JOURNAL OF VIROLOGY, Feb. 1997, p. 1292–1300 Vol. 71, No. 2
0022-538X/97/$04.00+0

Mother-to-Infant Transmission of Human Immunodeficiency Virus Type 1 Involving Five Envelope Sequence Subtypes

CHRISTOPHER H. CONTAG,[1] ANNEKA EHRNST,[2] JENNIFER DUDA,[3] ANN-BRITT BOHLIN,[4] SUSANNE LINDGREN,[5] GERALD H. LEARN,[6] AND JAMES I. MULLINS[6,7]*

Department of Pediatrics[1] and Department of Medicine,[3] Stanford University School of Medicine, Stanford, California; Department of Immunology, Microbiology, Pathology and Infectious Diseases,[2] Department of Pediatrics,[4] and Department of Obstetrics and Gynecology,[5] Huddinge University Hospital, Karolinska Institutet, Stockholm, Sweden; and Department of Microbiology[6] and Departments of Medicine and Laboratory Medicine,[7] University of Washington, Seattle, Washington

Received 29 May 1996/Accepted 17 October 1996

Genetic analysis of human immunodeficiency virus type 1 (HIV-1) from cases of mother-to-infant transmission were analyzed in an effort to provide insights into the viral selection that may occur during transmission, as well as the timing and source of transmitted viruses. HIV-1 *env* genes obtained from seven mothers and their perinatally infected infants in Sweden were studied. Five envelope sequence clades (A to E) were found to be represented. We used a heteroduplex tracking assay (HTA) to assess the genetic relatedness between early viral isolates from the infants and serial maternal virus populations taken during pregnancy and at delivery. HTA findings were used to select for DNA sequence analysis maternal virus populations that were either closely or more distantly related to the infant virus. In each case, nucleotide sequence analysis confirmed the genetic relationships inferred by the HTA. Only maternal peripheral blood was sampled, and large sets of maternal specimens throughout pregnancy were generally not available. However, no consistent correlation was found to support the hypothesis that infant viruses should match blood-derived maternal virus genotypes found early in pregnancy if infants were found to be infected at birth or, conversely, that infant viruses should match blood-derived maternal virus genotypes found at delivery if infants were found to be infected only some time later.

Transmission from mothers to infants is the most common source of human immunodeficiency virus type 1 (HIV-1) infection of children (2), estimated to occur in from 15% (30) to 43% (6) of births, the latter including transmission through breast-feeding. Transmission rarely occurs during the first half of pregnancy (4, 12), although recent studies suggest that in utero transmission corresponds to 30 to 40% of all mother-to-infant transmissions (33). The proportion of infants that can be diagnosed as infected rises from around 40% during the first week of life to around 90% by the second week (11). The major risk of maternal HIV-1 transmission thus appears to occur late during pregnancy or at delivery (12, 29). It has been possible to lower this frequency by two-thirds as a result of zidovudine therapy during the latter half of pregnancy, during delivery, and to the newborn (3, 5, 18).

The first HIV sequences detectable in infants have often been reported to be minor variants compared to the maternal viral DNA (1, 23, 38) or RNA sampled (26), suggesting selective transmission or selective growth following transmission. However, due to the sometimes high turnover rate in infected individuals (9), potential for compartmentalization of virus variant populations in different fluids or tissues involved in transmission (14), and the simple lack of adequate samples for analysis, it may be difficult to pinpoint the virus populations from which transmission occurred.

Here we report an analysis of HIV-1 strains in seven cases of mother-to-child transmission in Sweden. We used the heteroduplex tracking assay (HTA) (10), a sensitive means to discriminate between virus populations (9), to identify maternal virus populations with envelope gene sequences most closely related to those of initial infant virus isolates. HTA detects short deletions and insertions, as well as base sequence mismatches, generating diversity above 1 to 2%. It has proven useful for envelope sequence subtype determination (10) as well as for tracking virus populations and individual variants over time between (7) and within (9) infected individuals.

To further understand the nature of the variants detected in infants, their relationship to transmitted variants, and the reliability of HTA in predicting these relationships, DNA sequences of comparable regions of viral DNA taken from the infants and from selected maternal viruses were determined. We found that HTA did accurately discriminate among these closely related virus populations. Surprisingly, five of the nine known *env* sequence clades (28) were represented in this group of seven HIV-1 infected infants from Sweden. Furthermore, in contrast to some earlier reports (38), no shared genetic features within the V3 to V5 region of *env* were identified among the transmitted viruses. Lack of available, complete specimen sets from the pregnant women, however, may have prevented a rigorous evaluation of the timing of virus transmission.

MATERIALS AND METHODS

Patients and samples. Pregnant women identified as or who previously knew they were infected with HIV-1 participated by informed consent in a study of HIV during pregnancy approved by the ethics committee of the Karolinska Institutet. Breast-feeding was discouraged for all and was denied by all of the women described in this study. HIV isolation from the mothers was conducted during the three trimesters, a few days after delivery (referred to as delivery in this study), and 6 months after parturition. For inclusion in the present study, the women must have had (i) at least one positive virus isolation from peripheral blood mononuclear cells (PBMC) or plasma during pregnancy and (ii) an infected child with at least one positive virus isolation. These children were born between June 1989 and August 1992 and were followed by virus isolation at birth

* Corresponding author. Mailing address: Department of Microbiology, University of Washington, Seattle, WA 98195-7740. Phone: (206) 616-1851. Fax: (206) 616-1575. E-mail: jmullins@u.washington.edu.

and at 3, 6, 12, and 18 months of age unless circumstances called for an extra sample.

Virus isolation and DNA purification. EDTA-blood collected was received by the laboratory within 24 h from all participants, and 4 ml was separated by Ficoll-Isopaque gradient centrifugation. Undiluted plasma supernatant was subjected to ultracentrifugation at 55,000 × *g*, and virus was cultured from pelleted virus and washed PBMC as previously described (13). Cultures were maintained for 6 weeks, with changes of half of the medium twice weekly and testing for the presence of p24 antigen by enzyme-linked immunosorbent assay every week. Supernatants of cultures reaching an absorbency of >2.0 were stored frozen at −70°C and later recultured to produce specimens for this study. DNA was prepared from uncultured PBMC and cultured cells by standard methods including proteinase K digestion and phenol-chloroform extraction.

PCR. Seminested, diagnostic PCR for HIV-1 *gag* and *pol* genes was performed on uncultured patient PBMC in Sweden as described previously (34). *env* regions encoding V1 to V5 or V3 to V5 from both mother and infant blood samples and derived virus isolates were amplified by nested PCR and analyzed by the HTA (9, 10). First-round primers were ED3 and ED14. Second-round reactions used 2 μl of first-round product as the template and either primers ES7 and ES8 for analysis of the V3 to V5 regions of *env* or ED5 and ES8 for analysis of the V1 to V5 region.

HTA. HTA analyses containing a template copy number of 50 or more is desired for maximum sensitivity of detection of viral variants (9). To determine if threshold levels of viral template were present, 10 ng (and, if necessary, 30, 100, and 300 ng) of PBMC DNA was examined in a nested PCR. All virus isolates had a minimum of 30 templates used for the HTA. No quantitation was performed when PBMC DNA was the template, since only first-round PCR products were available for HTA. Hence, the product derived from PCR of 1 μg or 1.5×10^5 cell equivalents of DNA was used in each HTA of PBMC DNA.

HTAs were conducted essentially as previously described (9, 10). Concentrations of PCR products were estimated by ethidium bromide staining of agarose gels. Infant probe DNA was generated in labeling reactions of second-round PCR products involving five cycles of PCR in a 50-μl final volume containing 3.3 mM [^{32}P]TTP (10 mCi at a specific activity of 3,000 Ci/mmol), 33 mM each dATP, dCTP, dGTP, and TTP, and standard PCR buffer containing 7.5 pmol of each second-round primer and 1.25 U of *Taq* DNA polymerase. Labeled probe concentrations were estimated by assuming a doubling of the template concentration during each of the five cycles of amplification. After electrophoresis, gels were stained for 5 min with ethidium bromide, photographed, dried onto 3MM paper (Whatman) on a gel dryer for 1 h at 80°C, exposed, and scanned with a Molecular Dynamics (Sunnyvale, Calif.) PhosphorImager. Images were then transferred to disks and printed from a Macintosh computer, using Adobe (Mountain View, Calif.) Photoshop and Deneba (Miami, Fla.) Canvas software.

Cloning and DNA sequencing. *env* fragments of selected maternal samples for all cases and for infant samples from cases 2 and 5 were obtained for cloning by a second nested round of PCR amplification using the first-round PCR products described above as templates and primers ES7X (5′-actgagtcctcgag**ctgttaaatggc agtctagc**, containing an *Xho*I restriction endonuclease site [underlined] and an HIV-1-specific sequence [in boldface], corresponding to positions 7001 to 7020 of the HIV-1 HXB2 genome [GenBank accession no. K03455]) and ES8S (5′-g agtcgagctc**cacttctccaattgtccctca**, containing a *Sac*I site [underlined] and an HIV-1-specific sequence [in boldface] corresponding to positions 7667 to 7647 of HIV-1 HXB2). Sodium dodecyl sulfate and EDTA were added to final concentrations of 0.2% and 5 mM, respectively, to stop the PCRs, and then nucleic acids were precipitated with ethanol and 0.3 M sodium acetate.

PCR products were redissolved and digested with *Xho*I and *Sac*I, and the desired products were purified from agarose gels (using GeneClean [Bio 101, La Jolla, Calif.]), ligated with pUC19 digested with *Sal*I and *Sac*I, precipitated in the presence of linear polyacrylamide (19), and then used to transform the *Escherichia coli* J5, using a Bio-Rad (Richmond, Calif.) electroporator. Plasmids with an insert of the correct size were assessed by agarose gel electrophoresis, purified, and used as templates for DNA sequencing.

Viral DNA sequences from infants in cases 1, 3, 4, 6, and 7 were determined directly from PCR products generated with primers ES7 and ES8, using the 0.7-kb products purified from agarose gels as described above. All DNA sequences were determined by using dye-labeled universal and reverse M13 primers and the cycle sequencing method on an ABI model 370A automated sequencer.

Analysis of sequence data. Overlapping regions were identified and sequences were edited by using the Intelligenetics (Mountain View, Calif.) program GEL. Sequences were aligned with examples of the known envelope sequence clades of HIV-1 obtained from the Los Alamos database (28) by using the program GENALIGN (Intelligenetics), and the alignments were refined by eye, using MASE (15). Phylogenetic relationships were reconstructed by using PHYLIP software (17) (DNADIST to estimate the matrix of pairwise sequence distances [21] followed by NEIGHBOR to construct phylogenetic trees using the neighbor-joining method [32]) through Genetic Data Environment software (35). The degree of confidence in phylogenetic branching was assessed by using 1,000 bootstrap resamplings (16).

Nucleotide sequence accession numbers. GenBank accession numbers for the sequences reported in this study are U56263 to U56335.

RESULTS

Seven mother-child pairs participated in this study. Only two women were likely to have become infected in Sweden, while four were probably infected in Africa and one was probably infected in Southeast Asia (Table 1). All together, 67 HIV-1 virus isolates were obtained from these mother-infant pairs (Table 1). In experiments to be detailed below, nucleotide sequences of a segment of the *env* gene from maternal and infant viruses were determined. When aligned with each other and with representative viruses from the available database of HIV-1 *env* genes (28) and subjected to phylogenetic analysis using the neighbor-joining method (32), the mother-infant pairs clustered together as expected (Fig. 1). The seven mother-infant pairs were found to represent five of the nine known (28) envelope sequence subtypes of the main group of HIV-1 (Table 1 and Fig. 1).

Though obtained from Sweden, where *env* clade B viruses are expected to predominate, viruses from only mother-infant cases 2 and 6 had *env* genes that belonged to this clade. Viral sequences from cases 5 and 7 grouped with *env* clade C, and cases 1, 3, and 4 grouped with clades E, D, and A, respectively. While striking, this diversity was not unexpected, since each of the *env* sequence subtypes found was known to be prevalent in the country or geographic region (28) in which each mother was thought to have been infected (Table 1).

We sought to address the hypothesis that the timing of transmission can be inferred by genetic linkage of the first virus to appear in the infant with viruses in the mother during her pregnancy. In an effort to expedite and extend the analysis to more viral molecules than are traditionally examined by DNA sequencing, we used the HTA followed by targeted DNA sequencing. For the HTA, PCR-amplified, radiolabeled infant virus gene fragments were used as the probe against a driver corresponding to a 100-fold excess of the corresponding fragments amplified from maternal virus populations. We focused most of our analysis on PBMC- and plasma-derived virus isolates. PBMC-derived virus isolated likely correspond to the actively replicating fraction of virus in the PBMC at the time of sampling (9). Similarly, plasma viral sequences have been found to be produced and cleared more rapidly than the bulk of cell-associated proviral DNA (36) and hence may be predicted to be both more likely to be infectious than the majority of viruses integrated as proviruses in the PBMC as well as being most representative of actively replicating HIV strains at any given time point.

Transmission involving clade B sequences. The child in case 6 was HIV infected, as evidenced by PCR-positive PBMC and virus isolation from both cells and plasma, at the first time point examined, 2 days after birth (Table 1), consistent with the assumption of infection in utero. We examined *env* gene fragments encompassing the 0.7-kb C2 to V5 region in all of the following transmission cases. When case 6 was examined, however, little distinction could be made between the infant virus probe and the maternal quasispecies throughout pregnancy except at the time of delivery (Fig. 2). This is illustrated by the finding of heteroduplex bands migrating with the mobility of the reannealed probe, that is, with homoduplexes. In contrast, when 1.1-kb V1 to V5 fragments were examined, the number of lanes in which the heteroduplex mixtures comigrated almost fully with the probe homoduplexes was reduced from seven to four (Fig. 2). This greater sensitivity likely reflects a substantial number of point mutations as well as probable length changes in the first and/or second variable region encompassed within the larger fragment.

The maternal virus sequences most closely related to that

TABLE 1. Demographics of maternal HIV infections, envelope sequence subtypes, diagnostic (*gag* and *pol*) PCR from PBMC, and virus isolation from PBMC and plasma of mothers and their children[a]

Case	Maternal infection		*env* sequence subtype	Sample	Assay	Maternal samples				Infant samples								
	Probable route of transmission	Probable origin				1st trimester	2nd trimester	3rd trimester	Delivery	Earliest time sampled (days after birth)[b]	Postpartum	1.5 mo	3 mo	6 mo	9 mo	12 mo	18 mo	24 mo
1	Sexual	Southeast Asia	E	PBMC	PCR					3				+		+		
				PBMC	VI						−			C		+		+
				Plasma	VI				[+]		−			+		+		−
2	Intravenous drug use	Sweden	B	PBMC	PCR	+	+	+	+	3	−			−	−	+	+	+
				PBMC	VI	+	−	+	+		−			−	+	−		−
				Plasma	VI	−	+	−	+		−			−	+	−		−
3	Sexual	Uganda	D	PBMC	PCR					1	−		−	−	+	+	+	
				PBMC	VI		+	+	+		−		−	−	−	+	+	
				Plasma	VI		+	+	[+]		−		−	−	+	+	+	
4	Sexual	Uganda	A	PBMC	PCR		+	+	−	4	+		+	+			+	
				PBMC	VI		+	+	+		+		+	+		+	+	
				Plasma	VI		[+]	−	−		+		+	+		+	+	
5	Sexual	Kenya	C	PBMC	PCR					4	+		+	+				
				PBMC	VI				[+]		−		+	+		+		
				Plasma	VI				+		−		+	+		−		
6	Sexual	Sweden	B	PBMC	PCR					2	+	+	+	+		+		
				PBMC	VI	+	+	+	+		+	+	+	−		+		
				Plasma	VI	+	+	[+]	+		+	+	+	−		+		
7	Sexual	Zimbabwe	C	PBMC	PCR					3	+	+						
				PBMC	VI			−	+		+	+						
				Plasma	VI			+	[+]		+	+						

[a] +, positive result of test; −, negative result of test; VI, virus isolation; C, contaminated culture, no result. The maternal variant most closely linked to the variant present in the child is boxed.
[b] Time of postpartum sample.

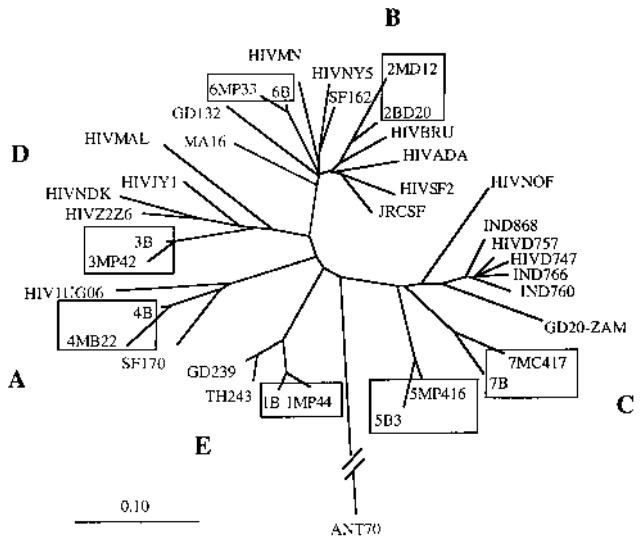

FIG. 1. Phylogenetic reconstruction demonstrating relationships of maternal and infant viruses to HIV-1 reference sequences from six different *env* sequence clades. Two sequences from each of seven mother-infant transmission cases (boxed) were compared to selected HIV-1 reference sequences from five envelope sequence subtypes, A to E, from the HIV-1 main group as well as the group O virus ANT 70 (28). The mother and infant sequences are designated as follows. The first number indicates the case number; next, a letter indicates maternal (M) or baby (B) origin. If more than one sequence was obtained, the next position is a letter indicating plasma (P)- or PBMC (C)-derived virus isolates, or an isolate taken directly from the patient's PBMC (D). The next position is a number indicating when the sample was obtained from the mother (1, 2, or 3, first, second, or third trimester; 4, delivery), and the final number indicates the clone number. A bar representing 10% sequence divergence is shown for this and each of the other phylogenetic reconstructions in subsequent figures.

present in the infant in case 6 were found to be abundant in the PBMC-derived isolates from the second trimester through delivery, along with the plasma isolate from the third trimester. Thus, DNA sequences were determined for the two maternal isolates from the third trimester to both confirm inferences from the HTA data and determine whether DNA sequence analysis would distinguish genetic relationships not readily determined by HTA.

DNA sequences of mother- and infant-derived viruses from case 6 formed a closely related monophyletic group within *env* clade B (Fig. 1 and 2), consistent with the homogeneity inferred by HTA. The ranges of variation within the maternal virus populations were 0.3 to 1.8% (plasma isolate, $n = 5$) and 2.2 to 3.4% (PBMC isolate, $n = 4$). Compared to the infant viral sequence, the average divergences of the maternal viral sequences were 2.8 and 2.0% for PBMC and plasma isolates, respectively. These ranges are at or near the lower limits of discrimination that have been reported for the HTA (10). Despite the limited variation within this set of sequences, phyletic distinctions were observed for both the PBMC- and plasma-derived sequences, although upon repetitive bootstrap resampling, the infant virus did not group preferentially with either (Fig. 2). While DNA sequence analysis did not reveal consistent relationships that were not evident by HTA, viruses sampled from this plasma isolate were on average slightly more closely related than PBMC-derived virus from the same time point.

Viruses from case 2 also belonged to a monophyletic group within *env* clade B; however, in contrast to case 6, each maternal virus isolate and uncultured PBMC sample demonstrated significant diversity as well as divergence from the infant probe (Fig. 2). Nearly all of the maternal-infant heteroduplexes migrated more slowly than homoduplexes, indicating a relatively distant genetic relationship to the infant viruses. This child was virus isolation and PCR negative at birth and when next tested at age 6 months (Table 2). PCR remained negative but virus isolation was positive from both plasma and cells at 9 months but only infrequently thereafter. PCR was first positive at 10.5 months (data not shown) and thereafter. However, only the age 24-month sample was positive with the *env* PCR primers used in the present study. Thus, in this child, the initial level of HIV was low and/or potentially divergent from the primers used here and appeared unusually late. Nonetheless, breast-feeding was denied by this mother and all mothers evaluated in this study, and hence postpartum transmission appears unlikely.

The case 2 maternal PBMC sample from the first trimester was found by HTA to have some sequences closely related to the infant probe. Sequences were therefore determined from PBMC taken from the mother during the first and second trimesters and from the infant at 24 months of age (Fig. 2). Sequence variation among five clones derived from the infant was minimal (0.15 to 1.4%), whereas the mean sequence distance between mother and infant sequences was 6.6% (range, 4.7 to 8.9%), only slightly greater than the distances between sequences within a given maternal specimen (mean, 5.8%; range, 1.8 to 8.3%). No reproducible phyletic distinctions were drawn between the two sets of maternal sequences, and the infant sequences were equidistant from the two maternal sequence sets. Thus, a maternal variant closely matching the infant sequence was not identified by HTA or DNA sequence analysis. Since we have only a limited sample of the maternal virus diversity, it is possible that we missed maternal sequences more closely related to infant sequences. The closest sequence from this pair to those found in the AIDS sequence database (28) was 6.7% divergent (a typical value for most closely related epidemiologically unlinked sequences in the database); it is therefore highly unlikely that any of the divergent sequences represent contamination from a known source (22, 24). Rather, if, as expected, transmission occurred at or before birth, then the 2 years of virus evolution in the child before sampling could explain the large degree of divergence between maternal and infant virus populations. That the infant virus population appeared relatively homogeneous might suggest that the child was infected for less than the 2 years between birth and sampling. However, a low viral template input in the PCR used to generate the infant probe could also have accounted for the lack of evident diversity (25), with *Taq* polymerase error then potentially accounting for some of the distinctions between sequences within the two infant virus clusters.

Transmission involving other HIV-1 envelope sequence clades. Sequences from case 3 belonged to *env* clade D (Fig. 1). The infant in this case was virus negative by PCR and by virus isolation at birth and at 3 and 6 months of age. Viral sequence from the first infant isolate obtained at 9 months of age was used to probe the maternal isolates, and only the plasma isolate from delivery formed heteroduplexes that migrated close to the homoduplex position (Fig. 3). The plasma and PBMC isolates generated at delivery were therefore selected for sequence analysis. As predicted by HTA, phylogenetic analysis had the infant sequence grouped more closely with the plasma than with the PBMC isolate sequences (in 97% of bootstrap repetitions). The average sequence diversities between the infant sequences, derived from the consensus sequence of bulk PCR products from a viral isolate, were 3.3% compared to the maternal PBMC-derived isolate and 2.3% compared to the

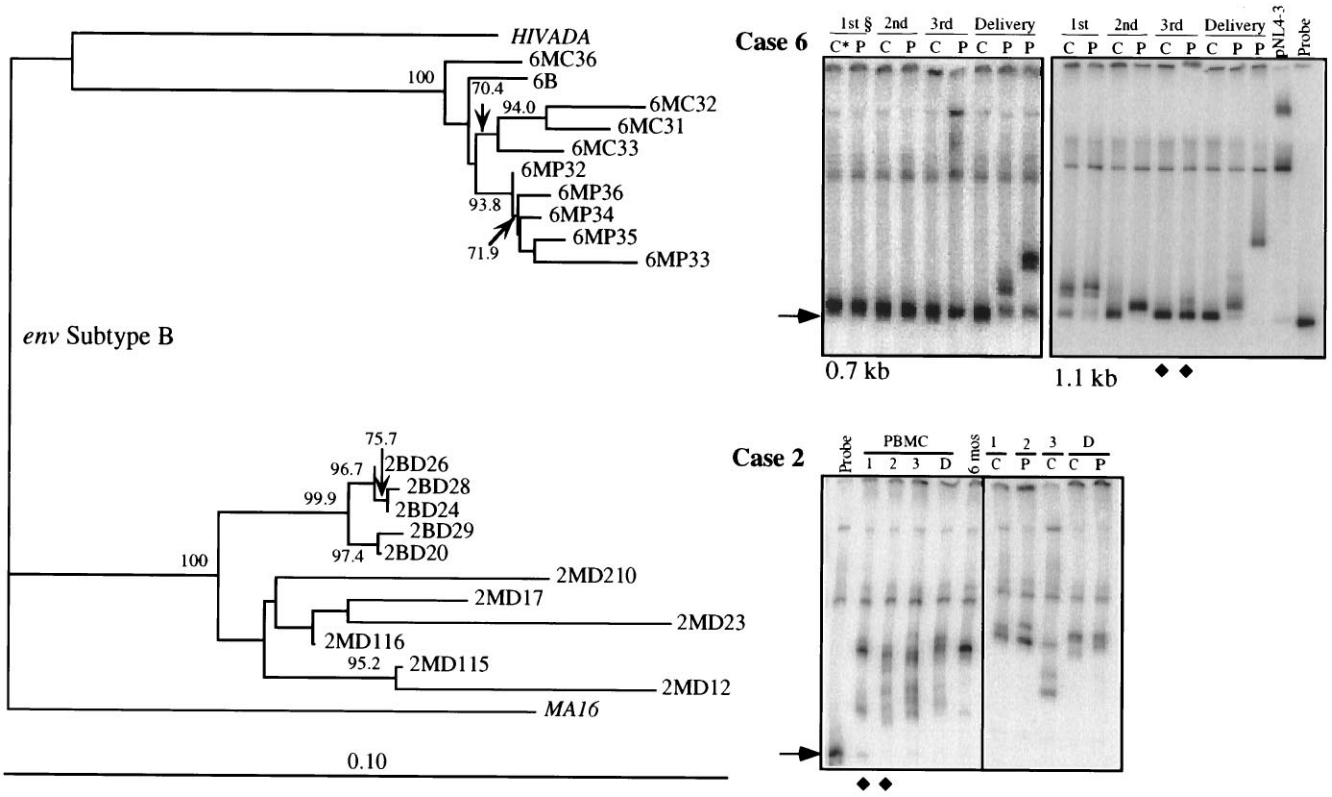

FIG. 2. Heteroduplex tracking and phylogenetic analyses of maternal and infant variants for cases in *env* clade B, as well as two reference sequences from *env* sequence subtype B. See the legend to Fig. 1 for explanations of the nomenclature and phylogenetic analysis. The numbers at branch points are the percentages of 1,000 bootstrap resamplings in which isolates in the group to the right were clustered together (all clusters with frequencies >70% are shown). The probe lane in each panel contains the labeled infant viral probe; the arrow indicates the position of homoduplex migration. The pNL4-3 lane in the upper panel corresponds to the comparison of the infant probe sequence to an epidemiologically unlinked clade B *env* gene used as positive control in the PCR experiments reported in this study. HTA for case 6 includes analysis of 0.7-kb (V3 to V5) and 1.1-kb (V1 to V5) fragments. HTA using the 0.7-kb fragment is shown for case 2. The infant probe in case 6 was an isolate obtained at birth, and the infant probe for case 2 was from PBMC taken at 24 months of age. Outgroup sequences taken from the HIV sequence database are shown in italics. C*, PBMC; P, plasma; §, trimester of pregnancy; ◆, samples chosen for sequencing.

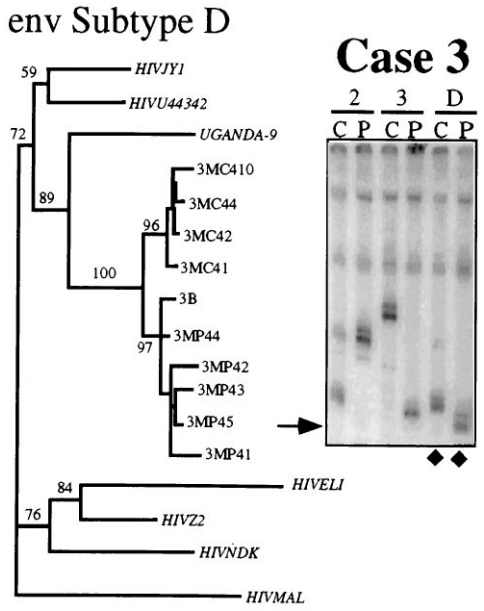

FIG. 3. Heteroduplex tracking and phylogenetic analyses of case 3 variants from *env* clade D. See the legends to Fig. 1 and 2 for explanations of the nomenclature, markers, and phylogenetic analysis. The infant was positive at 8 months of age, and the plasma virus isolate from this time was used as the probe.

plasma-derived isolate. Thus, the distinctions apparent among these sequences by HTA were also apparent in phylogenetic reconstructions of DNA sequence relationships.

The infant in case 4, involving transmission of a clade A virus (Fig. 1), was virus positive by all criteria 4 days after birth (Table 2), possibly indicating in utero infection. Similar to what was found for case 2, none of the maternal virus populations from case 4 matched closely with the PBMC-derived virus from 3 months of age used as the infant probe in HTA experiments (Fig. 4). Maternal viruses within the PBMC from the distinctly different second and third trimesters were selected for sequence analysis. Alignment of these sequences indicated that nucleotide substitutions and length changes in V4 largely accounted for the differences observed by HTA; length variations were not found in the other regions analyzed (data not shown). The infant sequence had a unique 9-nucleotide deletion not found in the maternal sequences; moreover, relative to the maternal second trimester sequences, the infant variant had a second, 12-nucleotide deletion and an 18-nucleotide deletion compared to the maternal variant population at the third trimester. The infant sequence was 3.9 to 5.9% divergent from the second-trimester clones and 4.3 to 5.7% divergent from the third-trimester clones. Thus, the size of the deletion correlated more closely with the HTA results than did the minor differences in nucleotide sequence divergence, but neither technique could identify a close relative to that infecting the infant.

Case 7 involved infection and transmission with HIV of

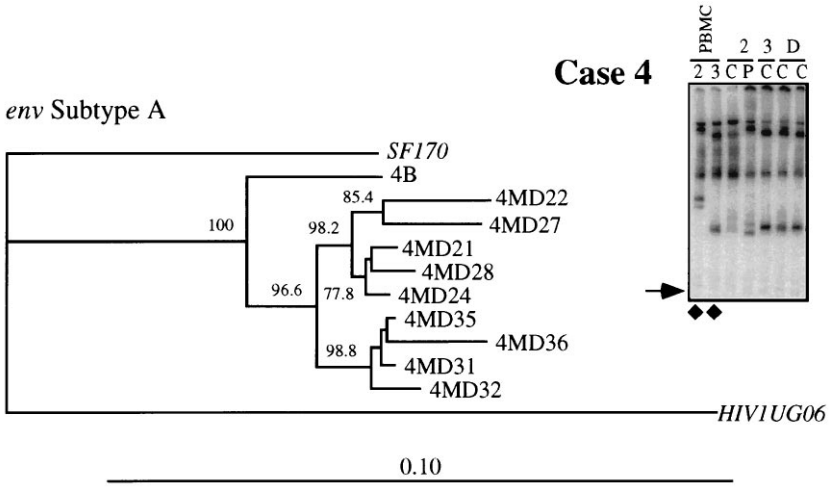

FIG. 4. Heteroduplex tracking and phylogenetic analyses of case 4 sequences from *env* clade A. See the legends to Fig. 1 and 2 for explanations of the nomenclature, markers, and phylogenetic analysis. The infant was HIV-1 positive by culture at birth, and the probe was from PBMC at 3 months of age.

envelope sequence clade C. HIV was detected in the infant by virus isolation from PBMC and PCR 3 days after birth, and comparison of this isolate to maternal isolates from the third trimester and delivery indicated that the infant sequence was closely related only to the plasma isolate from delivery (Fig. 5). Viral DNA from delivery cell- and plasma-derived isolates were sequenced, and as expected from HTA, the infant variant grouped with the sequences from this plasma virus population (Fig. 5). The infant-derived sequence diverged by 2.7 to 4.5% from six maternal plasma sequences and by 8.9 to 11.3% from six PBMC-derived sequences. A seventh PBMC-derived sequence (7MCD5) grouped with the plasma and was 2.9% divergent from the infant sequence.

Case 5 also involved clade C. This infant was first positive by PCR 4 days after birth but first positive by virus isolation at 3 months of age. In contrast to the other infant virus isolates examined, HTA analysis of the PBMC-derived virus isolate from 3 months of age revealed heterogeneity, individual variants of which were then cloned. Comparison of six infant-derived clones to the two available maternal isolates by HTA (Fig. 5 and data not shown) revealed that each was more closely related to the maternal PBMC-derived isolate at delivery than the plasma-derived isolate from the same time, with the exception of clone 5B.4, which was equidistant from the maternal isolates. Genetic relationships among the six infant clones were determined in a pairwise fashion using HTA, which indicated that three (5B.2, 5B.3, and 5B.8) were closely related to each other and distinct from the other three (5B.4, 5B.7, and 5B.9) (data not shown).

The genetic differences between infant and maternal sequences observed by HTA for case 5 were also evident in phylogenetic analyses of these sequences (Fig. 5), with a mean distance between all infant and maternal PBMC-derived virus sequences of 3.9%, compared with 5.7% between infant and maternal plasma-derived sequences. Sequence alignments also revealed significant length variability in the V4 and V5 regions that accounted for much of the reduced mobility of the mother-infant heteroduplexes (data not shown).

Viruses from case 1 belonged to *env* clade E (Fig. 1); this infant was negative by virus isolation at birth and positive by PCR and virus isolation from plasma starting at the next available specimen at 6 months. Only a single maternal sample was available and analyzed for this case. The virus sequence from the maternal plasma-derived isolate taken from the time of delivery was distinct but related to the infant virus sequence by both HTA and DNA sequence analysis (3.86% sequence divergence).

Patterns of potential N-linked glycosylation sites. Loss of the N-linked glycosylation consensus sequence (N-X-T/S) overlapping the first cysteine of the V3 loop of gp120 (usually NCT) has been correlated as a rare variant transmitted to infants in some studies of *env* clade B *env* sequences (1, 37, 38). Despite the prevalence of the potential N-linked glycosylation site at this position (amino acid 259) in clade B viruses, the infant and all but one maternal sequence from case 6 lacked this site but had another site six amino acid residues upstream (underlined in Fig. 6A).

Infant and maternal sequences from case 7 also lacked the glycosylation site at position 259, with an apparent shift to the upstream position found in the clade B viruses in both infant and maternal plasma-derived sequences. However, as shown in Fig. 6B, the glycosylation pattern of case 7 sequences is common among other *env* clade C viruses; therefore, the absence of an N-linked glycosylation site at position 253 appears characteristic of this clade and not of transmission in case 7. Thus, for all seven transmission cases under study, the pattern of N-linked glycosylation sites near the V3 loop reflected either the clade or that which was present in maternal sequences and did not seem to correlate with mother-to-infant transmission. A transmission-linked pattern of N-linked glycosylation sites at other locations in the *env* region analyzed was also not apparent in this study.

DISCUSSION

In the seven transmission cases evaluated, viral genetic variation found within the PBMC and virus isolates from the plasma and PBMC of HIV-1 infected women was usually sufficient to distinguish viral populations throughout pregnancy by the HTA technique. Moreover, the genetic relationships between infant-derived and mother-derived HIV-1 sequences predicted by HTA were substantiated by analysis of DNA sequences. Therefore, we have validated the use of the HTA as a rapid assay for distinguishing viral genomes present in serial samples taken at short time intervals, as well as for identifying viruses that are epidemiologically linked in cases of mother-

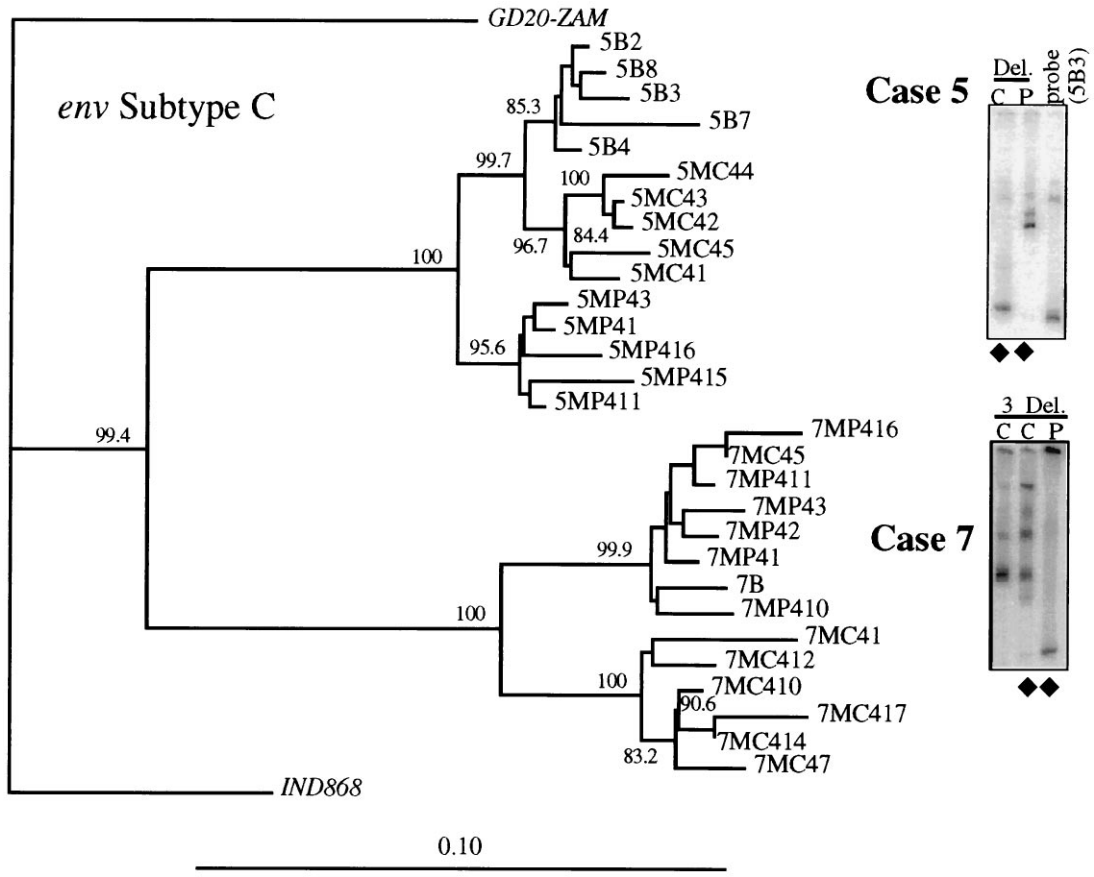

FIG. 5. Heteroduplex tracking and phylogenetic analyses of case 5 and 7 sequences from *env* clade C. See the legends to Fig. 1 and 2 for explanations of the nomenclature, markers, and phylogenetic analysis. Infants in cases 5 and 7 were both HIV 1 positive by PCR at birth, and the probes were PBMC isolates from birth for case 7 and clone 5B3 derived from the PBMC virus isolate from 3 months of age for case 5.

to-infant transmission. The use of this assay for establishing more temporally distant links between viruses transmitted by parenteral or sexual contact has also been documented (7).

In a recent large study (11), the frequency of detection of virus by PCR was found not to change during the first week of life (staying level at 40%) but rose to 90% by the second week of life in infants later shown to be infected with HIV. Hence, infants that are PCR positive during the first several days of life (as in cases 4, 6, and 7 here) are likely to have been infected in utero. The data in favor of most HIV transmissions occurring during the latter part of pregnancy are also substantiated by the fact that only 2% of fetuses from the second trimester have been found to be HIV positive (4) and by the reduction of maternal HIV-1 transmission by zidovudine treatment during the latter part of pregnancy (3, 5, 18). Nevertheless, it cannot be excluded in individual cases that HIV genomes can be detected in blood during the first week after infection. The infant in case 5 studied here was PCR positive at 4 days of life but not virus isolation positive until next tested at 3 months. This result may therefore correspond to early detection of virus transmitted during delivery or to transmission close to but prior to delivery.

Recent data suggest that virus in plasma and isolates derived from PBMC are distinguishable from the bulk of the proviruses integrated in PBMC in that the former correspond to younger, newly emerging and actively replicating viruses, whereas the PBMC proviral DNA turns over more slowly and in part represents an archeologic record of previously replicating variant populations (9, 36). We examined these populations in a test of the hypothesis that infants positive at birth would have been infected in utero whereas infants found to be infected later would more likely have been infected during delivery. Our results, however, fail to lend support to this hypothesis. In the three cases with a late diagnosis of HIV infection, no maternal virus population (case 2) or maternal virus populations from the second trimester (case 1) or delivery (case 3) were, based on the abundance of sequences related to that found in the infant, implicated as the more likely sources of transmission. Similarly, among the three cases with probable virus infection in utero (cases 4, 6, and 7), the most likely candidates were among viruses from the second and third trimesters and at birth. Only in the one case in which transmission seemed to occur at or just prior to delivery (case 5) were linked viruses most abundant in the mother at delivery. Finally, the fact that two of the seven children had a very late diagnosis despite repeated earlier testing was also unusual (11), especially since breast-feeding was denied.

Several factors may have influenced our failure to link the timing of maternal variant abundance to transmission: the restriction to mothers who had positive virus isolations during pregnancy could have biased the data, and myriad potential sources of virus were not available for evaluation. For example, the collections of blood-derived specimens available to this study were not comprehensive. Furthermore, HIV can poten-

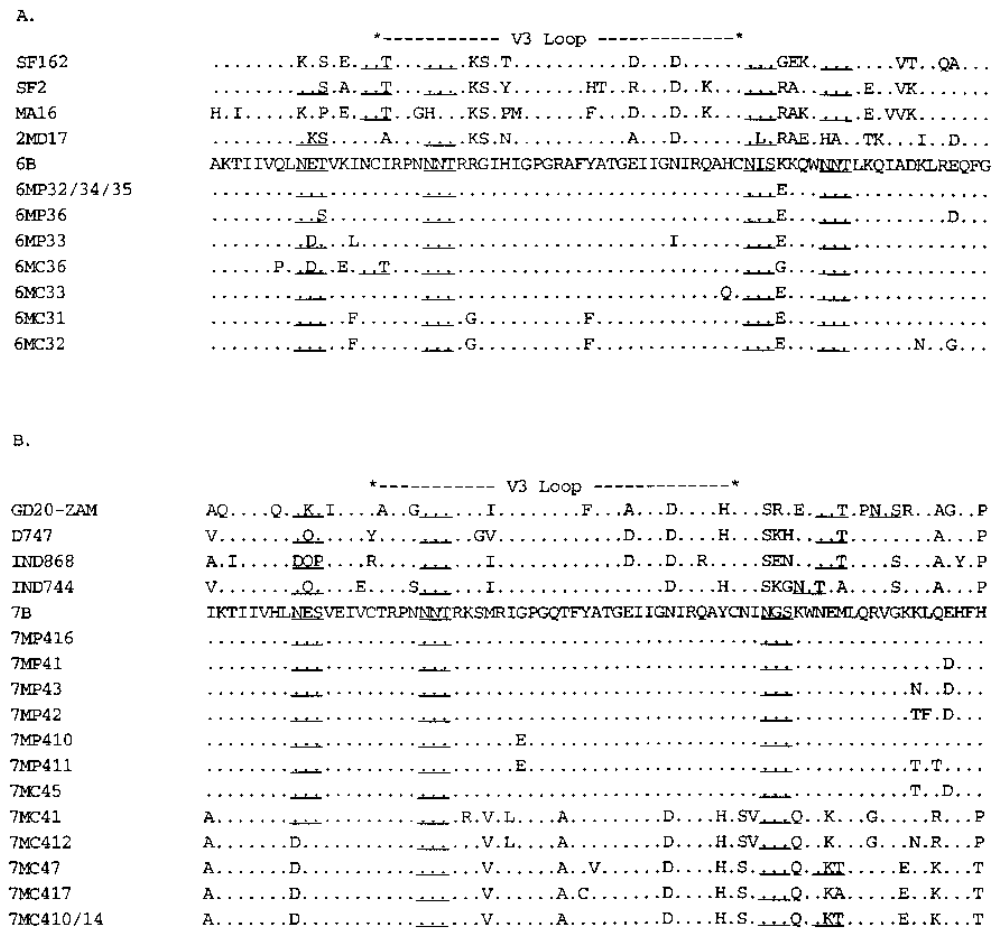

FIG. 6. N-linked glycosylation sites in the V3 loop region of viruses from cases 6 (A) and 7 (B). The sequences from case 7 are aligned with reference sequences from *env* clade C, and sequences from case 6 are aligned with reference sequences from *env* clade B (28). The V3 loop is indicated above each alignment. Potential N-linked glycosylation sites are underlined.

tially be transmitted across the placenta in utero, from blood and cervical-vaginal secretions during delivery, later from breast milk, and as cell-free as well as cell-associated virus. Indeed, because of the temporal fluctuations of viral variant populations that occur in different fluids and tissues within the body (8, 14, 20, 31), viruses in the peripheral blood may not adequately represent the transmitting virus population. Detailed study of virus populations in the maternal placenta as well as within the vaginal or uterine compartment and breast milk, as well as the blood, may yield a closer linkage to the transmitted forms. An alternative hypothesis for a delayed appearance of virus in the infant, other than infection during delivery, that also was not refuted in this study is that replication could somehow be delayed in the infant following transmission in utero, perhaps due to the status of the maternal immune response and passive transfer of antibody.

Our results are consistent with those of Mulder-Kampinga et al. (27), who also failed to support the timing hypothesis through the analysis of viral sequences from infants and infected mothers. It is possible that viral genetic studies will be unable to pinpoint the time of transmission with any certainty. First, as shown here, even though the majority of virus effectively turns over or evolves to unique forms in a short period of time, small amounts of preceding forms of the virus may reside in the PBMC as replication-competent provirus for extended periods of time. Second, because of the possible compartmentalization of virus variants in different parts of the body, the availability of maternal specimens harboring the transmitted variant populations for testing is not certain.

Any genetic features of *env* genes that correlate with transmission, throughout virus populations as diverse as five *env* sequence clades, would be strong evidence for involvement of the observed feature in transmission from mothers to their infants. However, neither signature sequences nor patterns of potential N-linked glycosylation or charged amino acids were apparent among the transmitted variants.

ACKNOWLEDGMENTS

The first two authors contributed equally to this work.

The expert technical assistance of Helena Strand, Vanja Sundberg, and Ulla Lips is gratefully acknowledged.

This work was supported by scholar award 700281-10-RF to C.H.C. from the American Foundation for AIDS Research, grants 50170-11-PG and 55510-2-ARI from the Pediatric AIDS Foundation, and grants B92-16H-9181-03, B93-16H-10395-01A, and B95-16X-10395-03A from the Swedish Medical Research Council.

REFERENCES

1. **Ahmad, N., B. M. Baroudy, R. C. Baker, and C. Chappey.** 1995. Genetic analysis of human immunodeficiency virus type 1 envelope V3 region isolates from mothers and infants after perinatal transmission. J. Virol. **69:**1001–1012.

2. **Ammann, A., P. Bolme, W. Borkowsky, Y. Bryson, G. Castelli, L. C. Bianchi, P. D'Argenio, A. Delgado, A. DeMaria, M. Duse, A. Ehrnst, E. M. Fenyö, J. Goedert, E. Guzzanti, E. Halapi, S. Livadiotti, B. Mathieson, V. Moschese, G. Mulder, J. I. Mullins, P. Nara, B. Nkowane, G. Pardi, C. Peckman, A. Rubinstein, P. Rossi, G. Scarlatti, A. Semprini, M. Stegnano, M. St. Louis, P. A. Tovo, R. Viscarello, H. Wigzell, and S. Wolinsky.** 1992. Maternal factors involved in mother-to-child transmission of HIV-1: report of the Consensus Workshop held in Siena, Italy, January 17–18, 1992. J. Acquired Immune Defic. Syndr. **5:**1019–1029.
3. **Boyer, P. J., M. Dillon, M. Navaie, A. Deveikis, M. Keller, S. O'Rourke, and Y. J. Bryson.** 1994. Factors predictive of maternal-fetal transmission of HIV-1: preliminary analysis of Zidovudine given during pregnancy and/or delivery. JAMA **271:**1925–1930.
4. **Brossard, Y., J. T. Aubin, L. Mandelbrot, C. Bignozzi, D. Brand, A. Chaput, J. Roume, N. Mulliez, F. Mallet, H. Agut, F. Barin, C. Brechot, A. Goudeau, J.-M. Huraux, J. Barrat, P. Blot, J. Chavinie, N. Ciraru-Vigneron, P. Engelman, F. Herve, E. Papierik, and R. Henrion.** 1995. Frequency of early in utero HIV-1 infection: a blind DNA polymerase chain reaction study on 100 fetal thymuses. AIDS **9:**359–366.
5. **Connor, E. M., R. S. Sperling, R. Gelber, P. Kiselev, G. Scott, M. J. O'Sullivan, R. VanDyke, M. Bey, W. Shearer, R. L. Jacobson, et al.** 1994. Reduction of maternal-infant transmission of HIV-1 with zidovudine treatment. N. Engl. J. Med. **331:**1173–1180.
6. **Datta, P., J. E. Embree, J. K. Kreiss, J. O. Ndinya-Achola, M. Braddick, M. Temmerman, N. J. Nagelkerke, G. Maitha, K. K. Holmes, P. Piot, H. O. Pamba, and F. A. Plummer.** 1994. Mother-to-child transmission of HIV-1: report from the Nairobi study. J. Infect. Dis. **170:**1134–1140.
7. **Delwart, E. L., M. P. Busch, M. L. Kalish, J. W. Mosley, and J. I. Mullins.** 1995. Rapid molecular epidemiology of HIV transmission. AIDS Res. Hum. Retroviruses **11:**1181–1193.
8. **Delwart, E. L., H. Pan, H. W. Sheppard, W. A. Paxton, R. A. Koup, D. Wolpert, A. U. Neumann, B. Korber, and J. I. Mullins.** HIV-1 quasispecies evolution during progression to AIDS. Submitted for publication.
9. **Delwart, E. L., H. W. Sheppard, B. D. Walker, J. Goudsmit, and J. I. Mullins.** 1994. Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. J. Virol. **68:**6672–6683.
10. **Delwart, E. L., E. G. Shpaer, F. E. McCutchan, J. Louwagie, M. Grez, H. Rübsamen-Waigmann, and J. I. Mullins.** 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 *env* genes. Science **262:**1257–1261.
11. **Dunn, D. T., C. D. Brandt, A. Krivine, S. A. Cassol, P. Roques, W. Borkowsky, A. De Rossi, E. Denamur, A. Ehrnst, C. Loveday, J.-A. Harris, K. McIntosh, A. M. Comeau, T. Rakusan, M.-L. Newell, and C. S. Peckham.** 1995. The sensitivity of HIV-1 DNA polymerase chain reaction in the neonatal period and the relative contributions of intra-uterine and intra-partum transmission. AIDS **9:**F7–F11.
12. **Ehrnst, A., S. Lindgren, M. Dictor, B. Johansson, A. Sönnerborg, J. Czajkowski, G. Sundin, and A.-B. Bohlin.** 1991. HIV in pregnant women and their offspring: evidence for late transmission. Lancet **338:**203–207.
13. **Ehrnst, A., A. Sönnerborg, S. Bergdahl, and Ö. Strannegård.** 1988. Efficient isolation of HIV from plasma during different stages of HIV infection. J. Med. Virol. **26:**23–32.
14. **Epstein, L. G., C. Kuiken, B. M. Blumberg, S. Hartman, L. R. Sharer, M. Clement, and J. Goudsmit.** 1991. HIV-1 V3 domain variation in brain and spleen of children with AIDS: tissue-specific evolution within host-determined quasispecies. Virology **180:**583–590.
15. **Faulkner, D. V., and J. Jurka.** 1988. Multiple aligned sequence editor (MASE). Trends Biochem. Sci. **13:**321–322.
16. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791.
17. **Felsenstein, J.** 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. University of Washington, Seattle, Wash.
18. **Frenkel, L. M., L. E. Wagner II, L. M. Demeter, S. Dewhurst, R. W. Coombs, B. L. Murante, and R. C. Reichman.** 1995. Effects of zidovudine use during pregnancy on resistance and vertical transmission of HIV-1. Clin. Infect. Dis. **20:**1321–1326.
19. **Gaillard, C., and F. Strauss.** 1990. Ethanol precipitation of DNA with linear polyacrylamide as carrier. Nucleic Acids Res. **18:**378.
20. **Goulston, C., E. Stevens, D. Gallo, J. I. Mullins, C. V. Hanson, and D. Katzenstein.** 1996. HIV in plasma and genital secretions during the menstrual cycle. J. Infect. Dis. **174:**858–861.
21. **Kimura, M.** 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16:**111–120.
22. **Korber, B. T. M., G. Learn, J. I. Mullins, B. H. Hahn, and S. Wolinsky.** 1995. Protecting HIV sequence databases. Nature **378:**242–243.
23. **Lamers, S. L., J. W. Sleasman, J. X. She, K. A. Barrie, S. M. Pomeroy, D. J. Barrett, and M. M. Goodenow.** 1994. Persistence of multiple maternal genotypes of HIV-1 in infants infected by vertical transmission. J. Clin. Invest. **93:**380–390.
24. **Learn, G. H., B. T. M. Korber, B. Foley, B. H. Hahn, S. M. Wolinsky, and J. I. Mullins.** 1996. Maintaining the integrity of human immunodeficiency virus sequence databases. J. Virol. **70:**5720–5730.
25. **Liu, S.-L., A. G. Rodrigo, R. Shankarappa, G. H. Learn, L. Hsu, O. Davidov, L. P. Zhao, and J. I. Mullins.** 1996. HIV quasispecies and resampling. Science **273:**415–416.
26. **Mulder-Kampinga, G. A., C. Kuiken, J. Dekker, H. J. Scherpbier, K. Boer, and J. Goudsmit.** 1993. Genomic HIV-1 RNA variation in mother and child following intra-uterine virus transmission. J. Gen. Virol. **74:**1747–1756.
27. **Mulder-Kampinga, G. A., A. Simonon, C. L. Kuiken, J. Dekker, H. J. Scherpbier, P. van de Perre, K. Boer, and J. Goudsmit.** 1995. Similarity in *env* and *gag* genes between genomic RNAs of human immunodeficiency virus type 1 (HIV-1) from mother and infant is unrelated to time of HIV-1 RNA positivity in the child. J. Virol. **69:**2285–2296.
28. **Myers, G., B. Korber, B. H. Hahn, K.-T. Jeang, J. W. Mellors, F. E. McCutchan, L. E. Henderson, and G. N. Pavlakis.** 1995. Human retroviruses and AIDS 1995. A compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, N. Mex.
29. **Newell, M. L., et al., and the European Collaborative Study.** 1994. Natural history of vertically acquired HIV-1 infection. Pediatrics **94:**815–819.
30. **Newell, M. L., D. Dunn, C. S. Peckham, A. E. Ades, G. Pardi, and A. E. Semprini, and the European Collaborative Study.** 1992. Risk factors for mother-to-child transmission of HIV-1. Lancet **339:**1007–1012.
31. **Poss, M., H. L. Martin, J. K. Kreiss, L. Granville, B. Chohan, P. Nyange, K. Mandaliya, and J. Overbaugh.** 1995. Diversity in virus populations from genital secretions and peripheral blood from women recently infected with human immunodeficiency virus type 1. J. Virol. **69:**8118–8122.
32. **Saitou, N., and M. Nei.** 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
33. **Simonon, A., P. Lepage, and E. Karita.** 1994. An assessment of the timing of mother-to-child transmission of HIV-1 by means of polymerase chain reaction. J. Acquired Immune Defic. Syndr. **7:**952–957.
34. **Sönnerborg, A., T. Bergström, B. Johansson, C. Håkansson, I. Julander, G. Norkrans, B. Svennerholm, and Ö. Strannegård.** 1991. Demonstration of HIV-1 DNA by polymerase chain reaction in immunocompetent HIV-1 antibody positive, but not in seronegative, homosexual men. Immunol. Infect. Dis. **1:**85–89.
35. **Smith, S. W., R. Overbeek, C. R. Woese, W. Gilbert, and P. M. Gillevet.** 1994. The genetic data environment: an expandable GUI for multiple sequence analysis. CABIOS **10:**671–675.
36. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw.** 1995. Viral dynamics of HIV-1 infection. Nature **373:**117–122.
37. **Wike, C. M., B. T. Korber, M. R. Daniels, C. Hutto, J. Munoz, M. Furtado, W. Parks, A. Saah, M. Bulterys, and J. B. Kurawige.** 1992. HIV-1 sequence variation between isolates from mother-infant transmission pairs. AIDS Res. Hum. Retroviruses **8:**1297–1300.
38. **Wolinsky, S. M., C. M. Wike, B. T. M. Korber, C. Hutto, W. P. Parks, L. L. Rosenblum, K. J. Kunstman, M. R. Furtado, and J. L. Muñoz.** 1992. Selective transmission of HIV-1 variants from mother to infants. Science **255:** 1134–1137.