*Proc. Natl. Acad. Sci. USA*
Vol. 90, pp. 1721–1725, March 1993
Medical Sciences

# 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

GABRIELLA SCARLATTI*†, THOMAS LEITNER‡, EVA HALAPI§, JOHAN WAHLBERG‡, PAOLA MARCHISIO¶, MARIA ANGELA CLERICI-SCHOELLER‖, HANS WIGZELL§, EVA MARIA FENYÖ*, JAN ALBERT**, MATHIAS UHLÉN‡, AND PAOLO ROSSI§††

Departments of *Virology and §Immunology, Karolinska Institute, Stockholm, Sweden; ‡Department of Biochemistry and Biotechnology, Royal Institute of Technology, Stockholm, Sweden; Departments of ¶Pediatrics IV and ‖Pediatrics I, University of Milan, Milan, Italy; **Department of Virology, National Bacteriological Laboratory, Stockholm, Sweden; and ††Department of Pediatrics, University of Rome "Tor Vergata," Rome, Italy

*Communicated by Sune Bergström, November 9, 1992*

ABSTRACT **We have compared the variable region 3 sequences from 10 human immunodeficiency virus type 1 (HIV-1)-infected infants to virus sequences from the corresponding mothers. The sequences were derived from DNA of uncultured peripheral blood mononuclear cells (PBMC), DNA of cultured PBMC, and RNA from serum collected at or shortly after delivery. The infected infants, in contrast to the mothers, harbored homogeneous virus populations. Comparison of sequences from the children and clones derived from DNA of the corresponding mothers showed that the transmitted virus represented either a minor or a major virus population of the mother. In contrast to an earlier study, we found no evidence of selection of minor virus variants during transmission. Furthermore, the transmitted virus variant did not show any characteristic molecular features. In some cases the transmitted virus was more related to the virus RNA population of the mother and in other cases it was more related to the virus DNA population. This suggests that either cell-free or cell-associated virus may be transmitted. These data will help AIDS researchers to understand the mechanism of transmission and to plan strategies for prevention of transmission.**

Human immunodeficiency virus type 1 (HIV-1)-infected mothers transmit the virus to their offspring in 13–40% of their pregnancies (1, 2). It has been recently suggested that transmission takes place at or shortly before delivery (3, 4) and that mothers with low $CD4^+$ lymphocyte counts and p24 antigenemia (2, 5) have an increased risk of transmission.

Recent data obtained by comparing HIV-1 DNA sequences from three mother–child pairs suggested that one single genotype is selected from the genetically heterogeneous virus population of the mother and is transmitted to the child (6). Moreover, the loss of an N-linked glycosylation site (7) immediately amino terminal to the variable region 3 (V3) loop of the viral envelope was suggested as a general characteristic of the transmitted virus (6). Mutations in this region could potentially affect mother-to-child transmission, since the V3 loop is an important determinant for cellular tropism and virus neutralization (8, 9). However, Wolfs *et al.* (10) recently reported that HIV-1 transmission by the sexual or parenteral route is not associated with a shift in the virus population from transmitter to recipient.

To further elucidate the molecular mechanisms involved in mother-to-child transmission of HIV-1 we compared V3 sequences from 10 HIV-1-infected infants with sequences from the corresponding mothers by analyzing RNA from serum, DNA from uncultured peripheral blood mononuclear cells (PBMC), and DNA from virus isolates obtained at or shortly after delivery.‡‡ The maternal virus populations were highly heterogeneous. In contrast, the children were shown to harbor homogeneous virus populations. In some cases the transmitted virus showed close similarity to the maternal proviral DNA sequences (derived from PBMC), and in other cases it showed similarity to virion RNA sequences (derived from serum). The N-linked glycosylation site, immediately amino terminal to the V3 loop, was absent from only 1 child (out of 10) and did not appear to play a crucial role in transmission.

## MATERIALS AND METHODS

**Study Population.** Ten HIV-1-infected mother–child pairs were included in this study. The samples were provided by the National Care System for HIV-infected mothers, the Department of Gynecology and Obstetrics VII, and the Departments of Pediatrics I and IV of the University of Milan, Italy. All the mothers were drug users or sexual partners of drug users. Maternal blood samples were collected at delivery ($n = 6$) or between 1 and 4 months after delivery (two at 1 month, one at 3 months, and one at 4 months after delivery). For the children the first available virus-positive sample was used. One child was positive by HIV-1 polymerase chain reaction (PCR) and virus isolation (5) at birth, three at 1 month of age, one at 2 months, four at 3 months, and one at 4 months, when this child was tested for the first time. Nine children were asymptomatic at the time of sample collection, and one child (193) with hepatosplenomegaly and hepatitis was classified as P2AF according to the criteria of the Centers for Disease Control (11).

**DNA Preparation.** DNA for PCR amplification was prepared directly from uncultured patients' PBMC (12) and from virus isolates after one passage on phytohemoagglutinin (PHA)-stimulated PBMC from healthy blood donors as previously described (13, 14). Briefly, PBMC ($2–4 \times 10^6$) were resuspended in PCR-lysis buffer (10 mM Tris·HCl, pH 8.3/1 mM EDTA/0.5% Nonidet P-40/0.5% Tween 20, containing proteinase K at 300 $\mu$g/ml) at a concentration of $10^6$ cells per 100 $\mu$l of buffer and digested with proteinase K overnight at

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HIV-1, human immunodeficiency virus type 1; V3, variable region 3; PBMC, peripheral blood mononuclear cells.

†To whom reprint requests should be addressed at: Department of Virology, Karolinska Institute, Lundagatan 2, 105 21 Stockholm, Sweden.

‡‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L08277–L08372).

37°C. The proteinase K was then inactivated by heating for 15 min at 94°C.

**RNA Extraction and cDNA Synthesis.** HIV-1 RNA was extracted from the serum as previously described (15). In short, $(dT)_{25}$-coated magnetic beads (25 μl; Dynal AS, Oslo) were used to extract viral genomic RNA from 25 μl of serum. The extracted RNA (20 μl) was immediately reverse transcribed into cDNA by using the primer JA 12 (12), which is complementary to the gp120 V3 region of the *env* gene of HIV-1. The reverse transcription mixture (30 μl) contained (final concentration) 50 mM Tris·HCl at pH 8.3, 8 mM $MgCl_2$, 30 mM KCl, 10 mM dithiothreitol, 1.7 mM each deoxynucleoside triphosphate (dATP, dCTP, dGTP, and dTTP; Pharmacia), 0.5 μM primer JA 12, 3 units of Moloney murine leukemia virus reverse transcriptase (Pharmacia), and 16 units of RNAguard (Pharmacia). Reverse transcription was performed for 1 hour at 37°C. Sera from blood donors and from HIV-1-infected patients (adults and children) were included as negative and positive controls, respectively.

**PCR.** DNA (10 μl) or cDNA (5-μl) samples were amplified by PCR with nested primers specific for the gp120 V3 region of the *env* gene as previously described (12, 14). The outer primers JA9 and JA12, and one of the inner primers, JA10, have been described previously (12). The sequence of the second inner primer, JA53, was 5′-AATTTCTGGGTC-CCCTCCTG-3′, nucleotides 6930–6911 relative to the MN strain of HIV-1 (16). Briefly, the samples were first amplified for 24 cycles with the outer primers, then 1/10 (5 μl) of the product from the first PCR was amplified for 30 cycles with the inner primers. Negative controls were included in each run and consisted of lysis buffer alone and samples from healthy blood donors. Positive controls consisted of HIV-1-infected cells diluted in uninfected cells to contain 10 HIV-1 viral DNA copies.

**Cloning.** The PCR product amplified by the inner primers was purified by using the Qiagen spin 20 column (Diagen, Düsseldorf, Germany) and dissolved in 20 μl of Tris·EDTA. The PCR fragments were treated with the Klenow fragment of DNA polymerase I (Pharmacia) as well as T4 polynucleotide kinase (Pharmacia). After precipitation with ammonium acetate approximately 1/10th of the purified PCR product was ligated with 50 ng of *Sma* I-digested pUC18 vector (Pharmacia) by using the DNA ligation system (Amersham). The ligated vector was used to transform the competent cell JM105 (Pharmacia) by treatment at 42°C for 90 sec. Positive colonies were picked, resuspended in 10 μl of PCR buffer (10 mM Tris·HCl, pH 8.3/50 mM KCl/0.01% Tween 20) and treated at 95°C for 5 min. To confirm the presence of the insert in the plasmid, amplification with pUC18-specific primers (RIT 28/29) was performed (17).

**DNA Sequencing.** The PCR product amplified by the inner primers was used for direct solid-phase DNA sequencing as previously described (14, 18). Briefly, a further amplification step of the first PCR product (diluted 1:100) was performed with the primers RIT 124 and RIT 128. The amplified product was then purified by immobilization on magnetic beads (Dynabeads M280-streptavidin; Dynal AS) and subsequently denatured to obtain single-stranded DNA. Fluorescent primers [RIT 129 or RIT 43; Universal Forward Primer (Pharmacia) for the clones] were used for the sequencing reaction of both DNA strands; the product was then loaded on a 6% polyacrylamide gel in an automated laser fluorescent (A.L.F.) sequencing apparatus (Pharmacia LKB). Each sample was sequenced twice.

**Sequence Analysis and Tree Alignment.** Computer analysis was performed with the program TREEALIGN edited by Jotun Hein, November 1990 (19). Pairwise comparisons were performed to establish amino acid inter- and intrapatient distances, using the direct sequence of the children and the clones of the mothers. When no clones but direct sequences from DNA and RNA of the mothers were available, the amino acid variants derived from the direct sequence were used to create sequences with all possible different amino acid combinations. The consensus sequence was obtained from the clones by determining the most abundant amino acid at each position after alignment. Positions where an alignment gap was introduced were excluded from the evaluations. Cluster analysis was performed and dendograms were generated.

## RESULTS

The amino acid sequences of the V3 region of HIV-1 obtained by direct sequencing from the cellular DNA of the 10 children are presented in Fig. 1. Each child had a unique V3 sequence. Pairwise analysis of the entire available sequence of the children showed interpatient divergences ranging between 12% and 35%. Intrapatient sequence heterogeneity was observed in only two children (196 and 201). This indicates that the virus populations detected in very young children born to HIV-1-infected mothers are homogeneous.

```
                                  281       291       301       311       321       331       341       351       361       371
                                   |         |         |         |         |         |         |         |         |         |
consensus    age    EEEVVIRSENFTDNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLKQIVTKLREQF.?NKTIVFNQSSGG
           (months)

Child 204    (0)    --------A-------S-------T----------------P----------N-------L-----EG---K--I-------.K----N-T-----

Child 145    (1)    ----------N----------T----------------P---------.--------------------D------S-------.Q----------

Child 190    (1)    --------K--SH-HES---------A--------------P----------D----------L------D--R--AI-------..----------

Child 196    (1)    ----------SN---V------K-I--------------VTV----VW---------------L-GAQ---------V------GK----I-K-P---
                                                                                        V
Child 199    (2)    ----I---K-----T-V----------------G----------------A---------K----------D-----------Y.E----I-K-----

Child 115    (3)    ---I------IS--T-------K-------A--S----R----------------T-----------E------R--AG-------.E--------A---

Child 130    (3)    -----F--------T----------A----------R--S--------A--D-----------E----------K--G---.V-------H----

Child 136    (3)    --?--?------N--RS------Q--Q----------R-------K----.-----N-------L-----DD---R-AD-----Y.E----------

Child 201    (3)    --------A-L-----I--VH----IV------S----QG---------F-----T----K-Y--V--T---DT--K-AI-LG---.K----V-------
                                    x     v           r                          a    i  q    i         a
Child 193    (4)    -------SK???--T---------T--------------G----------A--D---N-------L------D------I--K---.G----A-K-PA--
```

FIG. 1. Amino acid sequences of the V3 regions from 10 HIV-1 infected children. The consensus sequence of the V3 loop is derived from the North American/European consensus (underlined; ref. 20), and the consensus sequence of the flanking regions is from 19 HIV-1 isolates in the Los Alamos data base (16). Sequence heterogeneity is indicated by assigning two amino acids at the same position; uppercase letters indicate amino acids in approximately equal amounts, while lowercase letters indicate minor variants. Hyphens indicate an amino acid identical to that present in the consensus sequence, periods indicate an amino acid deletion, X indicates a position where all amino acids are present in equal amounts, and ? indicates a position which could not be resolved.

In contrast, the maternal V3 loop sequences, whether derived from uncultured PBMC, from virus isolates passaged on blood donor PBMC, or from RNA extracted from serum, were heterogeneous (Fig. 2). In one case (199) the degree of heterogeneity was too high to allow proper interpretation of the direct sequence. Sequences derived from the cultured PBMC showed less heterogeneity than those derived from serum or uncultured PBMC. No correlation could be found between the degree of heterogeneity and the clinical stage of the mother (data not shown). The PBMC-derived PCR products from five mothers (190, 193, 196, 199, 204) were cloned, and 9–22 clones from each sample were sequenced (Fig. 3). The estimation of the degree of genetic polymorphism by

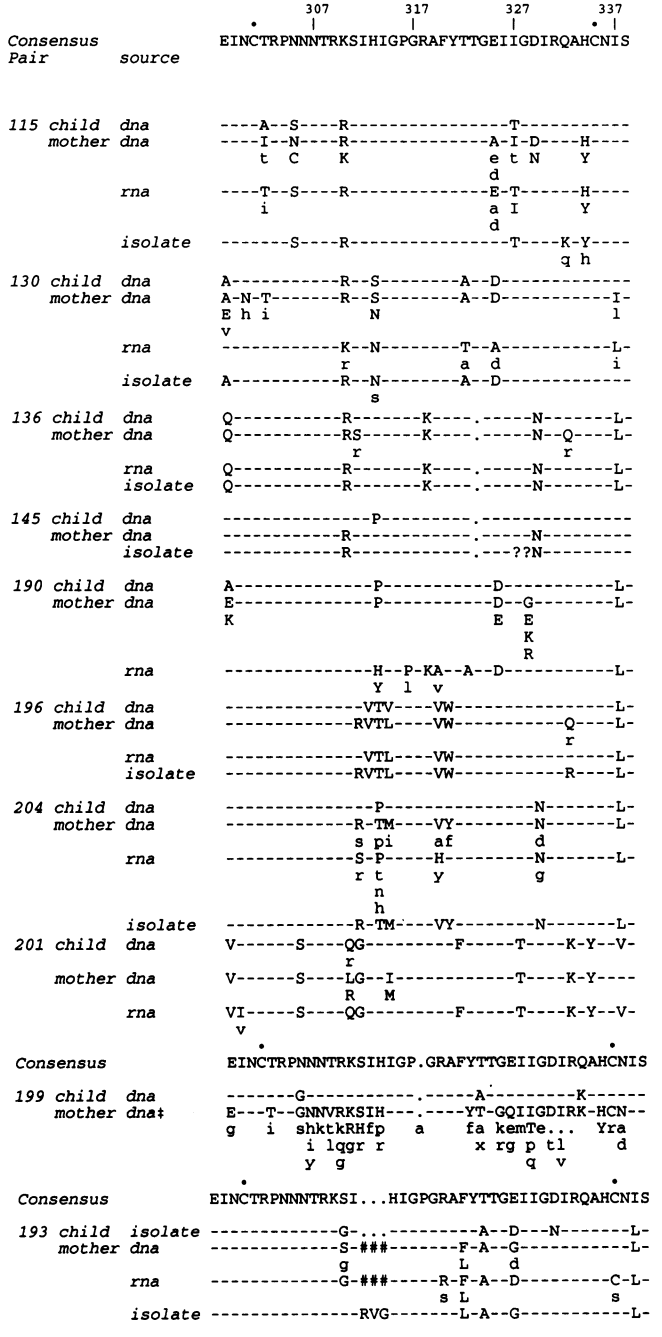


FIG. 2. Comparisons of the V3 loop amino acid sequences derived from the children to sequences obtained by direct sequencing of uncultured PBMC (dna), cultured PBMC (isolate), and serum (rna) of the corresponding mothers. Symbols are as in Fig. 1. ‡ indicates consensus sequence derived from DNA clones.

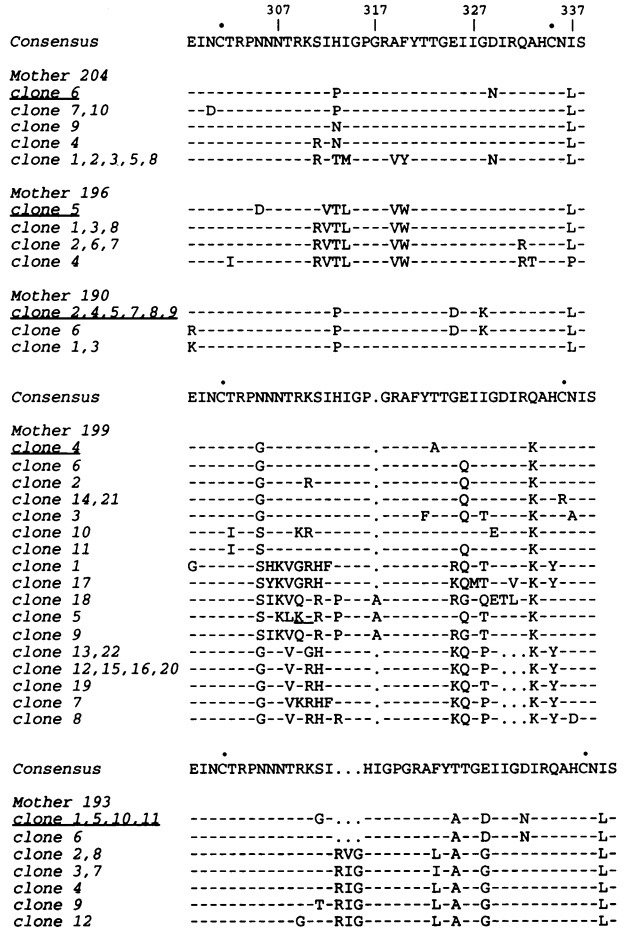


FIG. 3. Amino acid sequences of the V3 loops of clones derived from the uncultured PBMC (DNA) of five mothers. Symbols are as in Fig. 1. Underlined amino acid sequences indicate frame shifts. Underlined clones are identical or similar to the sequence of the child.

sequencing of multiple clones did not differ significantly from that obtained by direct sequencing.

Next we attempted to trace the origin of the children's sequences by comparing sequences obtained from the corresponding mother's PBMC, serum, and virus isolate. Phylogenetic tree analysis showed that the sequences of four children (115, 136, 196, and 201) were closer to the RNA-derived maternal sequence, whereas the sequences of the other four (130, 190, 193, and 204) were more related to the DNA-derived maternal sequence (uncultured PBMC or clones). Mother–child pairs 145 and 199 were not included in the evaluation because sequences derived from the maternal serum were not available. Two phylogenetic trees representative of the two patterns are shown in Fig. 4. No correlation was found between the presence of p24 antigen in the serum of the mother and the apparent origin of the virus transmitted to the child (data not shown).

The comparison of the children's sequences with sequences derived from the corresponding maternal DNA showed that the transmitted virus represented either a minor (in three cases) or a major (in two cases) virus variant of the mother (Fig. 3). The cysteine residues at positions 301 and 335 [numbered according to HIV-1 MN strain (16)] and the N-linked glycosylation sites at positions 300, 306, and 336, which are well conserved among available HIV-1 sequences, were usually conserved in the sequences from both children and mothers. Three mothers (115, 130, and 199), however, harbored sequences with heterogeneity at position 302, indicating the presence of a minor virus variant without the

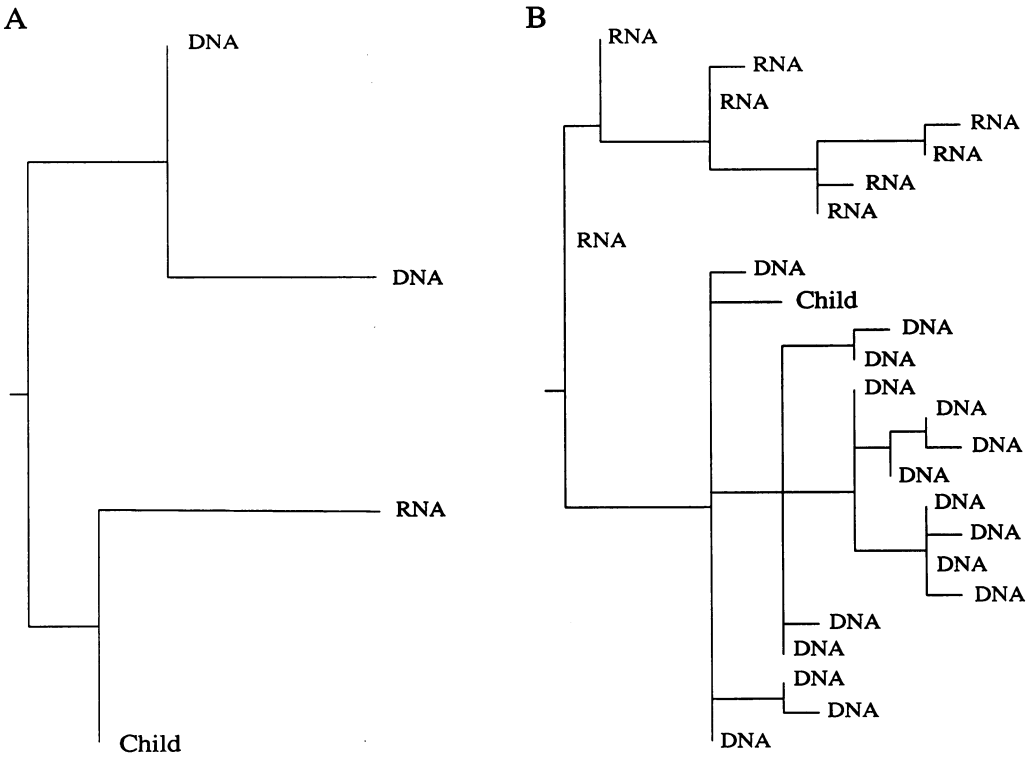


FIG. 4. Dendogram generated from the phylogenetic tree analysis of mother–child pairs 196 (*A*) and 190 (*B*). The sequence of the child was derived from the DNA from uncultured PBMC. The sequences of the mother were derived from uncultured PBMC (DNA) and from serum (RNA). When several amino acid variants were detected by direct sequencing we created sequences with all possible different amino acid combinations (see *Sequence Analysis and Tree Alignment*).

N-linked glycosylation site in this position. Importantly, this virus variant lacking the glycosylation site was transmitted in only one of those three cases (child 115). No inactivating mutations were observed, but two mother–child pairs displayed deletions of single amino acids in the V3 loop.

## DISCUSSION

We have compared the V3 sequences from 10 HIV-1-infected infants with virus sequences derived from DNA of uncultured PBMC, DNA of cultured PBMC, and RNA from serum of the corresponding mothers. The virus population harbored by the mothers showed a significant degree of heterogeneity. In contrast, the children, who were evaluated within 4 months from birth, displayed highly homogeneous virus populations. This agrees with the findings of previous studies (6, 10) and suggests either that a single genotype is transmitted or that there is a selective outgrowth of certain genotypes during primary infection. In this study we could not discriminate between these two alternatives, but in a heterosexual transmission case we have found evidence for the latter explanation (J.A., J.W., and M.U., unpublished results).

The present study and that by Wolinsky *et al.* (6) are in agreement concerning the low heterogeneity in newly infected children, but there are also important differences between the two studies. First, transmission of virus lacking the N-linked glycosylation site amino terminal to the first cysteine of the V3 loop seems to be a rare event, since only 1 of 10 children in our material presented this genotype. The fact that only 1 of 3 mothers harboring mixed virus populations, with and without this glycosylation site, transmitted the latter variant further suggests that this mutation does not confer selective advantage for transmission. The distinct results of the two studies may reflect differences in time of sampling, clinical stage, geographic origin, or risk group of the two populations.

Second, our results also demonstrate that the virus transmitted to the child may be a minor as well as a major variant in the virus population of the mother. Transmission of a minor variant, capable of escaping the maternal immune response, does not appear to be a general rule applicable to all cases. However, virus neutralization may still be of importance, since we have observed that mothers who transmit the virus to their children frequently lack neutralizing antibodies against autologous and heterologous primary isolates (G.S., J.A., P.R., V.H., P.B., L.M., and E.M.F., unpublished data). Furthermore, the transmitted virus appeared in some cases to be more related to the virus RNA population and in other cases to the virus DNA population of the mother. Thus, by studying only viral DNA populations Wolinsky *et al.* (6), may have biased their data in favor of transmission of minor variants if the transmitted viruses were more related to the RNA population of viruses in the mother.

Recent studies (3, 4) indicate that mother-to-child transmission occurs at or close to delivery. However, little is known about the mode of transmission—i.e., if transmission occurs by means of cell-free or cell-associated virus. Our comparisons of the virus present in the child to the RNA and DNA-derived sequences of the mother indicate that both pathways may be operative. It is also possible that similarities of the children's virus to the maternal DNA- or RNA-derived sequences may reflect differences in the time gap between transmission and sampling.

In conclusion, we have found no evidence of selection of minor virus variants during transmission. No characteristic molecular features of the transmitted virus variants could be identified. Furthermore, our data indicate that either cell-free or cell-associated virus may be transmitted. These data will help AIDS researchers to elucidate the mechanism of transmission and to choose strategies for preventing transmission.

Special thanks are given to our colleagues from the Divisions of

Obstetrics and Gynecology and Neonatology and Pediatrics at hospitals "L. Mangiagalli," "L. Sacco," and "G. D. De Marchi." This work was supported by the Istituto Superiore di Sanitá—Progetto AIDS, the Pediatric AIDS Foundation, the Swedish Medical Research Council, Nutek, and the U.S. Army Medical Research and Development Command (Grant DAMD 17-89-Z-9038). G. S. was supported by a fellowship from the Istituto Superiore di Sanitá.

1. Ryder, R. W., Nsa, W., Hassig, S. E., Behets, F., Rayfield, M., Ekungola, B., Nelson, A. M., Mulenda, U., Francis, H., Mwandagalirwa, K., Davachi, F., Rogers, M., Nzilambi, N., Greenberg, A., Mann, J., Quinn, T. C., Piot, P. & Curran, J. W. (1989) *N. Engl. J. Med.* **320**, 1637–1642.
2. European Collaborative Study (1992) *Lancet* **339**, 1007–1012.
3. Ehrnst, A., Lindgren, S., Dictor, M., Johansson, B., Sönnerborg, A., Czajkowski, J., Sundin, G. & Bohlin, A.-B. (1991) *Lancet* **338**, 203–207.
4. Goedert, J. J., Duliège, A. M., Amos, C. I., Felton, S. & Biggar, R. J. (1991) *Lancet* **338**, 1471–1475.
5. Scarlatti, G., Lombardi, V., Plebani, A., Principi, N., Vegni, C., Ferraris, G., Bucceri, A., Fenyö, E. M., Wigzell, H., Rossi, P. & Albert, J. (1991) *AIDS* **5**, 1173–1178.
6. Wolinsky, S. M., Wike, C. M., Korber, B. T. M., Hutto, C., Parks, W. P., Rosenblum, L. L., Kunstman, K. J., Furtado, M. R. & Munoz, J. L. (1992) *Science* **255**, 1134–1137.
7. Rademacher, T. W., Parekh, R. B. & Dwek, R. A. (1988) *Annu. Rev. Biochem.* **57**, 785–838.
8. Shioda, T., Levy, J. A. & Cheng-Mayer, C. (1991) *Nature (London)* **349**, 167–169.
9. Palker, T. J., Clark, M. E., Langlois, A., Matthews, T. J., Weinhold, K. J., Randall, R. R., Bolognesi, D. P. & Haynes, B. F. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1932–1936.
10. Wolfs, T. F. W., Zwart, G., Bakker, M. & Goudsmit, J. (1992) *Virology* **189**, 103–110.
11. Centers for Disease Control (1987) *Morbid. Mortal. Wkly. Rep.* **36**, 225–235.
12. Albert, J. & Fenyö, E. M. (1990) *J. Clin. Microbiol.* **28**, 1560–1564.
13. Åsjo, B., Morfeldt-Månson, L., Albert, J., Biberfeld, G., Karlsson, A., Lidman, K. & Fenyö, E. M. (1986) *Lancet* **ii**, 660–662.
14. Wahlberg, J., Albert, J., Lundeberg, J., von Gegerfelt, A., Broliden, K., Utter, G., Fenyö, E. M. & Uhlén, M. (1991) *AIDS Res. Hum. Retroviruses* **7**, 983–990.
15. Chiodi, F., Keys, B., Albert, J., Hagberg, L., Lundeberg, J., Uhlén, M., Fenyö, E. M. & Norkrans, G. (1992) *J. Clin. Microbiol.* **30**, 1768–1771.
16. Myers, G., Korber, B., Berzofsky, J. A., Smith, R. F. & Pavlakis, G. N. (1991) *Human Retroviruses and AIDS* (Los Alamos Nat. Lab., Los Alamos, NM).
17. Hultman, T., Murby, M., Ståhl, S., Hornes, E. & Uhlén, M. (1990) *Nucleic Acids Res.* **18**, 5107–5112.
18. Hultman, T., Ståhl, S., Hornes, E. & Uhlén, M. (1989) *Nucleic Acids Res.* **17**, 4937–4946.
19. Hein, J. (1990) *Methods Enzymol.* **183**, 626–645.
20. LaRosa, G. J., Davide, J. P., Weinhold, K., Waterbury, J. A., Profy, A. T., Lewis, J. A., Langlois, A. J., Dreesman, G. R., Shadduck, P., Holley, L. H., Karplus, M., Bolognesi, D. P., Matthews, T. J., Emini, E. A. & Putney, S. D. (1990) *Science* **249**, 932–935.