T--------N----KG-H--------A------N----------S------R--A--------------K-------- 2 2
                                                                                                          N

CB3.C    --------------------S-----N-T--------N----KG-H--------A------N----------S------R--A--------------N-------- 10 2
```

c. PAIR C

```
               270        280      290        300      310       320       330       340       350       360         n  Phyl. group
MC1.C1   GSLAEEEVVIRSENFTNNAKTIIVQLKESVEINCTRLSNNTRRSINIGPGRAFYTTGAIIGDIRQAHCNISRVKWNNTLKQIVRKLGEQFKNKTIAFNQSSGGDPE 4 1
MC1.C2   --------------------S-----N-T-------PN----KG-H--------A--D---N----------S---D--R--AI-------[_]---[_]------ 3 2

MC2.C    --------------------S-----N-T-------PN----KG-H--------A--D---N----------S---D--R--AI-------[_]-V-[_]------ 9 2
MC2.1    ----------------d-t-----------------pn----k--------------------------l--a---d-----a-------g[_]-v-[_]------ 1

MC3.C    --------------------S-----N-T-------PN----KG-H--------A--D--------------A---D--R--AI-------[_]-V-[_]------ 11 2

CC1.C1   ------DI----D---D---------NQ--V-----PN----KD-H--------A--E-----------L---Q-----E--AT--R---NK---V-K-------- 5 1a
CC1.C2   --------------I-D-T---------AI------PN----K--H------W-K--D--------------A------E---K--R----[_]-V-[_]------ 3 1b

CC2.C    --------------I-D-T---------AI------PN----K--H------W-K--D--------------A------R---K--R----[_]-V-[_]------ 5 1b

                                                                                                       V
CC3.C    --------------------S-----N-T-------PN----KG-H----K---A--D--------------A---D--R--AI-------[_]-A-[_]------ 7 2
```

d. PAIR D

```
               270        280      290        300      310       320       330       340       350       360           n  Phyl. group
                                                                                                      K
MD1.C    GSLAEEEVVIRSENFTNNAKTIIVQLKEPVEINCTRPNNNTRKGIHIGPGRAFYTTGEIIGNIRQAHCNLSRAKWNDTLNQIVVKLREQEQFKNTTIAFNQSSGGDPE 11 2

MD2.C    --------------L-------------S-------------------------------------------S------K---I---D..----K----[_]------ 8 2

MD3.C1   -------------------------------------------------------------------------------K--------..----K----[_]------ 6 2
                e                                     h                   a                    g     d    i
MD3.C2   -------v----------------------K------------S-p----K---A--D---D---p-------------K---N--r-..--k-K--v-K-------- 2 1

MD4.C1   ------------------------------K------------------------------------------------K---N----..----K----k-------- 8 2
                                                           r                     e                        i
MD4.C2   ------------------------------K------------S-P----k---A--D---D----------a------K---N--G-..----K--v-n-------- 2 1

CD1.C    ------------------------------K------------S-P----K---A--D---D-----------------K---N--G-..----K--V-K-------- 10 1

CD2.C    -------------------------------------------S------K---A--D---D-----------------K---N--G-..----K--V-K-------- 10 1
```

mothers and children. Comparison of mother and child amino acid sequences revealed common patterns of potential N-linked glycosylation in the four mother-child pairs (Fig. 4). Six potential glycosylation sites, located at positions 276, 295, 301, 331, 338, and 354, were perfectly conserved in all groups of sequences from pairs A, B, and D. However, in pair C, the potential glycosylation site at position 354 was absent in a small number of child sequences (Fig. 4c, CC1.C). Two additional sites at positions 289 and 360 differed between the mother and child. The sites appeared only in child late-time-point sequences in pair A (Fig. 4a, CA3.C). For pair B, the sites were present in the majority of mother and child amino acid sequences associated with phylogenetic group 2 (Fig. 4b), the main group present within the child. By contrast, the sites were absent in mother and child sequences associated with group 1 (Fig. 4b). For pair C, the potential N-linked glycosylation site at position 289, present in the main maternal group, child group 2, and child subgroup 1a sequences, was absent in sequences of maternal group 1 and child subgroup 1b. The site at position 360 was detected in the majority of mother and child sequences, with the exception of child subgroup 1a sequences (Fig. 4c). For pair D, the potential N-linked glycosylation site at position 360, which was absent in the child, was present in all maternal sequences with the exception of sequences MD3.C2 and MD4.C2 (Fig. 4d). These sequences were most closely associated with those of the child. The site at position 289 was absent in both the mother and child.

(ii) Potential phenotype of transmitted viral species. The potential phenotype of the amino acid sequence variants was predicted on the basis of the global net charge of the V3 loop and the degree of sequence divergence from the La Rosa subtype B consensus (11, 30). The majority of sequences from the mother-child pairs were predicted to be of the macrophage-tropic, non-syncytium-inducing (NSI) phenotype. However, a small number of sequences predicted to be of the T-cell-tropic, syncytium-inducing (SI) phenotype were observed to be present in pairs A and B. Five potential SI variants were observed in pair A, and these consisted of a single early-pregnancy maternal sequence (time point MA1) and four child sequences (time points CA1, CA2, and CA3). For pair B, four potential SI variants were observed to occur in the mother (time points MB1, MB2 and MB4) and a single predicted SI variant was observed at the first time point in the child (time point CB1).

DISCUSSION

In order to examine the molecular mechanisms involved in HIV-1 mother-to-child transmission, we have analyzed the genetic relationships among cell-associated viral populations detected during pregnancy in four HIV-infected mothers and in their respective children. This analysis was performed with a 313-bp fragment containing the V3 region and the highly informative flanking sequences. We have compared proviral DNA sequences obtained from the mothers at different time points during pregnancy and postdelivery with child sequences obtained from birth.

Genetic diversity and possible multiple transmission. Sequence data obtained from the V3 region for two of the four children sampled in this study (pairs A and C) revealed substantial levels of genetic heterogeneity in their cell-associated viral populations. Such heterogeneity has not been a prominent feature of earlier studies of mother-child transmission (31, 42, 50). Child B also showed significant viral heterogeneity when all samples from the child were included. Only child D showed substantially less diversity than found in its mother. The occurrence of such heterogeneity in the newly infected infants is quite unlike the situation in newly infected adults, several studies having shown very restricted levels of variation in the *env* gene in the peripheral blood of hemophiliacs and patients infected by sexual contact (4, 22, 45, 51). This has been interpreted as evidence for selection for particular viral variants from the heterogeneous pool present in most individuals later in infection. The striking difference between the situation described for these children and that observed for adults suggests that mother-child transmission may be a more complex process.

Phylogenetic analysis of the V3 sequence data has shed more light on the circumstances of transmission for each of the four mother-child pairs. For pair A, three groups of viral sequences within the child were identified, but as they were only weakly defined it was not possible to draw any specific conclusions regarding transmission. Nevertheless, the heterogeneity of the viral populations of both the mother and the child does not support the view that the infection of the child involved a single viral variant. It is interesting that the mother showed a high level of cell viremia throughout pregnancy (Table 1), which could have facilitated the transmission of multiple variants to the child.

For pair B, both the statistical analysis of the reconstructed phylogeny and the amino acid sequence alignments clearly indicated the occurrence of two very distinct populations within both the mother and the child. The simultaneous presence of both populations within the 1.5-month sample of the child, which was responsible for the particularly high within-sample diversity for this time point, clearly indicated that the infection within the child was the result of the transmission of two distinct maternal variants. We cannot infer from the data whether the two variants were transmitted simultaneously or at different times, as both maternal sequences with which the two child groups were associated were present throughout pregnancy.

The analysis of pair C also revealed an unusually high level of genetic diversity within the child sequences obtained within the first year of life (8 to 10%). Two statistically significant groups of child sequences were apparent from the phylogenetic tree, and specific maternal sequences were clearly associated with each group. The major child lineage was associated with maternal sequences from 5 and 6.5 months. Other child sequences were more closely related to maternal sequences obtained later. PCR and culture diagnosis for child C, performed at 12 days postdelivery, provided further evidence for early infection.

The results of the phylogenetic analysis are compatible either with early infection with multiple variants or with transmission of virus on more than one occasion, possibly during

FIG. 4. Amino acid sequence alignments. Consensus sequences were deduced for each phylogenetic group present within a time point and aligned in relation to the first time point consensus of the mother. The number of individual sequences within a consensus and the phylogenetic group to which each sequence corresponds are shown at the right. Consensus sequences are presented in uppercase, whereas individual clone sequences are given in lowercase. Amino acids are numbered according to their position in the HIV-LAI genome (33). Potential N-linked glycosylation sites conserved between mother and child sequences are indicated by shaded boxes, and variable N-linked glycosylation sites are underlined. A "•" indicates the deletion of a codon, "n" indicates the number of clones represented by each consensus, and "?" indicates that no consensus amino acid residue could be defined at a given position.

pregnancy and at delivery. Although transmission of more than one variant between adults appears to be rare, we note that there is evidence that it occurred for one patient of a Florida dentist (27, 36) and also for a victim of rape (1).

Evidence for selective mother-to-child transmission. In contrast to the situation for the other three cases we have studied, the cell-associated viral sequences detected in child D were highly homogeneous at all time points. These sequences clustered within a single group in the phylogenetic tree and were closely related to a small number of maternal sequences detected during late pregnancy and following delivery. This suggests that the infection within this child was the result of the transmission of a single maternal variant. The fact that these maternal genotypes were found only in late pregnancy and after delivery is consistent with a later infection, possibly at delivery. This hypothesis is supported by the fact that HIV infection within child D could not be detected at birth but was first diagnosed at 1 month of age (Table 1). It is interesting that the sequences observed to be present in this child were highly homogeneous and remained closely associated with sequences detected at the time of diagnosis for at least a year after birth, whereas samples obtained from infected children in the other transmission pairs at approximately 1 year of age were more heterogeneous. These results therefore support the hypothesis of selective transmission of a minor maternal variant in pair D, as indicated in other mother-child transmission studies (31, 42, 50).

Implications of V3 amino acid sequence variation in transmission. For the three mother-child pairs in which multiple variants appeared to be transmitted to the child (pairs A, B, and C), comparison of amino acid sequences from the mothers with sequences from their respective children did not identify any overall pattern distinguishing between transmitted and nontransmitted viral species. We did not observe the selective loss of any glycosylation site within the V3 region in variants transmitted to the children. In particular, the potential N-linked glycosylation site proximal to the first cysteine of the V3 loop (position 295), absent in the infant sequence sets described by Wolinsky et al. (50), usually remained conserved between mother and child sequences in our study. This observation has also been reported by other investigators (31, 42). Glycosylation site differences between mother and child were, however, observed to occur in the V3 flanking regions at positions 289 and 360. The GPGR motif at the crown of the V3 loop showed a high degree of conservation within the maternal and child sequences of pairs A, B, and C, probably because of the functional importance of the region. However, a conversion to GPGK was observed to occur in child C at 1 year of age.

In contrast, in pair D, where selective transmission of a minor maternal variant was observed, specific amino acid sequence variations were identified between the mother and child sequences. The N-linked glycosylation site at position 360, present in the majority of maternal sequences, was absent from the sequences of the child. Moreover, the GPGR motif which was present in most of the mother's sequences was converted to GPGK in all child sequences. These amino acid sequence variants were also observed to be present in the minor subset of maternal sequences highly related to the transmitted subtype. Genetic variations within V3 have been found to influence host immune responses to antibody (34, 52) and antibody titers, as well as affinities for epitopes within V3. These factors have been implicated as important in HIV-1 perinatal infection (10, 37). The mutations observed to occur between the mother and child of pair D may suggest that the transmitted variant had been subjected to immune selection within the mother, perhaps conferring a selective advantage for transmission. The observed sequence variations in pair D do not themselves, however, localize the determinants accounting for the immune escape of the virus. Other biological tests, particularly assays based on analysis of maternal neutralizing antibodies and maternal cellular immunity, would be required to confirm the immune escape hypothesis.

The V3 region has also been shown to influence the ability of the virus to replicate in macrophages and to grow in transformed T-cell lines (7, 17, 24, 43). Several articles have reported that the emergence of basic amino acids at specific positions within V3 is associated with the phenotypic shift of the virus from the NSI, macrophage-tropic form to the SI, T-cell-line-tropic form (7, 9, 11, 30). Amino acid sequences characteristic of the macrophage-tropic NSI phenotype predominated in the mother-child sequence sets. A few individual maternal and child sequences in pairs A and B exhibited basic amino acids at critical positions for tropism, potentially conferring a T-cell-tropic, SI phenotype. Macrophage-tropic viruses have been shown to be transmitted sexually and to be responsible for establishing chronic infection (51).

It is interesting that child A rapidly showed neurological symptoms and developed AIDS within 24 months. A small number of sequences with a potential SI phenotype were detected in this child at 1 month of age and remained when the child was 16 months old. Whether these variants were transmitted from the mother or whether they emerged from sequence variations of transmitted viral species within the child could not be determined. Overall, no association between HIV sequence diversity and disease evolution has been established. However, the rapid progression to AIDS in child A might be associated with the presence of T-cell-tropic, SI variants soon after birth, as has been reported for adult infections (26, 47, 48). This observation is in good agreement with the rapid decline in $CD4^+$ T-cell count and progression to AIDS reported previously to occur in patients showing a phenotypic shift of the virus from the NSI phenotype to SI.

In conclusion, our analysis has shown that the genetical processes involved in vertical HIV-1 transmission may be more complex than those found to occur in transmission between adults. In one case from our study, the selective transmission of a single maternal variant appeared to have occurred. However, in two other pairs, infection of the child by at least two maternal variants was demonstrated. These multiple transmissions may have occurred at different times during pregnancy. Analysis of mother-to-child transmission of HIV infection in relation to maternal HIV antibodies as well as comparison of biological properties of variants transmitted and not transmitted to the child could lead to a better understanding of viral determinants involved in HIV-1 vertical transmission.

ACKNOWLEDGMENTS

Special thanks are given to J. Tricoire from the division of Pediatrics at CHU Purpan and A. Berrebi from the division of Gynecology-Obstetrics at La Grave Hospital, who provided maternal and child blood samples. We also acknowledge the help provided by Denis Lobidel at the Centre for HIV Research, Edinburgh, United Kingdom. We also thank J. F. Magnaval at CHU Purpan for helpful assistance.

L.B. is a fellow of the Agence Nationale de Recherches sur le SIDA. This work was supported by the Agence Nationale de Recherches sur le SIDA, the Medical Research Council AIDS Directed Programme, and the Conseil Régional de Midi-Pyrénées.

REFERENCES

1. **Albert, J., J. Wahlberg, and M. Uhlén.** 1993. Forensic evidence by DNA sequencing. Science **361:**595–596.
2. **Alizon, M., S. Wain-Hobson, L. Montagnier, and P. Sonigo.** 1986. Genetic variation of the AIDS virus: nucleotide sequence analysis of two isolates

from African patients. Cell **46:**63–74.
3. **Auger, I., P. Thomas, V. de Gruttola, D. Morse, D. Moore, R. Williams, B. Truman, and C. E. Lawrence.** 1988. Incubation periods for pediatric AIDS patients. Nature (London) **336:**575–577.
4. **Balfe, P., P. Simmonds, C. A. Ludlam, J. O. Bishop, and A. J. Leigh Brown.** 1990. Concurrent evolution of human immunodeficiency virus type 1 in patients infected from the same source: rate of sequence change and low frequency of inactivating mutations. J. Virol. **64:**6221–6233.
5. **Burger, H., B. Weiser, K. Flaherty, J. Gulla, P. Nguyen, and R. Gibbs.** 1991. Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity among close contacts. Proc. Natl. Acad. Sci. USA **88:**11236–11240.
6. **Centers for Disease Control and Prevention.** 1987. Classification system for human immunodeficiency virus (HIV) in children under 13 years of age. Morbid. Mortal. Weekly Rep. **36:**225–235.
7. **Chesebro, B., K. Wehrly, J. Nishio, and S. Perryman.** 1992. Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical amino acids involved in cell tropism. J. Virol. **66:**6547–6554.
8. **Courgnaud, V., F. Lauré, A. Brossard, C. Bignozzi, A. Goudeau, F. Barin, and C. Bréchot.** 1991. Frequent early *in utero* HIV-1 infection. AIDS Res. Hum. Retroviruses **7:**337–341.
9. **De Jong, J. J., A. De Ronde, W. Keulen, M. Tersmette, and J. Goudsmit.** 1992. Minimal requirement for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. J. Virol. **66:**6777–6780.
10. **Devash, Y., T. A. Calvelli, D. G. Wood, K. Reagan, and A. Rubinstein.** 1990. Vertical transmission of human immunodeficiency virus is correlated with the absence of high affinity/avidity maternal antibodies to the gp120 principal neutralizing domain. Proc. Natl. Acad. Sci. USA **87:**3445–3449.
11. **Donaldson, Y. K., J. E. Bell, E. C. Holmes, E. S. Hughes, H. K. Brown, and P. Simmonds.** 1994. In vivo distribution and cytopathology of variants of human immunodeficiency virus type 1 showing restricted sequence variability in the V3 loop. J. Virol. **68:**5991–6005.
12. **Ehrnst, E., S. Lindgren, M. Dictor, B. Johanson, A. Sonnerborg, J. Czajkowsky, G. Sundin, and A. B. Bohlin.** 1991. HIV in pregnant women and their offspring: evidence for late transmission. Lancet **338:**203–206.
13. **Felsenstein, J.** 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17:**368–376.
14. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791.
15. **Felsenstein, J.** 1993. PHYLIP manual version 3.52c. Berkeley University Herbarium, University of California, Berkeley.
16. **Fitch, W. M., and E. Margoliash.** 1967. Construction of phylogenetic trees. A method based on mutation distances as estimated from cytochrome c sequences is of general applicability. Science **155:**279–284.
17. **Fouchier, R. A. M., M. Groenink, N. A. Koostra, M. Tersmette, H. G. Huisman, F. Miedema, and H. Schuitemaker.** 1992. Phenotype-associated sequence variations in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. J. Virol. **66:**3183–3187.
18. **Gupta, P., M. Brady, M. Raabe, and A. Urbach.** 1991. Detection of human immunodeficiency virus by virus culture and polymerase chain reaction in children born to seropositive mothers. J. Acquired Immune Defic. Syndr. **4:**1004–1009.
19. **Hahn, B. H., M. A. Gonda, G. M. Shaw, M. Popovic, J. A. Hoxie, and R. C. Gallo.** 1985. Genomic diversity of the acquired immunodeficiency virus HTLV-III: different viruses exhibit greatest divergence in their envelope genes. Proc. Natl. Acad. Sci. USA **82:**4813–4817.
20. **Higgins, D. G., A. J. Bleasby, and R. Fuschs.** 1992. CLUSTAL V: improved software for multiple sequence alignment. Comput. Applic. Biosci. **8:**189–191.
21. **Ho, D., T. Moudgil, and M. Alam.** 1989. Quantitation of human immunodeficiency virus type 1 in the blood of infected persons. N. Engl. J. Med. **321:**1621–1625.
22. **Holmes, E. C., L. Q. Zhang, P. Simmonds, C. A. Ludham, and A. J. Leigh Brown.** 1992. Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. Proc. Natl. Acad. Sci. USA **89:**4835–4839.
23. **Holmes, E. C., L. Q. Zhang, P. Simmonds, A. Smith Rogers, and A. J. Leigh Brown.** 1993. Molecular investigation of human immunodeficiency virus (HIV) infection in a Baltimore surgeon. J. Infect. Dis. **169:**1411–1414.
24. **Hwang, S. S., T. J. Boyle, H. K. Lyrly, and B. R. Cullen.** 1991. Identification of the envelope V3 loop as a primary determinant of cell tropism in HIV-1. Science **253:**71–74.
25. **Kishino, H., and M. Hasegawa.** 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data and the branching order in Hominoidea. J. Mol. Evol. **4:**406–425.
26. **Koot, M., A. H. V. Vos, R. P. M. Keet, R. E. Y. de Goede, M. Wouter Dercksen, F. G. Terpstra, R. A. Coutinho, F. Miedema, and M. Tersmette.** 1992. HIV-1 biological phenotype in long term infected individuals evaluated with an MT-2 cocultivation assay. AIDS **6:**49–54.
27. **Korber, B., and G. Myers.** 1992. Signature pattern analysis: a method for assessing viral sequence relatedness. AIDS Res. Hum. Retroviruses **8:**1549–1560.
28. **Krivine, A., G. Firtion, L. Cao, C. Francon, R. Heurion, and P. Lebon.** 1992. HIV replication during the first week of life. Lancet **339:**1187–1189.
29. **Kusumi, K., B. Conway, S. Cunningham, A. Berson, C. Evans, A. K. N. Iversen, D. Colvin, M. V. Gallo, S. Coutre, E. G. Shaper, D. V. Faulkner, A. de Ronde, S. Volkman, C. Williams, M. S. Hirsch, and J. Mullins.** 1992. Human immunodeficiency virus type 1 envelope gene structure and diversity in vivo and after cocultivation in vitro. J. Virol. **66:**875–885.
30. **Milich, L., B. Margolin, and R. Swanstrom.** 1993. V3 loop of the human immunodeficiency virus type 1 *env* protein: interpreting sequence variability. J. Virol. **67:**5623–5634.
31. **Mulder-Kampinga, G. A., C. Kuiken, J. Dekker, H. J. Scherpbier, K. Boer, and J. Goudsmit.** 1993. Genomic human immunodeficiency virus type 1 RNA variation in mother and child following intra-uterine virus transmission. J. Gen. Virol. **74:**1747–1756.
32. **Mullis, K. B., and F. A. Faloona.** 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods Enzymol. **155:**335–350.
33. **Myers, G., B. Korber, J. A. Berzofsky, R. F. Smith, and G. N. Pavlakis.** 1991. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
34. **Nara, P. L., L. Smith, N. Dunlop, W. Hatch, M. Merges, D. Waters, J. Kelliher, R. C. Gallo, P. J. Fischinger, and J. Goudsmit.** 1990. Emergence of viruses resistant to neutralization by V3-specific antibodies in experimental human immunodeficiency virus type 1 IIIB infection of chimpanzees. J. Virol. **64:**3779–3791.
35. **Oram, J. D., R. G. Downing, M. Roff, N. Serwankambo, J. C. S. Clegg, A. S. R. Featherstone, and J. C. Booth.** 1991. Sequence analysis of the V3 loop regions of the env genes of Ugandan human immunodeficiency proviruses. AIDS Res. Hum. Retroviruses **7:**605–614.
36. **Ou, C.-Y., C. A. Ciesieleski, G. Myers, C. I. Bandea, C.-C. Luo, B. T. M. Korber, J. I. Mullins, G. Schochetman, R. L. Berkelman, A. N. Economou, J. J. Witte, L. J. Furman, G. A. Satten, K. A. MacInnes, J. W. Curran, H. W. Jaffe, the Laboratory Investigation Group, and the Epidemiologic Investigation Group.** 1992. Molecular epidemiology of HIV transmission in a dental practice. Science **256:**1165–1171.
37. **Rossi, P., V. Moschese, P. A. Brolinden, C. Fundaro, I. Quinti, A. Plebani, C. Giaquinto, P. A. Tovo, K. Ljunggren, J. Rosen, H. Wigzell, M. Jondal, and B. Wahren.** 1989. Presence of maternal antibodies to human immunodeficiency virus 1 envelope glycoprotein gp120 epitopes correlates with the uninfected status of children born to seropositive mothers. Proc. Natl. Acad. Sci. USA **86:**8055–8058.
38. **Ryder, R. W., W. Nsa, S. E. Hassig, F. Behets, M. Rayfield, B. Ekungda, A. M. Nelson, U. Mulenda, H. Francis, K. Mwanda-Galirwa, F. Davachi, M. Rogers, N. Nzilambi, A. Greenberg, J. Mann, T. C. Quinn, P. Piot, and J. W. Curran.** 1989. Perinatal transmission of the human immunodeficiency virus type 1 to infants of seropositive women in Zaire. N. Engl. J. Med. **320:**1637–1642.
39. **Saag, M. S., B. H. Hahn, J. Gibbons, Y. Li, E. S. Parks, W. P. Parks, and G. M. Shaw.** 1988. Extensive variation of the human immunodeficiency virus type-1 *in vivo*. Nature (London) **334:**440–444.
40. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
41. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
42. **Scarlatti, G., T. Leitner, E. Halapi, J. Wahlberg, P. Marchisio, M. A. Clerici-Schoeller, H. Wigzell, E. M. Fenyo, J. Albert, M. Uhlen, and P. Rossi.** 1993. Comparison of variable region 3 sequences of human immunodeficiency virus type 1 from infected children with the RNA and DNA sequences of the virus populations of their mothers. Proc. Natl. Acad. Sci. USA **90:**1721–1725.
43. **Shioda, T., J. A. Levy, and C. Cheng-Mayer.** 1991. Macrophage and T cell-line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene. Nature (London) **349:**167–169.
44. **Simmonds, J., P. Balfe, C. A. Ludlam, J. O. Bishop, and A. J. Leigh Brown.** 1990. Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1. J. Virol. **64:**5840–5850.
45. **Simmonds, P., L. Q. Zhang, F. McOmish, P. Balfe, C. A. Ludlam, and A. J. Leigh Brown.** 1991. Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 *env* sequences in plasma viral and lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. J. Virol. **65:**6266–6276.
46. **Soeiro, R., A. Rubinstein, W. K. Rashbaum, and W. D. Lyman.** 1992. Maternofoetal transmission of AIDS: frequency of human immunodeficiency virus type 1 nucleic acids sequences in human foetal DNA. J. Infect. Dis. **166:**699–703.
47. **Tersmette, M., R. E. Y. de Goede, M. Al, I. N. Winkel, R. A. Gruters, H. T. M. Cuypers, J. G. Huisman, and F. Miedema.** 1988. Differential syncytium-inducing capacity of human immunodeficiency virus isolates: frequent detection of syncytium-inducing isolates in patients with acquired immunodeficiency syndrome (AIDS) and AIDS-related complex. J. Virol. **62:**2026–2032.

48. **Tersmette, M., R. A. Gruters, F. de Wolf, R. E. Y. de Goede, J. M. A. Lange, P. T. A. Schellekens, J. Goudsmit, H. G. Huisman, and F. Miedema.** 1989. Evidence for a role of virulent human immunodeficiency virus (HIV) strains in pathogenesis of acquired immunodeficiency syndrome obtained from a panel of sequential HIV isolates. J. Virol. **63:**2118–2125.
49. **Tricoire, J.** Personal communication.
50. **Wolinsky, S. M., C. M. Wike, B. T. M. Korber, C. Hutto, W. P. Parks, L. L. Rosenblum, K. Kunstman, M. R. Furtado, and J. L. Munoz.** 1992. Selective transmission of human immunodeficiency virus type-1 variant from mothers to infants. Science **255:**1134–1137.
51. **Zhang, L. Q., P. MacKenzie, A. Cleland, E. C. Holmes, A. J. Leigh Brown, and P. Simmonds.** 1993. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J. Virol. **67:**3345–3356.
52. **Zwart, G., H. Langedijk, L. Van der Hoek, J. J. de Jong, T. F. Wolfs, C. Ramautarsing, M. Bakker, A. de Ronde, and J. Goudsmit.** 1990. Immunodominance and antigenic variation of the principal neutralisation domain of HIV-1. Virology **181:**481–489.