INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

CHARACTERIZATION OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 ASSOCIATED WITH AND WITHOUT VERTICAL TRANSMISSION

by

Tobias Hahn

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2002

THE UNIVERSITY OF ARIZONA ©
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Tobias Hahn

entitled Characterization of Human Immunodeficiency Virus Type 1 Associated With and Without Vertical Transmission

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Nafees Ahmad	4/9/02 Date
Harris Bernstein	4/9/02 Date
Janet Decker	4/9/02 Date
Mark Dodson	April 9, 2002 Date
	Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Nafees Ahmad — 4/9/02

Dissertation Director Nafees Ahmad — Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Tors Ibl

ACKNOWLEDGMENTS

I wish to thank my advisor, Dr. Nafees Ahmad, for guiding me through my graduate career. I was lucky to have worked under him and benefited from his teachings. I am grateful for his patience and for always supporting me during the stay in his laboratory.

I also would like to thank my committee, Dr. Bernstein, Dr. Decker, Dr. Dodson, and Dr. Hall for their advice, support and patience over the past years.

Additionally, I would like to thank the members of my lab, past and present, for their assistance and support.

In addition, I would like to acknowledge Dr. Colombe Chappey for statistical analysis on the p17MA mother-infant project. I also wish to thank Dr. Raymond C. Baker at the Children's Hospital Research Center in Cincinnati, Ohio and Dr. Ziad Shehab at the University of Arizona Medical Center for providing blood samples which are the basis of this work. Additionally, I wish to thank the staff of the Biotechnology Computing Facility for their help with the computer systems used for the sequence analysis. I also would like to acknowledge the Department of Microbiology and Immunology, its faculty and administrative staff, for their support. I thank the Department of Veterinary Science and Microbiology and the Department of Biochemistry for financial assistance in the form of teaching assistantships, without I would not have been able to complete this study. This project would not have been possible without the funding from the National Institutes of Health and the Arizona Disease Control Research Commission.

DEDICATION

I would like to dedicate this dissertation to my parents, Barbara and Meinrad Hahn, and my future wife Melissa Grover.

Barbara und Meinrad, ohne eure Unterstüzung wäre es mir niemals möglich gewesen diese Arbeit zu vollbringen. Ihr habt immer an mich geglaubt und mich immer ermutigt meinen Träumen nachzugehen. Ich möchte euch von ganzem Herzen danken, daß ihr mir diesen Traum ermöglicht habt.

Melissa, thank for your support, patience and love. I knew, I could always come to you for a shoulder to lean on. The success of this work is just as much yours as it is mine. Thank you for everything, I would have not been able to do it without you.

TABLE OF CONTENTS

TABLE OF CONTENTS - *Continued*

TABLE OF CONTENTS - *Continued*

TABLE OF CONTENTS - *Continued*

TABLE OF CONTENTS - *Continued*

LIST OF TABLES

LIST OF ILLUSTRATIONS

ABSTRACT

Vertical transmission of human immunodeficiency virus type 1 (HIV-1) occurs at an estimated rate of 30% and accounts for 90% of all HIV-1 infections in children. Increased risk of vertical transmission correlates with advanced maternal disease status, low $CD4^+$ lymphocyte count, and high viral load. However, the molecular mechanisms of vertical transmission are poorly understood, making it difficult to design effective strategies for prevention and treatment. Our hypothesis is that specific molecular and biological properties of HIV-1 are critical determinants of vertical transmission. We characterized the HIV-1 *gag* p17 matrix (MA) and *nef* genes associated with and without vertical transmission. In addition, we determined the effect of *env* gp120 from mother-infant pairs and from infected mothers who failed to transmit the virus to their infants (non-transmitting mothers) on HIV-1 replication, cellular tropism, cytopathic effects and co-receptor utilization. Our data indicate that the open reading frames and the functional domains of both the *gag* p17MA and *nef* genes were highly conserved in isolates from mothers and their infants. While there was no significant difference in the maintenance of open reading frames and the conservation of functional domains between isolates from transmitting and non-transmitting mothers, we found that the non-transmitting mothers' *gag* p17MA sequences were more homogenous compared with the transmitting mothers' sequences. In addition, we were able to associate several motifs in p17MA with either transmitting or non-transmitting status. To study the effect of gp120 on HIV-1 biology, we reciprocal inserted the gp120 from mother-infant pairs and non-transmitting mothers into a T-tropic infectious clone and found that the

chimeras were unable to replicate in T-cell lines and did not form syncytia in MT-2 cells. Moreover, these chimeras used the CCR5 co-receptor for entry in the U373MAGI-CD4-CCR5 cell line. Both the mother-infant pairs' and the non-transmitting mothers' gp120 chimeras replicated well in primary peripheral blood lymphocytes (PBL) with no significant difference in replication kinetics. These results may be helpful in the understanding of the association of viral determinants and molecular mechanisms of vertical transmission, which may contribute towards the development of new strategies for treatment and prevention of HIV-1 infection in children.

1 INTRODUCTION

The description of a novel retrovirus as the cause of a severe immunodeficiency in humans in the early 1980s started a new era in the biomedical sciences. Soon after the human immunodeficiency virus type 1 (HIV-1) was identified as the etiologic agent of the acquired immunodeficiency syndrome (AIDS) (13, 68), a massive research effort sought to understand the biology and pathogenesis of the virus. The AIDS pandemic initiated a concerted research effort whose impact may be compared to the "Race for the Moon" of the 1960s. Just as the space program rapidly advanced the fields of physics, engineering, material- and computer sciences – advances that influenced science and industry far into the following decades – so has HIV research advanced the fields of virology, biochemistry, molecular biology, immunology, and epidemiology. However, despite the massive effort and the progress made, it took years until the first new drugs were developed that specifically targeted the virus. It took almost a decade until an at least temporarily effective therapy was developed which made it possible to manage the disease.

HIV-1 is an enveloped retrovirus that is transmitted by virus-containing body fluids, so viral spread includes transmission through blood and blood-derived products, sexual transmission, intravenous drug use, and perinatal transmission (mother-to-infant transmission). HIV-1 infects cells of the immune system and causes a severe immunodeficiency as well as other tissue pathologies, such as thymic atrophy and dementia. Due to the diminishing immune response, the infected individual is less able to fight off secondary infections, which would be only a minor challenge to an

immunocompetent host. Ultimately the patient is overcome by secondary, mostly opportunistic, infections. The most prevalent conditions HIV-1 infected individual surrender to are *Pneumocystis carinii* pneumonia (PCP), chronic tuberculosis (TB), malignancies, such as Karposis Sarcoma (KS), and a variety of viral infections. This includes reactivation of latent infections of viruses of the herpes family. On a further note, recent studies found evidence that KS may be the result of a herpes virus that is latent and benign under normal circumstances. Despite the progress made in treatment of HIV-1 infection and AIDS disease, currently there is no effective vaccine available for the prevention of HIV-1 infection. Hence, the AIDS pandemic is far from subsiding, to the contrary, in developing countries HIV-1 infections are still rapidly rising and we may have not yet seen the full socio-economic consequences of the pandemic. In addition, there is still much to be learned about the virus and discoveries made in the HIV field may also be useful in the understanding of other processes in the life sciences.

Soon after the identification of HIV-1, a second closely related human immunodeficiency virus (HIV-2) was described. In addition, other closely related primate retroviruses were found and classified as simian immunodeficiency viruses (SIV). This classification was based on molecular phylogeny, a field that greatly helped to unravel the mysteries of the retroviruses but was itself improved and expanded due to the demands of HIV research. HIV-1 was found to belong to the family *Retroviridae* and the genus *Lentivirus*. Phylogenetic analysis was used in conjunction with classical virology field work, epidemiology, and molecular methods to identify how HIV

emerged. It has been shown that SIVcpzUS, a chimpanzee strain of the simian immunodeficiency virus (SIV), is the phylogenetic precursor to HIV-1 (70) and it is now believed that HIV-1 evolved from a simian lentivirus and entered the human population from african primate species (154).

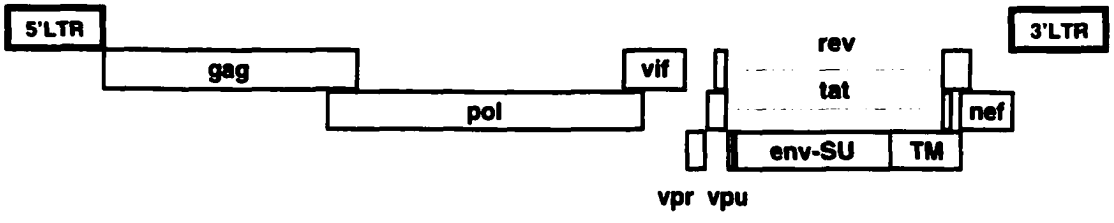
Linking HIV-1 to AIDS and the identification of the virion structure as well as the fundamental processes of the viral lifecycle were greeted with great optimism and the belief that a cure and a vaccine would soon be found. However, the more that was revealed about the virus, the more problems were identified that would complicate the development of therapeutics and vaccines. It was discovered that the HIV-1 genome is highly plastic as it displays an extremely high degree of variability. For instance, despite the production of neutralizing antibodies to the viral envelope, the virus is able to evade the host immune response. Moreover, the viral genome integrates into the host cell genome and may be clinically latent. This feature, in conjunction with the fact that the virus infects the cells of the immune system, makes it difficult for the immune system to eliminate the virus. Viral integration also provides a viral reservoir within the patient, that may never be eradicated by drug intervention. This was illustrated by the failure of the first combination drug therapy trials in the late 1990s. Therefore, despite the early victories in the battle against the virus, the prevelance of HIV-1 is still increasing. We must expand our understanding of the virus and its interplay with the host to be able to develop new treatments and, at some point, a protective vaccine.

1.1 HIV-1 Biology and Genetics

1.1.1 Genome Structure

Like the genome of all retroviruses, the HIV-1 genome consists of single-stranded positive sense RNA molecules. The virion harbors two copies of the RNA genome, that are each approximately 9.6 kb in size, therefore the virus is diploid. Since the viral genome is not very well protected from harsh environmental influences, the nearly identical copies of the genome may allow for repair by recombination during reverse transcription (37). The HIV-1 genome (Fig. 1) encompasses the common retroviral coding regions *gag* (encoding the main structural proteins), *pol* (encoding the viral enzymes), and *env* (encoding the viral envelope proteins). In addition, HIV-1 encodes two regulatory genes, *tat* and *rev*, and four accessory genes *vif*, *vpr*, *vpu* and *nef*. The coding region of the genome is flanked on each side by an untranslated long terminal repeat (LTR) that contains the viral promoter as well as other regulatory signals.

Figure 1: HIV-1 Proviral Genome

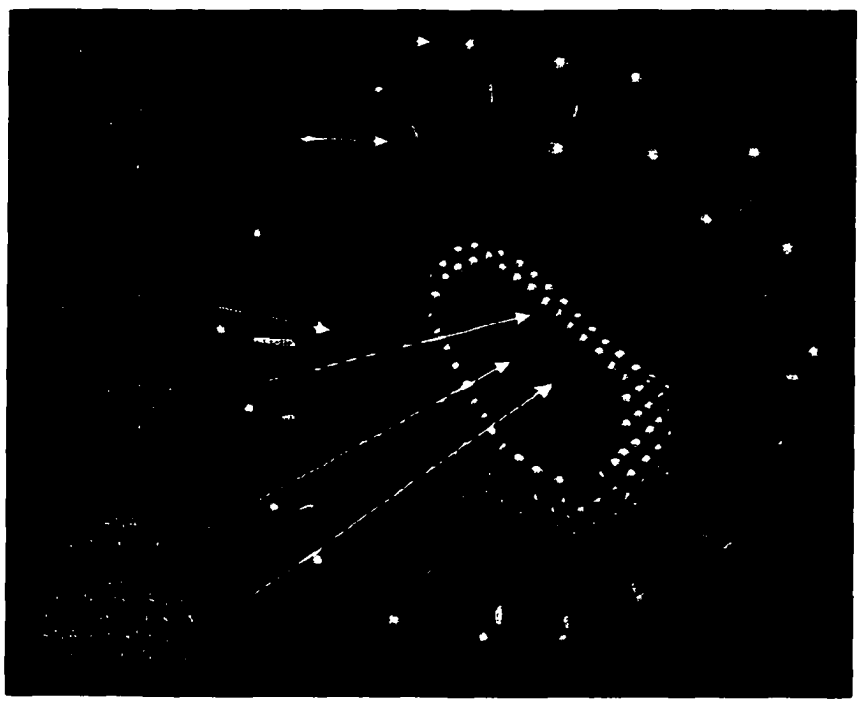


1.1.2 Virion Structure

In the mature virion (Fig. 2) the genomic RNAs are tightly associated with the nucleocapsid protein (NC). This complex is packaged into a conical shaped core that

consists of the p24 capsid protein (CA). The core contains the three viral enzymes reverse transcriptase (RT), protease (PR) and integrase (IN), several of the viral accessory gene products (Vif, Vpr, and Nef) as well as cellular proteins. The core complex itself is surrounded by a shell of the p17 matrix (MA) protein, which is the major structural protein of the virus and gives the virion its overall round shape. The outside of the MA shell is lined with the viral envelope, that consists of the host-cell derived lipid bilayer membrane. Incorporated into this membrane are the two viral envelope proteins, gp41 (TM) and gp120. The transmembrane protein (TM) spans the membrane and connects the MA shell with the viral surface glycoprotein (gp120). In addition to the viral envelope proteins, one can find numerous host-cell derived membrane proteins, such as molecules of the major histocompatibilty complex (MHC).

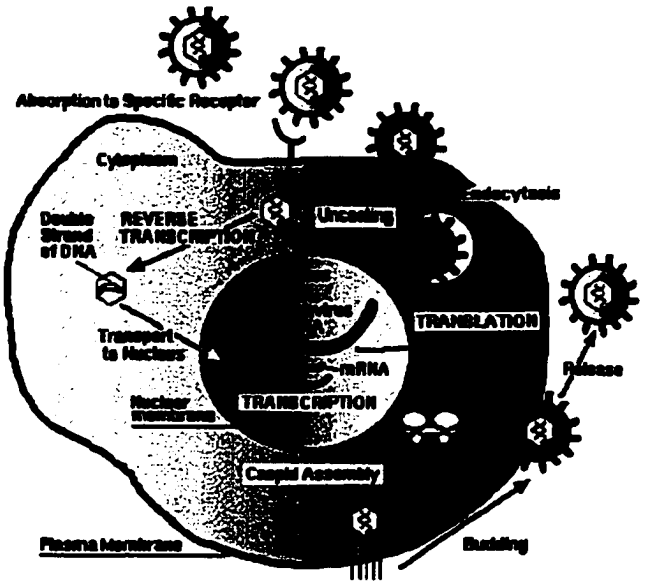
Figure 2: HIV-1 Virion



1.1.3 HIV-1 Life Cycle

The HIV-1 life cycle (Fig. 3) initiates with the attachment of the gp120 surface glycoprotein to a target-cell CD4 molecule. The CD4 receptor is a cellular glycoprotein found on the surface of helper T-cells as well as on the surfaces of cells of the monocyte/macrophage lineage. A distinctive feature of HIV-1 is that it requires more than one cell-surface receptor for infection. Several strains of HIV-1 have been characterized, that all use CD4 as the primary attachment protein but can be distinguished based on their unique requirement for a secondary attachment protein or co-receptor. HIV-1 utilizes several different co-receptors, which are members of a family of cell-surface receptors involved in cell-signaling pathways. The natural ligands of these cell-surface receptors are signaling molecules called chemokines (45, 49). The two main strains of HIV-1 use either the chemokine receptor CCR5 or the chemokine receptor CXCR4. This use of different co-receptors also correlates with the tropism of the virus. Viruses that use the CCR5 co-receptor, expressed on cells of the monocyte/macrophage lineage and $CD4^+$ T-cells, are termed R5 viruses or macrophage-tropic. Similarly, viruses which use the CXCR4 co-receptor, expressed mainly on $CD4^+$ T-cells, are termed X4 viruses or T-lymphotropic. The X4 viruses efficiently infect T-cells and T-cell lines but are unable to infect monocyte-derived macrophages (MDM). This is in contrast to R5 viruses that are able to infect cells of the monocyte/macrophage lineage and primary $CD4^+$ T-cells but are unable to infect T-cell lines. The cellular tropism of HIV-1 is largely governed by the interaction of the variable V3 region within the gp120 surface protein with the different co-receptors (25, 26, 32).

Figure 3: HIV-1 Life Cycle



After adsorption of the virus to the target cell receptor a conformational change of the TM protein initiates fusion of the viral envelope with the host cell membrane. Subsequently, the viral core is released into the cytoplasm where the viral structural proteins dissociate from the genome. A Lys-tRNA, which is attached to the 5' LTR of the genomic RNA, acts as a primer for the viral reverse transcriptase (RT). The RT first synthesizes one DNA strand while simultaneously RT's RNaseH activity degrades the RNA template. In a second round, RT synthesizes another copy of viral DNA using the first DNA copy as template, producing a double stranded DNA molecule termed provirus. Due to the error-prone nature of the viral RT, which lacks a proofreading function, a number of mutations may be introduced at this step. The provirus circularizes and complexes with p17MA, Vpr and the viral integrase (IN) to form the preintegration complex (PIC) that is transported into the nucleus. Subsequently, the IN catalyzes the intergration of the provirus into the host genome. This integration is an

undirected process in the sense that the location of the integration is site-independent and nearly random. The viral promoter, which is located at the 5' LTR and contains several recognition sites for cellular transcription factors, including the T-cell specific factors NF-κB, SP1 and others, drives transcription. The viral mRNA transcripts are synthesized by the host RNA polymerase II that requires the HIV-1 Tat protein for efficient elongation. Tat acts as a transactivator that by recruiting cellular factors and greatly enhances elongation of the transcribing RNA polymerase II complex. The viral mRNA transcript contains multiple splice sites that can be used in a complex way to yield alternate splice products. Translation of multiply spliced mRNA yields the viral gene products Tat and Rev. Rev binds to the Rev responsive element (RRE), which is a complex stem-loop structure located in the *env* gene. This interaction changes the splicing pattern, so singly spliced and unspliced viral mRNAs are allowed to be transported into the cytoplasm for translation. The *gag* gene, which encompasses the p17MA, p24CA and NC proteins, is translated as a p55gag polyprotein. A ribosome slippage region at the end of *gag* leads to the expression of a gag-pol polyprotein that includes all of *gag* fused to the PR, RT and IN regions about 10% of the time. The regulatory and accessory genes *vif*, *vpr*, *tat*, and *rev* are expressed, depending on the usage of several alternate splice sites. Only the accessory protein Nef is translated from a separate reading frame on the 3' end of the genome. The *env* genes are also translated as a polyprotein consisting of the vpu, p120 SU, and p41 TM regions and are cleaved by a cellular protease. This is in contrast to the gag and gag-pol polyproteins that are cleaved by the viral protease during virion assembly and maturation. The *env*

polyprotein is heavily glycosylated in the Golgi network before it reaches the cell membrane. Once all the viral proteins and the genomic transcripts are produced, the virion is assembled at the cell membrane. This step is mainly directed by the Gag polyprotein. The nucleocapsid (NC) domain of Gag interacts with a packaging signal, found near the 5' end of the full-length transcript, and directs the incorporation of genomic RNA into the virion. Polymerization of the MA protein domain of the gag-pol p160 activates the protease (PR) domain of Pol. During budding, the viral protease cleaves the viral polyproteins into a mature virion containing fully processed Gag (MA, CA, NC, p6, p2, p1) and the pol gene products (PR, RT, IN). In addition, the viral proteins Vif, Vpr, Tat and Nef are packaged into the maturing virus particle. The virion acquires the cellular lipid bilayer membrane containing oligomeric gp120 as well as various host cell-surface proteins during budding (134).

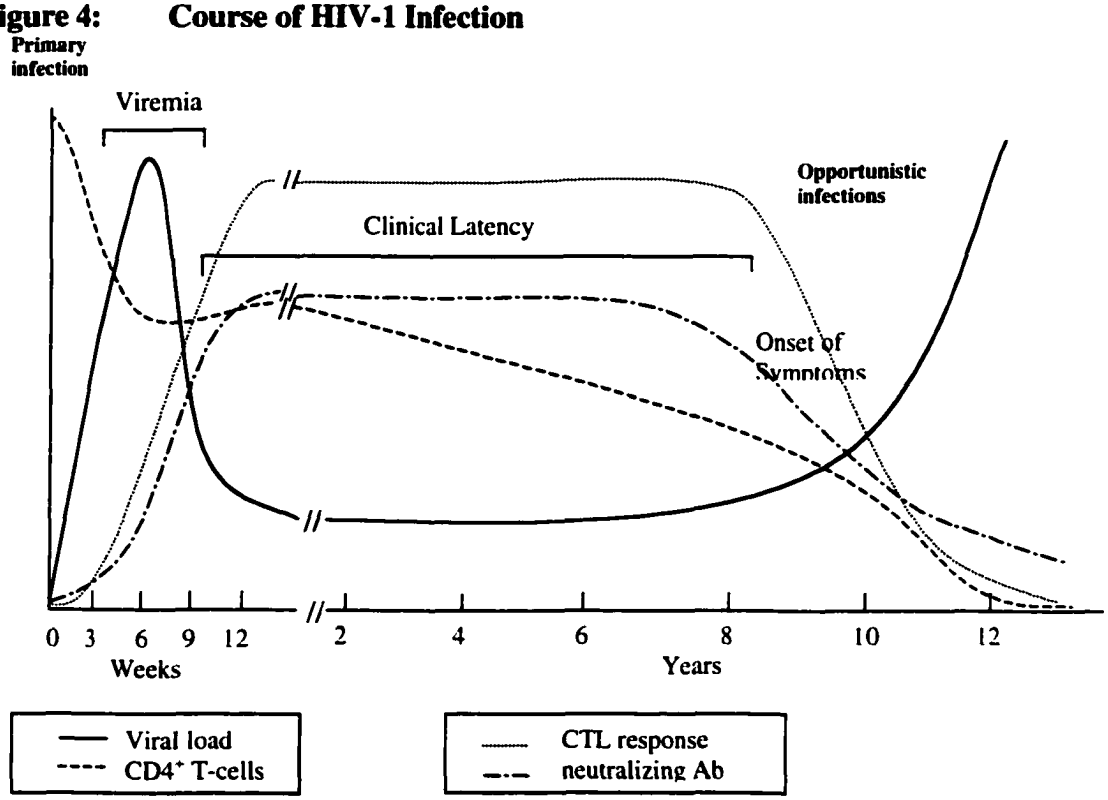
1.2 HIV-1 Infection

1.2.1 Adult HIV-1 Infection

Over the course of a typical HIV-1 infection, an adult individual goes through an initial systemic viremia at about three weeks after transmission. This high-grade viremia is accompanied by a short, very high peak of viral load that is associated with an "acute retroviral syndrome". The clinical presentation may vary greatly from individual to individual but typically involves symptoms that are reminiscent of mononucleosis. About nine weeks after infection the viremia declines and the infection may enter a clinically latent phase. The peak viral load is associated with the initial

rapid drop of $CD4^+$ T-cells (Fig. 4). After fading of the initial viremia, the $CD4^+$ T-cell count may stay nearly constant but eventually declines during the period of clinical latency that may last from one year to up to ten years. During this period the viral load increases gradually and the CD4 count decreases until it drops under a critical threshold. This is associated with a rapid increase of the viral load and the onset of symptoms of AIDS accompanied by opportunistic infections. During the clinically latent phase, an infected individual can be classified as either a slow progressor, a normal progressor or a fast progressor. Slow progressors display fairly stable $CD4^+$ cell counts, whereas intermediate progressors display gradually declining $CD4^+$ cell counts. Fast progressors are characterized by rapidly declining $CD4^+$ cell counts and development of symptoms of AIDS (184).

Figure 4: Course of HIV-1 Infection



1.2.2 Pediatric HIV-1 Infection

Perinatal HIV-1 infection generally follows a more rapid course than infection acquired in adulthood. The time frame is much longer in adults than in children - years compared to months. The clinical course of perinatal HIV-1 infection is highly variable from individual to individual but approximately 20% of infants experience a rapid progression to disease. Several studies have shown that by 18 to 24 months of age, nearly all perinatally infected children were found to have some degree of symptomatic AIDS (102, 189, 214) and more than one third develop full-blown AIDS within three years of age (146).

Neonates may test positive for HIV-1 either at birth or within the first 4 weeks of life. During this period of an infant's life, viral load rapidly rises and reaches levels comparable to the highest levels seen in adults in the initial burst of viremia (91). Virus copy numbers stabilize around $7*10^5$ /ml by age 2 month and stay constant for 9-12 months (1, 162). This is in contrast to adults where after the peak, the viral load significantly declines over a 2-12 month period after seroconversion (111).

1.3 Viral Genetic Diversity in HIV-1 Infection

1.3.1 Nature of Viral Diversity

HIV-1 is characterized by a very high degree of viral replication and genetic variability that results in an extensive viral heterogeneity in a given host. This extreme diversity within a host leads to the concept of viral quasispecies (51), which is defined as a large and complex population of genetically diverse but related viral genomes. The

development of quasispecies is the result of the high mutation rate (47, 94) during reverse transcription by the viral RT as well as during synthesis of genomic RNA by the host RNA polymerase II. Neither of these enzymes exhibit 3' exonuclease proofreading activity, and each therefore displays high rates of nucleotide misincorporation. In addition, the RNA genome has a spontaneous mutation rate due to the deamination of C to U residues. The rate of misincorporations has been estimated around $3x10^{-5}$ errors per base per replication cycle (138), a rate that is five to six orders of magnitude higher than the average error rate of a eukaryotic nuclear gene. Others estimate the *in vivo* mutation rate of HIV-1 in the 10^{-2}-10^{-3} error range per average nucleotide position per year with much higher mutation rates in the hypervariable regions of the viral genome (240). This high error rate and the high viral turnover lead to the accumulation of mutations in the viral genome. With more than 10^9 viral particles made every day, it is estimated that every single point mutation possible in the viral genome occurs between 10^4-10^5 times per day in an infected individual (36, 164).

One of the most important differences between a species and quasispecies is that each individual genome has only a momentary existence within the swarm of mutant quasispecies. Also the population of genomes, not the individual genome, is the object of natural selection. The evolutionary properties of a quasispecies may explain the dynamics of genetic variation and the biological diversity of HIV-1. One of the main features of a quasispecies is its genetic instability. At any given time, the genome of a variant population may be at equilibrium in respect to the selective forces, but this equilibrium is inherently unstable and may shift rapidly if a more advantageous mutant

variant appears. Because of the large population size of HIV-1 genomes within one host, random events may cause fluctuations in the distribution of variant genomes. New variants can arise from one or few advantageous variants and may have profound biological effects. Phenotypic changes may include modifications in cell tropism, virulence, and the formation of escape mutants. Many studies over the last decade recognized patterns of mutations in the HIV-1 genome. The main findings were: 1) genetic changes result mainly from point mutations and only to lesser extent from insertions or deletions; 2) sequence variation is not uniform along the genome, 3) genetic diversity within a host varies in time and space, and 4) viral genomes from cultured specimens do not reflect the genetic changes *in-vivo*.

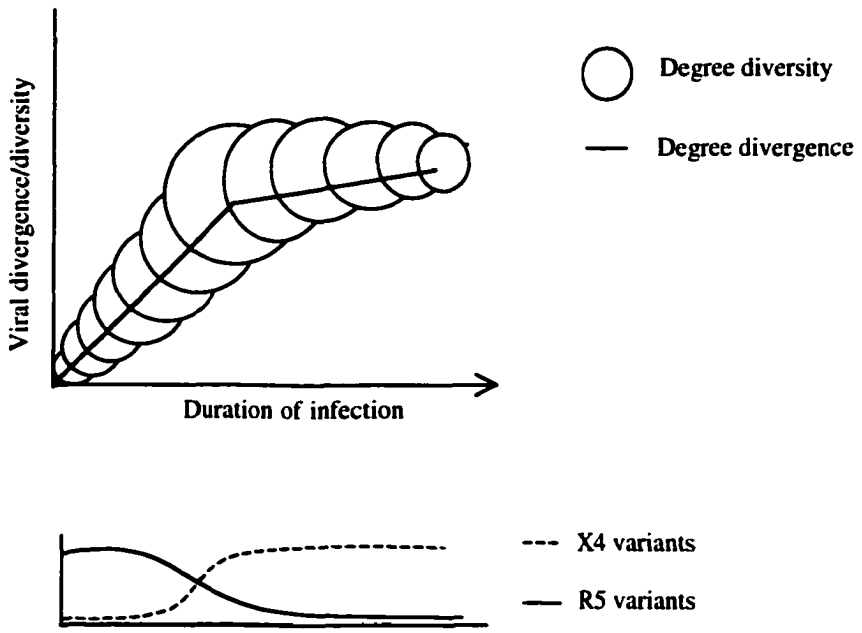
Therefore, the viral population in a host displays high genetic diversity and continues to evolve during asymptomatic, persistent infection. In both asymptomatic adults and infants, the accumulation of divergent phenotypes results in a very complex viral population of distinct quasispecies and small advantageous changes may rapidly shift the properties of the population of viral genomes.

1.3.2 Genetic Variation and Disease Progression

Tracking of viral evolution over time within infected individuals has revealed an association between viral heterogeneity and disease progression. Research showed that disease progression can be correlated with viral genetic diversity within a host. Several studies have found an inverse relationship between the rate of viral diversification and disease progression (43, 44, 69, 131, 135). Longitudinal studies of large populations of

infected adults have revealed a pattern of viral evolution that can be summarized as follows (Fig 5): At first, in a newly infected individual, the viral population displays low genetic diversity, defined as the degree of genetic difference among the present viral genomes. The genetic divergence, defined as the evolutionary distance form the founding (transmitted) variant, is also low. In addition, it has been shown that the transmitted variants have macrophage-tropic and non-syncytium inducing phenotypes (R5 virus) (143, 178, 217). As the infection goes on, both the viral diversity and divergence increase. At a certain point during the asymptomatic infection, viral variants emerge that can be characterized as X4 or T-lymphotropic. As mentioned earlier, these variants utilize the CXCR4 co-receptor and display a more cytopathic phenotype in comparison to macrophage-tropic variants. The appearance of X4 variants coincides with a decline of the rate of diversification and with an increase of the number of different X4 variants (Fig. 5). As the abundance of different X4 variants peaks, the rate of divergence from the founder variant decreases. In essence, a limited number of different X4 variants (110, 191, 192) outcompetes most other variants that differ in tropism and cytopathicity. The decrease in diversity and divergence as well as the outgrowth of X4 variants is coincidental with the onset of symptomatic AIDS.

Figure 5: Diversity, Divergence and X4 or R5 Variants during the Course of HIV-1 Infection



1.4 Vertical Transmission

Transmission of human immunodeficiency virus type 1 (HIV-1) from mother-to-infant is the primary mode of pediatric HIV-1 infection accounting for 90% of all HIV-1 infections in children. Since close to 40% of new cases of HIV-1 infection are among women of the childbearing age group, there is an increased risk of vertical transmission. During the year 2001, 2.7 million children were living with HIV/AIDS worldwide. It is estimated that 1,600 HIV-infected infants are born daily, accumulating close to 600,000 new infections each year worldwide. Due to the high death rates without antiretroviral treatment of the infant, generally not available in developing countries, an estimated 580,000 HIV-infected children die annually of AIDS related

complications. These deaths represent 19% of all worldwide HIV/AIDS related deaths (215). Since the beginning of the HIV-1 epidemic more than 15,000 infants have been born with HIV in the US. At the present time approximately 4,300 children are reported to The Centers for Disease Control and Prevention (CDC) (29) as living with HIV/AIDS.

In the US and Europe, rapid implementation of serologic testing, treatment of HIV-1 infected pregnant women, and changing of obstetrical practices have resulted in a reduction of transmission rates. However, serologic HIV screening of pregnant women is still not part of the routine prenatal care. Despite reduction of transmission rates in the US and Europe, the rate of maternal transmission remains high worldwide. The UNAIDS estimates that 90% of the world's perinatal infections have occurred in Sub-Saharan Africa. This is due to a combination of high seroprevalence - 30%-70% of pregnant women are HIV infected - high fertility, and poor access to information and health care services.

Perinatal HIV-1 infection and the mechanisms of transmission have become a major area of HIV-1 research in an effort to prevent pediatric HIV-1 infection. In addition, perinatal transmission provides a well-defined transmitter-recipient relationship that enables one to study the dynamics of HIV-1 infection.

1.4.1 Levels of Maternal Transmission and Associated Risks

Perinatal transmission can occur at three levels during gestation: prepartum (in-utero), intrapartum (during labor and delivery), and postpartum (through breastfeeding).

It has been very difficult to determine the exact timing of HIV-1 infection of infants, which has been closely tied to the technical capability of virus detection in newborns. Infants who are not breast-fed are generally considered infected in-utero if a sample taken 48h after delivery tests positive by PCR or virus culture (34). Accordingly, a child is classified as infected intrapartum, if the first sample that tests positive was taken later than one week. Using this classification, it is estimated that approximately 25-40% of infections occur in-utero and the remaining children are infected intrapartum. Without antiretroviral treatment (ART) of the infected mother, which is generally only available in developed countries, the vertical transmission rate is approximately 30% (2, 105, 166, 210). The overall risk for an infant to be in infected in-utero is approximately 5%-6% and the risk for intrapartum transmission is approximately 13%-18% (118, 151). Assessment of the risk of infection postpartum is more challenging, because breastfeeding may start soon after birth and the infant may be exposed to the virus continuously over several months. Therefore it is almost impossible to determine the timing of infection in these cases.

1.4.1.1 In-Utero Transmission

Detection of HIV-1 in fetal tissue of first and second trimester abortuses using *in-situ* hybridization, PCR, and virus culture provided the first direct evidence of in-utero infection. Most in-utero infected infants already have detectable HIV-1 in their peripheral blood at the time of delivery. This indicates that in-utero infection must have occurred at least several days or weeks before delivery. However, phylogenetic and

statistical studies showed that in-utero infection most likely occurs late in gestation (34, 108, 118).

Analysis of placentas of HIV-1 infected mothers showed that HIV-1 infected lymphocytes adhere to the chorionic villus stromal cells (186) and HIV-1 infected macrophage like cells were detected in full term placentas (145). Moreover, it has been demonstrated that HIV-1 can pass through the intact placental barrier (185). This indicates that in-utero transmission is most likely transplacental. The transplacental transmission may be enhanced by the increased presence of infected lymphocytes and macrophages due to inflammation of the placental membrane (203). Also, disruption of the placental membrane due to bacterial or fungal infection may facilitate the migration of the virus. However, HIV-1 may be maintained ex-vivo in the amniotic fluid (185) and some in-utero transmission may occur through the infant's mucosal membranes.

1.4.1.2 Intrapartum Transmission

Intrapartum transmission is due to contact of the infant's skin and mucous membranes with infected blood and maternal secretions during labor and delivery. The importance of intrapartum infection becomes apparent considering that only a minor proportion of infants is infected in-utero. Also, it has been shown that vaginal delivery and prolonged duration of ruptured membranes increase the risk of infection, which points to the intrapartum period.

1.4.1.3 Postpartum Transmission

Infants that are not infected in-utero or intrapartum may still be infected through breastfeeding since HIV-1 is commonly contained in breast milk of infected women. The infant is most likely infected by prolonged exposure of the oral and gastrointestinal mucous membranes to the virus. Breast feeding by HIV-1 infected women is not recommended in developed countries. However, in developing countries these women should still breast feed since the morbidity and mortality of malnutrion and diarrhea-associated complications with formula feeding outweigh the risk of HIV-1 infection through this route.

1.4.2 Factors Associated with Vertical Transmission

Despite substantial research, the determinants of perinatal transmission remain poorly understood. Maternal factors, fetal factors, mode of delivery, complications during labor and delivery, and viral factors play important roles in perinatal transmission. Most of these factors are interrelated such as maternal disease status, viral load and $CD4^+$ cell count. Nevertheless, the associated factors broadly fall in four categories: 1) fetal exposure, 2) maternal host factors, 3) fetal host factors, and 4) genetotypic and phenotypic characteristics of the virus.

1.4.2.1 Fetal Exposure

The risks of fetal exposure differ greatly for in-utero and intrapartum transmission. The risks of in-utero transmission are largely determined by maternal factors whereas the risks of intrapartum transmission are dependend on the mode of delivery. In-utero transmission may occur transplacentally or by ingestion of HIV-1

contaminated amniotic fluid. The fetus may be exposed to HIV-1 due to disruption of cervical, vaginal or placental tissues (90). This may be the case in chorioamnionitis or concurrent sexually transmitted diseases (STD). This not only decreases the placental barriers but may also increase local viral replication and therefore local viral load.

The likelihood of intrapartum transmission is increased by prolonged exposure of the infant to maternal vaginal and cervical secretions. This is especially apparent in cases of prolonged exposure due to ruptured membranes.

1.4.2.2 Host Factors of the Mother

The risk of vertical transmission correlates with several maternal factors: 1) maternal disease status, which includes low maternal $CD4^+$ lymphocyte count, and high viral load, 2) opportunistic infections of the mother and 3) anti retroviral therapy of the mother (17, 24, 40, 147, 166, 202).

The rates of vertical transmission fluctuate with the symptoms and severity of the disease in the mother. It is estimated that the transmission rate in symptom-free mothers is 24% but 65% in mothers with disease and/or who have had a previous child with AIDS (52, 105, 152, 175, 188). Several studies could also correlate vertical transmission with low maternal $CD4^+$ lymphocyte count. With a $CD4^+$ cell count of over 600 the risk of transmission is 15% but increases to 43% with a $CD4^+$ cell count of under 200 (147). In addition, high maternal viral load may be important in perinatal transmission because it provides a higher density of infectious virions that increases the likelihood of transmission. Although a correlation between high maternal viral load and transmission has been shown (9, 24, 211), it does not account for all cases of perinatal

infection. Several other studies have shown association of vertical transmission with low maternal viral load or no vertical transmission with very high maternal viral load thus providing evidence for a lack of correlation of transmission and viral load (88, 159). Therefore, maternal viral load alone does not fully explain vertical transmission and cannot be used as prognostic marker.

A clear correlation of increased risk of perinatal transmission and concurrent STDs could be established and is a result of viral, bacterial, and/or fungal infections as well as chorioamnionitis (190). STDs such as chlamydia and gonorrhea, cervical or vaginal ulcers increase cervical-vaginal shedding of HIV-1 (73). Viral shedding in genital secretions may be more relevant to transmission than the plasma viral load of the mother. This is apparent since most infants are infected in the intrapartum phase and not in-utero. On a further note, several studies have shown that the viral genetic variants in genital secretions differ from the variants found in peripheral blood (161, 168, 248). This may be important to vertical transmission, since the infant comes in contact with these viral populations first. Although the role of the placenta in transmission is not fully understood, vertical transmission with cell-free virus may occur due to contact of maternal blood with trophoblast or terminal villi cells. Both of these cell types have CD4 surface receptors and therefore may be susceptible to infection by HIV-1. Ex-vivo studies have shown that it is possible for HIV-1 to cross over intact placental tissue (15), however disruption of the placental membranes may greatly facilitate transfer of HIV-1 (74, 203).

Antiretroviral therapy (ART) administerd to the mother during pregnancy reduces the rate of maternal transmission of HIV-1, as the zidovudine (ZDV) trials demonstrated (202). The mechanism of reduction of transmission by ZDV monotherapy is not fully understood. However, the decrease of viral load in the mother cannot fully explain the reduction of transmission, considering that ZDV monotherapy only achieves a modest viral suppression. The AIDS Clinical Trials Group (ACTG) study revealed that the decrease of viral load by ZDV only accounts for a 17% decrease in transmission (202). One of the possible explanations may be a protective effect of ZDV in fetal tissues similar to the protecting ZDV therapy in health care and research personnel after HIV-1 exposure. However, another explanation may be that since ZDV reduces viral replication, it not only lowers viral load but also reduces viral heterogenity due to the reduction of active viral replication. This decrease of viral heterogeneity may reduce perinatal transmission. This is indicated by a study involving HIV-1 infected mothers that failed to transmitt the virus who displayed low viral hetereogeneity in the absence of ART (142).

Although ART during the pregnancy helped to reduce the frequency of perinatal transmission, a complication of maternal therapy is the fact that ZDV selects for the transmission of ZDV-resistant variants (39, 197). The transmission of drug resistant variants may potentially detrimental to the prognosis for the infants.

1.4.2.3 Host Factors of the Infant

After the discovery of the two required co-receptors for HIV-1, CCR5 and CXCR4, and their role in adult transmission, it was believed that the genetic make up of

the infant in respect to the co-receptors may be important in perinatal transmission. Initial reports showed that a 32 bp deletion in the CCR5 co-receptor may be protective in adult transmission to the infection with macrophage-tropic variants (96, 180). Currently, studies failed to find a correlation between the CCR5 mutation and prevention of perinatal transmission (50, 172). Moreover, a recent study indicates that infants who display the CCR5 mutation may be infected with multiple variants of both the R5 and X4 type. Thus, the CCR5 32 bp deletion does not protect from vertical HIV-1 infection, the contrary it may lead to infection with the more cytopathic X4 variant (178).

1.4.2.4 Genotypic and Phenotypic Characteristics of HIV-1 Involved in Transmission

The clinical parameters of infected mothers alone are insufficient to distinguish between transmitting mothers and non-transmitting mothers. Furthermore, infant host factors seem not to significantly modulate transmission. Therefore, host factors, maternal or fetal, fail to explain why more than half of HIV-1 infected mothers do not transmit the virus to their infants.

As mentioned earlier, HIV-1 is characterized by a very high degree of viral replication that results in an extensive viral heterogeneity in a given host. The reason for this genetic variability is the error-prone nature of the reverse transcriptase as well as mutations that are introduced into viral RNA transcripts. Newly infected individuals, adult or infant, generally display homogenous viral populations. This has been most extensively studied in the *env* gene. Several studies in adults have shown that the viral

populations are homogenous and that the viral variants are of the macrophage-tropic and non-syncytium inducing (NSI) type (R5 virus) (33, 58, 148). To determine if this is also true for perinatal transmission, several genetic characterizations (3, 183, 231) of HIV-1 envelope V3 regions from mother-infant isolates following perinatal transmission have revealed that the minor maternal genotypes are transmitted to the infant. In addition, it has been shown that the transmitted HIV-1 variant is of the macrophage tropic and NSI type (R5 virus) (143, 178, 183, 217). Selective transmission of the transmitter's minor subtype has also been shown during sexual transmission of HIV-1 (217, 229, 247). Moreover, transmission of the minor macrophage tropic variant was demonstrated in the presence of a major X4 variant (229). However, other studies have shown transmission of major, minor, and multiple variants, that may differ by the timing of transmission (46, 123, 163). Additionally, it has been demonstrated that *env* V3 region sequences from mothers who failed to transmit the virus in the absence of antiretroviral therapy (ART) are more homogenous in comparison to *env* V3 region sequences from transmitting mothers (142).

Several models have been suggested to explain the selective transmission of HIV-1. The first model predicts that the recipient is infected by a single or few variants because the inoculum is very small. The second model suggests that multiple variants are transmitted but only one is selectively amplified in the recipient because the particular variant has advantageous biological properties. The third model predicts that certain variants are predominantly transmitted due to their biological properties. For instance, a variant may have a selective advantage to infect cells at the point of

transmission. However, a combination of the second model (selective amplification) and the third model (selective transmission) may be most likely. One explanation of the predomination of macrophage-tropic variants in fetal recipients could be that maternal R5 strains replicating in the genital mucosa may preferentially infect fetal Langerhans cells in mucocutaneous fetal tissue. In addition, these variants may be able to replicate more efficiently in these macrophage lineage cells and therefore be selectively amplified. Then the amplified variants would establish the infection in the infant. Therefore, one should study what variants are transmitted in respect to selective transmission and establishment of infection in the fetal host.

1.4.2.5 Genetic Characterization of Other HIV-1 Genome Regions Involved in Perinatal Transmission

Several regions of the HIV-1 genome may be involved in vertical transmission but only limited information has been available. In this context, our laboratory has analyzed several HIV-1 genes, including *vif* (239), *vpr* (238), *vpu* (236) and *tat* (101). These studies have indicated that the reading frames of *vif*, *vpr*, *vpu* and *tat* are highly conserved in isolates from infected mothers and their infants. In addition, several motifs in the respective genes were identified to be associated with perinatal transmission. Furthermore, the functional domains of the *vif* and *vpr* genes were less conserved in non-transmitting mothers' isolates compared with transmitting mothers' isolates, indicating a correlation between low conservation of these genes and lack of vertical transmission (237). Moreover, the *vif* and *vpr* sequences from the non-transmitting mothers were more homogenous compared with *vif* and *vpr* sequences

from transmitting mothers, confirming similar results described for the *env* V3 region (142). In the analyses of the *vpu* and *tat* sequences, it was found that these open reading frames were highly conserved in mother-infant isolates (101, 236).

While perinatal transmission is multifactorial and highly complex, elucidation of the molecular mechanisms and identification of viral determinants involved in maternal transmission may advance the development of strategies for prevention of perinatal transmission and treatment of HIV-1 in infants. The possibility exists that several other regions of the HIV-1 genome may be involved in mother-infant transmission. The two structural genes *gag* p17 and gp120 as well as the accessory gene *nef* are important factors for viral replication. The gag p17 matrix protein is essential for the virus because it is involved in the targeting of Gag to the plasma membrane (20, 75, 243, 246), virus assembly and release (63, 75), envelope glycoprotein incorporation into virus particles (60, 242), virus entry (241) and localization of the virus preintegration complex to the nucleus of nondividing cells (21, 67, 220). The env gp120 is functionally key to virus infectivity (226), virus neutralization (107, 124, 174), replication efficiency (209), and host cellular tropism. In addition, the accessory gene *nef* is important for viral pathogenesis and its functions include the downregulation of CD4 (5, 136) and MHC class I (125, 187), stimulation of virion infectivity (82, 106) and alteration of the activation state of cells by interaction with cellular kinases (14, 177). Therefore, the characterization of the HIV-1 *gag* p17, *nef* and the *env* gp120 from mother-infant pairs as well as non-transmitting mothers may provide a better understanding of the molecular characteristics of the transmitted viruses.

1.5 Gag p17MA Matrix Protein

The *gag* p17 matrix (MA) protein is the major viral structural protein and gives the virion its overall spherical shape. In the mature virion, p17MA is a phosphorylated, N-terminally myristoylated (20, 75) protein that trimerizes and forms the spherical shell that lines the inner surface of the host-cell derived lipid membrane. The protein is expressed as the N-terminal domain of the $Pr55^{Gag}$ precursor and it is cleaved off during virion maturation by the viral protease. The general organization of the Gag precursor and its proteins, including MA, is highly conserved among the retroviruses (228). The *gag* p17 matrix protein plays a pivotal role in the life cycle of HIV-1 and is involved in several important events that are spatially and chronologically distinct. The HIV-1 p17MA protein has been shown to have four main functions: 1) targeting and binding of the Gag precursor to the plasma membrane (243, 246), 2) incorporation of the envelope glycoproteins into new virions (48, 60, 242) 3) direction of virion assembly and release (31, 63, 201), and 4) involvement in post-entry events (169, 241).

1.5.1 Gag p17 Matrix Protein Functions

In general, Gag proteins need to fulfill two opposing functions. During virus assembly they must assemble into a multimeric structure that is directed toward the cell-membrane for assembly. In contrast, following viral entry, they must disassemble and be directed toward the nucleus for integration. The problem of these apparently opposing properties is solved for p17MA by 1) expression of p17MA as part of the Gag precursor whose regions undergo conformational changes after PR processing as well as 2) posttranslational modifications in form of the addition of a myristic acid to p17MA.

The structure of the HIV-1 MA protein has been resolved (92, 140) and revealed that the protein folds into a compact core domain, consisting of five α-helices and a three stranded β-sheet (Fig. 6).

1.5.2 Membrane Binding and Gag Targeting

Plasma membrane localization of the Gag precursor is essential for correct assembly as well as adequate incorporation of the viral envelope proteins into new virions and therefore indispensable for the viral lifecycle. Binding to the membrane involves several domains of the protein. The p17 matrix is a membrane bound protein and membrane localization requires the cotranslational myristoylation of a N-terminal Glycine as well as a cluster of basic amino acids around residues 17-31 (243, 246). Experiments have shown that p17MA itself binds less efficiently to the plasma membrane as compared to the Gag precursor. This, together with the finding that certain deletion mutations throughout the protein enhance membrane binding of MA, has led to the idea of a myristic switch mechanism that has also been described for other proteins. In the Gag precursor the myristic acid moiety is highly exposed and ready to bind to the plasma membrane. In contrast, during maturation MA undergoes a conformational change that partly conceals the myristic acid moiety so MA can dissociate from the membrane during uncoating in the early steps of the viral lifecycle. In addition, this process is modulated by phosphorylation of several phosphorylation sites by cellular kinases (positions 110-114 and position 132). The introduction of negative charges into p17MA facilitates membrane dissociation due to the negatively charged plasma membrane.

1.5.3 Env Glycoprotein Incorporation

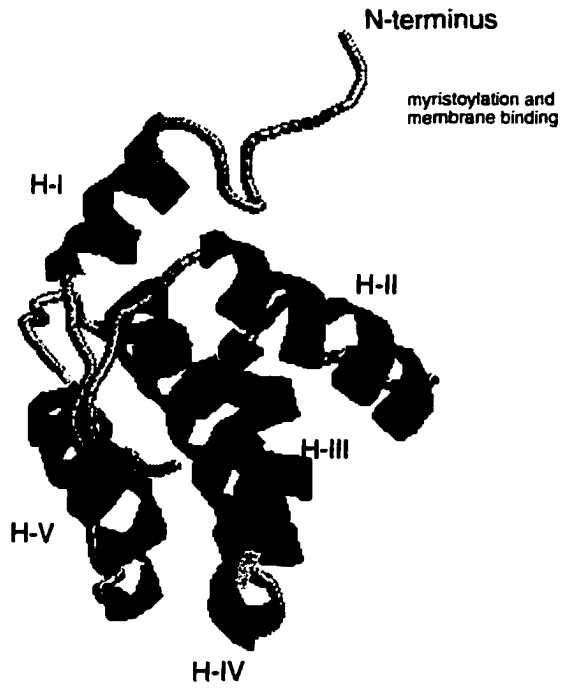
The incorporation of the viral envelope glycoproteins gp41(TM) and gp120SU is mediated by the Gag precursor (95). The close association of p17MA with the membrane in the mature particle implicates p17MA in the process of specific envelope incorporation. Indeed, the domains required for this interaction in p17 and gp41 have been identified (60, 242) as well as domains and amino acid residues that are essential for incorporation of gp120 into virus particles (63). However, much of the data involving *env* incorporation was obtained by *in-vitro* studies only and have been difficult to confirm by *in-vivo* experiments. One reason for the difficulties could be that *env* incorporation is dependent on the holes present in the lattice-like MA structure formed by oligomeric MA in the mature virion. This idea is supported by studies that introduced point mutations that disrupt multimerization of MA as well as *env* incorporation (60, 62).

1.5.4 Early Post-entry Steps

The p17 matrix protein is involved in the early stages of HIV-1 infection such as viral DNA synthesis, uncoating, and penetration (169, 241), suggesting a role in virus transmission. The close association of MA with the core structures implicates a role of MA after membrane fusion. Mutations in MA that do not impair assembly, virion release or *env* incorporation have been described that reduce the amount of viral DNA synthesis (27, 170, 241). In contrast to onco-retroviruses, which require that the host-cell undergoes mitosis to access the nucleus, lentiviruses may infect non-dividing cells efficiently. Moreover, *in-vitro* infection of macrophages with HIV-1 has been shown to

require one of the two independent routes of nuclear import, either through the nuclear import porperties of Vpr or the p17MA nuclear localization signal (NLS) (89). This is biologically important since it enables HIV-1 to infect terminally differentiated monocyte derived macrophages that are a major target for HIV-1 infection (71). Since HIV-1 in sexual (246) and vertical (143) transmission seems to be macrophage-tropic, p17 matrix should have a role in HIV-1 infection following mother-to-infant transmission. The presence of p17MA in the viral preintegration complex (PIC) that functions in translocation of the viral DNA into the nucleus, implicates a function of p17MA in this step. Based on mutational analysis, several studies (21, 220) have identified a nuclear localization signal (NLS) at the N-terminus of p17 that is required to import the preintegration complex in the nucleus of the nondividing cells that may be critical in the establishment of HIV-1 infection in quiescent T-cells and macrophages (21, 66, 150, 220), although this issue has been challenged (59).

Figure 6: p17MA Protein Structure



1.6 Nef Protein

Nef is a highly conserved accessory protein encoded by HIV-1 and is a critical virulence factor for the primate lentiviruses. Nef is incorporated into the virion as well as abundantly produced during the early phase of infection when its mRNA represents three quarters of the viral mRNA load of the cell (83, 114). Several studies characterized long term non-progressors that harbored viral variants with attenuated *nef* genes (42, 126, 139) suggesting that Nef is required for HIV-1 pathogenesis. Furthermore, in a transgenic mouse model, *nef* expressed in CD4 positive cells was the sole cause of pathogenesis (86). The importance of *nef* as a virulence factor is also

underscored by the observation that *nef*-deleted SIV variants lead to low viral loads in monkeys who also show a lack of disease progression (112).

1.6.1 Nef Functions

Nef is a phosphorylated (234) 206 aminoacid protein (~24 kDa) that is, due to a N-terminal myristoylation (158), associated with cellular membranes. Nef is also phosphorylated at serine and threonine (234) residues and forms homomeric oligomers and intramolecular disulfide bonds (83). Since the determination of the three dimensional structure of Nef, major progress has been made in understanding Nef (Fig. 7). The major findings were: Nef complexes with the SH3 domains of cellular tyrosine kinases (10, 129) and Nef binds to the cytoplasmic tail of CD4 (81). Extensive research uncovered three major *nef* functions: 1) downregulation of CD4 and MHC class I, 2) stimulation of virion infectivity and 3) alteration of the activation state of cells.

1.6.2 CD4 Downmodulation

Nef mediated down modulation of CD4 is accomplished by rerouting expressed CD4 at three levels. First, due to a N-terminal myristoylation, Nef is targeted to its primary site of action, the cell membrane. Here it binds to the cytoplasmic tail of CD4 and recruits cellular adaptor complexes that lead to the premature internalization of CD4 by endocytosis (5, 57, 136). Then, Nef targets CD4 from the endosome to the lysosome and is therefore responsible for CD4 degradation (113, 165). Apparently Nef acts as a connector between target molecules and the cellular protein trafficking machinery. This is best illustrated by the recognition of a dileucine motif in the cytoplasmic tail of CD4 by Nef and the subsequent recruitment of downstream cellular

components. One of the most important downstream component is the clathrin coated pit adaptor protein complex (AP). This complex usually directs clathrin to the cytoplasmic tail of surface receptors that contain an endocytosis signal. Thus, Nef links surface bound CD4 with an endocytosis signal targeting CD4 for degradation. This is demonstrated by *nef* mutants that lack the ability to internalize surface CD4, due to a defective AP binding region. Nef also interacts with the COP-I coatomer that is involved in the ER-Golgi and endosomal sorting. Again, Nef mutants defective in the COP-I binding domain are unable to direct CD4 to lysosomal compartments. Lastly, Nef redirects newly synthesized CD4 from the trans-Golgi network to the endosomal compartments and thus to degradation (136, 171). The biological significance of CD4 down modulation becomes apparent if one considers the affinity of *env* gp120 to the CD4 receptor. First, the lack of CD4 on the cellular surface prevents potentially dangerous superinfection. More importantly however, high levels of surface CD4 would hinder proper *env* incorporation during virion assembly (122) and may even prevent virion release (173).

1.6.3 MHC-I Downmodulation

The downregulation of MHC class I is less well understood than the downregulation of CD4, but generally follows a similar principle. Interaction of cytoplasmic Nef with surface expressed MHC-I molecules leads to its internalization and targeting to the trans-golgi and ultimately to its degradation (125, 187). However, the motifs and residues in Nef responsible for this interaction are independent from the motifs in Nef responsible for CD4 downregulation. In addition, phosphorylation of the

CD4 receptor does not seem to be necessary. Also, the AP recruiting motif in Nef is completely dispensable for MHC-I downregulation. However, regions within Nef that were implicated in alteration of cell signaling pathways seem to be involved. This suggests that Nef affects MHC-I downregulation -as with CD4 downregulation- with the help of cellular factors that remain to be identified. Whereas the mechanism is poorly understood, the biological significance of MHC I downregulation is apparent. Since MHC I presents intracellular antigenic peptides to effector cells of the immune system, the downregulation of MHC-I is a way for the virus to circumvent immune surveillance by cytotoxic T-cells (CTL) (41). This is further underscored by the fact that Nef only down regulates HLA-A and HLA-B but not HLA-C. Missing HLA-C on the cell surface would trigger destruction of the infected cell by Natural Killer cells (NK cells) (38).

1.6.4 Cellular Activation

Nef interacts with a multitude of host factors that are part of the cellular signaling machinery. A conserved PxxP motif lets Nef directly bind to SH3 domains of cellular kinases, such as Hck and Lyn (129, 177). In addition, Nef interacts with the T-cell specific Lck tyrosine kinase, that is involved in T-cell activation (14). Moreover, it has been reported that Nef may interact with members of other major signaling pathways such as the p21 activated kinases (PAK) (182), protein kinase C (199), and the rho-family GTPase guanine nucleotide exchange factor Vav (54). Several studies point towards the idea that Nef sensitizes T-cells to activation via the T-cell receptor (TCR). This state of hyperresponsiveness may enable peripheral blood mononuclear

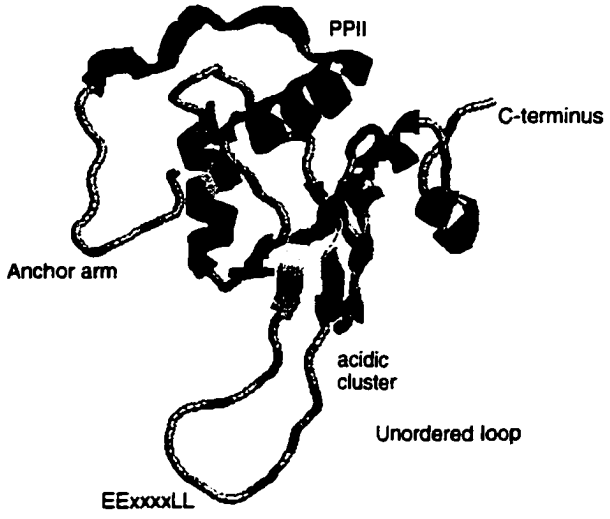
cells (PBMC) to replicate without exogenous stimulation and induces IL-2 production. Therefore Nef sensitizes T-cells to be activated by the TCR and makes them more susceptible to viral replication. Although less is known about the specific aspects of these interactions, binding of Nef to the TCR ζ may be another mechanism of immune evasion. Nef TCR ζ interactions results in the upregulation of FAS ligand, a part of the effector repertoire of CTLs that triggers apoptosis. So CTLs that come in contact with HIV infected cells are killed (41, 93, 232, 233). Apart from its influence on T-cells, Nef can promote the production of macrophage inflammatory proteins (MIP) in HIV-1 infected macrophages. MIPs are chemokines that act as chemotactic factors for T-cells and activate resting T-cells. Therefore, Nef may promote the infection of T-cells after an established infection in macrophages (207).

1.6.5 Virion Infectivity

Another key function of Nef is its ability to enhance virion infectivity. Virions with defective Nef are 3-10 times less infective than wildtype virions. This decrease in infectivity could be an effect of faulty virion assembly. Mutated *nef* that is defective in CD4 downregulation, may inhibit proper *env* incorporation. However, there seem to be additional CD4-independent mechanisms at play. One of these mechanisms may be the Nef mediated incorporation of cellular kinases of the MAPK family into the maturing virion (106). These kinases may phosphorylate the viral matrix protein (p17MA) which is involved in processes preceding proviral integration (206). It has also been reported that *nef* mutant viruses have decreased levels of viral DNA synthesis after infection (82). This effect can be overcome if wildtype Nef is supplied in trans to virus

producing cells but not in target cells, suggesting that Nef is needed in virion maturation and assembly.

Figure 7: Nef Protein Structure

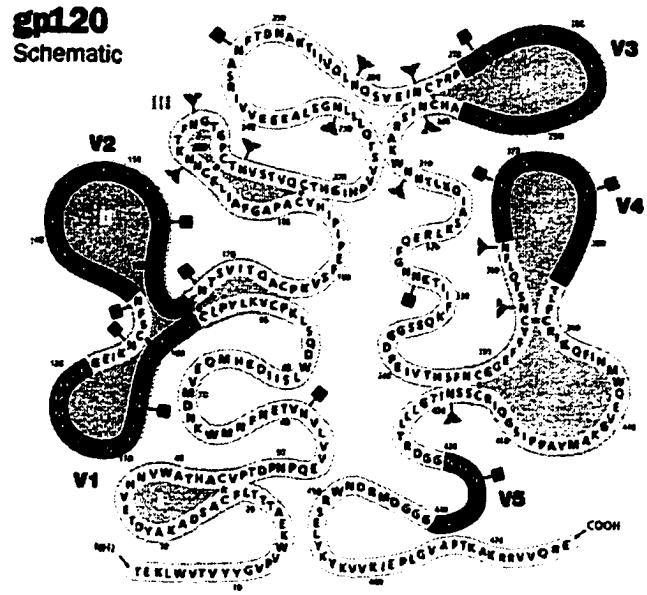


1.7 Env gp120 Surface Protein

The Env surface protein (SU) is synthesized as part of the gp160 precursor polyprotein and is postranslationally glycosylated. Approximately half of molecular mass of gp120 is composed of oligosaccharides. The gp160 protein is cleaved into the surface subunit gp120(SU) and the transmembrane subunit gp41(TM) by a cellular protease (134) in the Golgi during transport to the cell-surface. The two subunits stay non-covalently connected and trimerize on the surface of the virion. gp120 is the viral attachment protein and binds to the CD4 receptor as well as one of the co-receptors on the target cell. Sequence analysis of gp120 showed that gp120 is highly variable between viral isolates and that this variability is not uniform across the protein. This

lead to the designation of five variable regions (V1 through V5) which are interspersed with more constant regions (C1 through C5) (205). In addition, the protein forms five cystein loops (Fig. 8).

Figure 8: gp120 Surface Protein Structure



1.7.1 gp120 Protein Functions

In addition to binding to the main cell surface receptor CD4, gp120 also determines cell-type specificity or tropism. As mentioned earlier, HIV-1 either infects cells of the monocyte/macrophage lineage or T-cells. This specificity is dependent on the presence of the co-receptors CCR5 and CXCR4 for macrophage tropism and T-cell tropism respectively. Molecular epidemiological studies showed that the third variable region (V3) of gp120 is the most variable and largely governs this cell specificity. Analysis revealed that the sequence of V3 is responsible for the interaction with the different co-receptors (26, 103, 117, 194). This has been well demonstrated by the

introduction of the V3 sequences from macrophage tropic clones into T-cell tropic clones, that rendered the chimeric viruses macrophage tropic (103, 143, 194, 223). Although the V3 region is functionally important in virus infectivity (226), virus neutralization (107, 124, 174), replication efficiency (209), and host cellular tropism, it is not the only region of gp120 that determines viral properties. Studies have demonstrated that regions adjacent to V3 greatly enhance the ability of chimeric viruses to infect macrophages (224). Changes in C2 that dramatically change infectivity, could be reversed by changes in C1 and V3 (225), and studies showed a functional interaction of V1/V2 and C4 (61). The V1-V2 regions influence replication efficiency in macrophages by affecting virus spread (213). In addition, analysis of chimeric viruses constructed from distinct HIV-1 phenotypes indicates that a functional interaction between V3 and the second conserved region is important for infectivity as well as syncytium formation and cell tropism (115, 204, 227). Early studies of cell specificity focused on discreet regions within gp120, mostly the V3 loop, however it becomes increasingly clear that cell tropism requires regions within gp120 adjacent and nonadjacent to each other.

1.8 Hypothesis and Objectives

Our hypothesis is that there are specific molecular and biological properties of HIV-1 that are critical determinants of perinatal transmission. Better characterization of HIV-1 involved in mother-infant transmission can provide relevant information for the development of strategies for prevention and treatment of pediatric HIV-1 infection,

because such strategies should be targeted at the properties of the transmitted virus. In addition, the study of mother-infant transmission of HIV-1 is a well-defined model of a transmitter-recipient relationship and thus is ideal to study the dynamics of HIV-1 infection in-vivo. The molecular analyses of HIV-1 isolates from transmitting mothers and comparison with isolates from non-transmitting mothers may provide useful information about the viral determinants involved in vertical transmission. Therefore, we propose to characterize the important HIV-1 regions *gag* p17 matrix, *nef* and *env* gp120. The objectives are: 1) To molecularly characterize *gag* p17 matrix gene isolates from mother-infant pairs as well as from mothers who failed to transmit the virus. 2) To compare *nef* gene isolates from transmitting mothers and their infants, and 3) To determine the biological properties, including replication efficiency, cellular tropism and cytopathic effects of HIV-1 *env* gp120 isolates from mother-infant pairs and non-transmitting mothers that were previously molecularly characterized in our laboratory.

These detailed molecular and biological characterizations of HIV-1 *gag* p17 matrix, *nef* and *env* gp120 isolates from mother-infant pairs and non-transmitting mothers may provide a better understanding of the molecular mechanisms of perinatal transmission. This may enable AIDS researchers to develop new strategies for treatment and prevention of HIV-1 infection by means of new antivirals and preventative vaccines.

2 MATERIALS AND METHODS

2.1 Patient Population and Sample Collection

This study was approved by the Human Subjects Committee of the University of Arizona and the Institutional Review Board of the Children's Hospital Medical Center, Cincinnati, Ohio, and written informed consent was obtained from the study participants. The transmitter studies were conducted on eight mother-infant pairs including mother-H who had twins (pairs A, B, C, D, E, F, G, H). The non-transmitter studies were conducted on four HIV-1-infected asymptomatic mothers (Mnt, mothers who failed to transmit the virus to their infants), including a mother with multiple deliveries. Three mothers (Mnt-1, Mnt-2 and Mnt-3) did not receive any antiretroviral therapy, whereas one mother (Mnt-4) was administered zidovudine (ZDV) as per ACTG protocol. Blood samples were collected from infected mothers following childbirth, including Mnt-3 from whom we collected three samples, Mnt 3.2 and Mnt 3.5 following second and fifth delivery, respectively and a follow-up sample, Mnt 3.5F after one year of the fifth delivery. Mother 3 has given birth to five HIV-1 negative infants. Blood samples from mother 4 (Mnt-4) were collected 4 months after the delivery. These samples were collected between 1990 and 1995 and the children of these mothers were evaluated until diagnosed uninfected as per CDC recommendations. The demographics of patients used in this study are summarized in Table 1 (A and B) and Table 2.

Table 1: Patient Demographic, Clinical, and Laboratory Parameters of HIV-1 infected Mother-Infant Pairs

A

Patient	Age	Sex	Race[a]	$CD4^+$ cells/mm^3	Length of infection[b]	Antiviral Drug	Clinical evaluation[c]
				Mothers			
A	36 yr	F	B	706		None	Asymptomatic
B	28 yr	F	B	509	11 mo	None	Asymptomatic
C	23 yr	F	W	818	1yr 6 mo	None	Asymptomatic
D	31 yr	F	W	480	2yr 6 mo	None	Asymptomatic
E	26 yr	F	B	395	2 yr	ZDV[d]	Symptomatic AIDS
F	23 yr	F	B	692	2yr 10 mo	None	Asymptomatic
G	23 yr	F	W	480	10 mo	None	Asymptomatic
H	33 yr	F	B	538	5 mo	None	Asymptomatic

B

Patient	Age	Sex	Race[a]	$CD4^+$ cells/mm^3	Length of infection[b]	Antiviral Drug	Clinical evaluation[c]
				Infants			
A	6 wks	F	B	2994	6 wks	None	Asymptomatic; P1A
B	4.75 mo	M	B	1942	4.75 mo	None	Asymptomatic; P1A
C	14 mo	F	W	772	14 mo	ZDV	Symptomatic AIDS; P2A, D1, 3, F
D	28 mo	M	W	46	28 mo	ddC[e]	Symptomatic AIDS; P2A, B, F; failed ZDV therapy
E	34 mo	M	B	588	34 mo	ZDV	Symptomatic AIDS; P2A
F	1 wk	M	B	2953	1 wk	ZDV	Asymptomatic; P1A
G	24 mo	F	W	4379	24 mo	ZDV	Asymptomatic; P1B
HT1	7 mo	F	B	3157	7 mo	ACTG152	Hepatosplenomegaly, lymphadenopathy
HT2	7 mo	F	B	2176	7 mo	ACTG152	Hepatosplenomegaly, lymphadenopathy

[a]B, black, W, white

[b]Length of infection: The closest time of infection that we could document was the first positive HIV-1 serology date or the first visit of the patient to the AIDS treatment center, to which all the HIV-1 positive patients were referred as soon as an HIV-1 test was positive. Therefore, these periods of time may not reflect the exact dates of infection.

[c]Evaluation for infants is based on CDC criteria (28)[d]ZDV, Zidovudine [e]ddC, Zalcitibine

Table 2: Patient Demographic, Clinical, and Laboratory Parameters of HIV-1 infected Non-transmitting Mothers

Patient	Age	$CD4^+$ lymphocytes per mm^3	Length of infection at the time of sampling[a]	Antiviral Therapy	Clinical evaluation[b]
Mnt 1	32 yr	845	6 mo	None	Asymptomatic
Mnt 2	20 yr	698	1 yr	None	Asymptomatic
Mnt 3.2	20 yr	739	1 yr & 10 mo	None	Asymptomatic
Mnt 3.5	24 yr	609	4 yr & 10 mo	None	Asymptomatic
Mnt 3.5F	25 yr	515	5 yr & 10 mo	None	Asymptomatic
Mnt 4	28	490	9 mo	ZDV[c]	Asymptomatic

Abbreviations: Mnt nontransmitting mother

[a)] Length of infection: The closest time of infection that we could document was the first positive HIV-1 serology date or the first visit of the patient to the AIDS treatment center, to which all the HIV-1 positive patients were referred as soon as an HIV-1 test was positive. Therefore, these periods of time may not reflect the exact dates of infection.

[b)] Clinical evaluation: Clinical evaluation was made on the basis o CDC criteria (28)

[c]ZDV, Zidovudine

2.2 Isolation of Peripheral Blood Mononuclear Cells

Primary monocytes-derived macrophages (MDM) and peripheral blood lymphocytes (PBL) were obtained from peripheral blood mononuclear cells (PBMC) on a single-step Ficoll-Paque procedure (Pharmacia-LKB) gradient from the whole blood of HIV-1 positive transmitting mothers and their infants, HIV-1 positive non-transmitting mothers, and HIV-1 negative normal donors. We collected blood from all donors with their consent and the study was approved by the Human Subjects Committee of the University of Arizona. The blood was spun for 45 min. at 450 RCF with the brake off. The PBMC were carefully aspirated from the interphase of the plasma and ficoll, washed three times with phosphate buffered saline (pH 7.4) containing no Ca^{++} or Mg^{++}, and resuspended in culture medium.

2.3 Isolation of PBMC DNA

DNA from the uncultured PBMC from HIV-1 infected individuals was isolated according to a modified procedure described by Oram et al. (160). Approximately 10^6 PBMCs were centrifuged at 12,000 RPM for 2 min and the cell pellet was resuspended in 0.5 ml of TNE buffer (0.5 M Tris-HCl, pH 7.5; 0.1 M NaCl, 1 mM EDTA). The suspension was treated with 0.50% sodiumdodecyl sulfate (SDS) and 10 ug/ml proteinase K (Boehringer) at 60°C for 3 hours, followed by several extractions with phenol and chloroform. The DNA was precipitated with ethanol and dissolved in 50-100 µl TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA), and sheared by repeated pipetting.

2.4 Polymerase Chain Reaction Amplification of HIV-1 Genes

A two step polymerase chain reaction (PCR) amplification, first with outer primers and then with nested or inner primers, was performed to detect the presence of HIV-1 in infected patients' PBMC. Equal amount of HIV-1 PBMC DNA (approximately 25 to 50 copies minimum) was used from each patient as determined by end point dilution (65, 69). The PCRs were performed according to the procedure of Ahmad et al.(3), in a 25 µl reaction mixture containing 2.5 µl of 10X buffer (25mM TAPS (tris-(hydroxymethyl)-methyl-amino propanesulfonic acid, sodium salt) pH 9.3, 50 mM KCl, 2mM MgCl2, 1mM 2-mercaptoethanol, 400 µM each of dATP, dCTP, dGPT, and TTP, 0.2 µM of each outer primers, and 2.5 units of different high-fidelity thermo stable DNA-polymerases according to the respective manufacturers specifications. The PCR products were analyzed by electrophoresis on a 1% agarose gel. Negative controls consisting of DNA from PBMC of sero-negative individuals were included in each set of reactions, which were negative in all the assays. Multiple (four to six) independent PCRs were performed to obtain clones that were then sequenced and analyzed. To avoid contamination, the samples, reagents, and first and second round PCR products were kept separately and dispensed in a separate room free from all laboratory used DNAs. We also included the known HIV-1 sequence of the molecular clone NL4-3 for PCR amplifications as a control to assess errors generated by UlTma DNA polymerase or TaKaRa LA *Taq* polymerase.

2.4.1 PCR Amplification of *gag* p17MA

The HIV-1 *gag* p17MA gene from infected patients' PBMC DNA was amplified using the following primer pairs: GAG1 (5'ACAGGGACTTGAAAGCGAAAGTA, 646 to 668, sense), GAG2 (5'TACATGTAGTTCCTGCTATGACA, 1491 to 1513, antisense) as outer primers, and GAG3 (5'TCTCTCGACGCAGGACTCGGCTT, 682 to 704, sense) and GAG4 (5'CAGTCATCTGGCCTGGTGCAATA, 1455 to 1477, antisense) as inner primers (all positions according to published HIV-1 sequences of NL4-3 (153)). The PCRs were performed according to the procedure described by Ahmad et al. (3) for the mother-infant pair study using 2.5 U of UlTma DNA polymerase (Perkin-Elmer Cetus, Norwalk, Conn.) in concordance with the manufacture's protocol. For the non-transmitting mother study TaKaRa LA *Taq* polymerase (TaKaRa Biomedicals, Shiga, Japan) was used. The reactions were carried out at 95°C for 30 sec, 50°C for 45 sec, and 72°C for 3 min for 35 cycles. After the first round of PCR, 1µl of the product was amplified for 35 cycles with the corresponding inner primers at 95°C for 30 sec, 55°C for 45 sec and 72°C for 3 min.

2.4.2 PCR Amplification of *nef*

The HIV-1 *nef* gene from infected patients' PBMC DNA was amplified using the following primer pairs: NEF1 (5'GTAGCTGAAGGGACAGATAGGGTTAT, NL4-3 position 8687 to 8712, sense) and NEF2 (5'GCACTCAAGGCAAGCTTTATTGAGGC, positions 9631 to 9605, antisense) as outer primers; and NEF3 (5'CGTCTAGAACATACCTAGAAGAATAAGACAGG, 8748 to 8768, sense), NEF4 (5'CGGATTCCGTCCCCAGCGGAAAGTCCCTTGTA,

9552 to 9429, antisense) were used as inner primers. The PCRs were performed according to the procedure described by Ahmad et al. (3) using 2.5 U of TaKaRa LA Taq polymerase (TaKaRa Biomedicals, Shiga, Japan) in concordance with the manufactures protocol. The first reaction was carried out at 94°C for 30 sec, 50°C for 45 sec, and 72°C for 90 sec for 35 cycles using the outer primer pair. After the first round of PCR, 1µl of the product was amplified for 35 cycles with the corresponding inner primers and the same concentrations of other PCR ingredients at 95°C for 30 sec, 55°C for 45 sec, and 72°C for 90 sec.

2.4.3 PCR Amplification of *env* gp120

The HIV-1 *env* gene from infected patients' PBMC DNA was amplified using the following primer pairs: Env20 (5'-GAAGTAGTATTGGTAAATGTGACAG, 6467 to 6491, sense), Env21 (5'-CCACTCTTCTCTTTGCCTTGGTGGG, 7706 to 7731, antisense), Env22 (5'-CGGATTTCTTAACATGTGGAAAAATGACATGGT, 6489 to 6522, sense), and Env23 (5'-GCTCTAGATTTATATAATTCACTTCTCCA ATTG, 7645 to 7678 antisense). The PCRs were performed according to the procedure described by Ahmad et al. (3) using 0.5 U of TaKaRa LA Taq polymerase (TaKaRa Biomedicals, Shiga, Japan) in concordance with the manufactures protocol. The reactions were carried out at 94°C for 45 sec. , 45°C for 30 sec., and 72°C for 3 min for 35 cycles. After the first round PCR, 1 ul of the product was amplified for 35 cycles with the corresponding inner primers at 94°C for 30 sec, 50°C for 45 sec, and 72°C for 3 minutes.

2.5 Cloning of HIV-1 Genes

2.5.1 Cloning of *gag* p17MA

The products from multiple independent PCRs (four to six) amplified by the inner primer pair GAG3/GAG4 were blunt ended by DNA polymerase I (Gibco-BRL, Gaithersburg, MD), kinased by T4 polynucleotide kinase (Gibco-BRL, Gaithersburg, MD), and cloned into the Sma I site of pGem 3Zf (+) vector (Promega Corp., Madison, WI). Individual bacterial colonies were screened for the presence of recombinants by restriction enzyme analysis of plasmid DNA. The clones with the correct size of inserts were selected for DNA preparation followed by nucleotide sequencing of 8 to 15 clones for each patient.

2.5.2 Cloning of *nef*

Most of the PCR products obtained from multiple independent PCRs (four to six) were directly cloned in the pCR2.1TOPO TA-cloning vector, following the manufacturers instruction for the TOPO-TA cloning Kit, Version K2 (Invitrogen, Carlsbad, CA). In some experiments, the PCR products were treated with DNA polymerase I (Gibco-BRL, Gaithersburg, MD, now a subdivision of Invitrogen) to make the DNA ends blunt and then kinased using T4 polynucleotide kinase (Gibco-BRL). Then, the DNA was cloned into the Sma I site of the pGEM 3Zf (+) vector (Promega Corp., Madison, WI.). Using either cloning procedure, individual bacterial colonies were screened for the presence of correct size of fragments by restriction enzyme digestion of recombinant plasmid DNA. The positive clones were selected and

propagated for DNA isolation followed by nucleotide sequencing of 6-18 clones from each patient.

2.5.3 Cloning of *env* gp120

The PCR products amplified by inner primer Env22/Env23 were blunt ended by DNA polymerase I (BRL, Gaithersburg, MD), kinased with T4 polynucleotide kinase (BRL, Gaithersburg, MD), and cloned into the Sma I site of pGem 3Zf (+) vector (Promega Corp., Madison, WI). Individual bacterial colonies were screened for the presence of recombinants by restriction enzyme analysis of plasmid DNA. Cloning and genetical characterization of the *env* region of mother-infant pairs as well as non-transmitting mothers was performed by Erik Matala, PhD (141).

2.6 DNA Isolation

DNA was isolated according to the miniprep alkaline lysis protocol outlined in Sambrook (179). Bacteria containing the desired plasmid were grown in 3ml of LB medium containing 100 µg/ml ampicillin (Sigma) overnight and pelleted in a microfuge tube. The pellet was then resuspended in 100 ul of solution I (25mM Tris-HCl pH 8.0, 10mM EDTA, 50mM glucose) followed by 200 ul of solution II (1% SDS, 0.2 M NaOH) and mixed by inversion. After a 5 min. incubation, 150 ul of solution III (3 M sodium acetate pH 4.6) was added and mixed by inversion. This was followed by the addition of 450 ul of phenol:chloroform, and mixed vigorously. After a 5 min. spin at 15,000 r.p.m. in an Eppendorf microfuge, the aqueous layer was removed and transferred to a fresh microfuge tube. The DNA was then precipitated from the

supernatant by the addition of 800 µl of ice cold ethanol and incubated at –20° C for 1 hour. The DNA is pelleted by spinning in a microfuge at 15,000 r.p.m. for 20 min. After removal of the supernatant the DNA pellet was air dried DNA and resuspended in 50 µl of Rnase-water.

2.7 DNA Sequencing

DNA was sequenced with the Sequenase sequencing kit v. 2.0 (U.S. Biochemical, Cleveland, Ohio) according to the Sanger method of di-deoxy termination described in the Sequenase user manual (181). Briefly, DNA templates are denatured in a volume of 200 µl containing 10% denaturing solution (2 M NaOH, 2 mM EDTA) for 45 min. at 37^0 C. The denaturant was then neutralized with 1/10th volume of 3 M sodium acetate and extracted with an equal volume of chloroform. The DNA was then precipitated with 2 volumes of ice cold ethanol and incubated at –20^0 C for 1 hour and pelleted at 15,000 r.p.m. for 20 min in an Eppendorf microfuge. The liquid was aspirated and the DNA pellet was dried and resuspended in 15-20 µl of water. For sequencing, 3.5 ul of denatured DNA was used following the protocol outlined in the Sequenase Version 2.0 DNA sequencing kit manual (U.S. Biochemicals).

2.7.1 Sequencing of *gag* p17MA

Sequencing was performed using the Sequenase protocol as described above. The reactions were then run on a 6% polyacrylamide gel, the gel was transferred to blotting paper, and dried on a gel-dryer (BioRad). The gel was exposed to BioMax film (Kodak) for 24-28 hours, developed, read and entered manually into the sequence editor

of the Wisconsin Package Version 10.2, Genetics Computer Group (GCG), Madison, WI.

2.7.2 Sequencing of *nef*

Sequencing was performed using the Sequenase protocol as described above and, after polyacrylamide gel electrophoresis, read and entered manually into the sequence editor. Alternatively, prepared DNA was sequenced using an ABI PRISM® 3700 (Applied Biosystems, Foster City, CA) DNA automated sequencing system.

2.8 Phylogenic Reconstruction

2.8.1 Phylogenic Reconstruction of *gag* p17MA Isolates

The nucleotide sequences of the *gag* p17 gene (~396 bp) were aligned using ClustalX program (212), adjusted by hand, and then translated to corresponding amino acid sequences (~132 amino acids). The phylogenetic analysis was performed by using the PHYLIP (56) software package, version 3.5.7, and the Wisconsin Package version 10.2, Genetics Computer Group (GCG), Madison, WI. The tree was built from a nucleotide distance matrix (function DNADIST) by using the neighbor-joining method (function NEIGHBOR). The robustness of the neighbor-joining tree was assessed by bootstrap resampling of the multiple alignment (function SEQBOOT).

2.8.2 Phylogenic Reconstruction of *nef* Isolates

The nucleotide sequences of the *nef* gene (~618 bp) were aligned using ClustalX (212) adjusted by hand, and then translated to corresponding amino acid sequences (~206 amino acids). A model of evolution was optimized for the entire nucleotide

sequence data set using the approach outlined by Huelsenbeck and Crandall (99). Likelihood scores for different models of evolution were calculated using PAUP* (208), and a chi-square test was performed by Modeltest 3.06 (167). The model of choice was incorporated into PAUP* to estimate a neighbor-joining tree (176). Bootstrap values were based on 1,000 neighbor-joining searches. The tree was generated for the nucleotide sequences, and the reference HIV-1 sequence NL4-3 was used as an outgroup for the tree display. Using Modeltest and the Akaike Information Criterion (AIC) (6), all the null-hypotheses were rejected except for a TVM+G model with equal transition rates and no invariable site variation. The five rate categories were: R(A-C) = 1.2829, R(A-G) = 4.7709, R(A-T) = 1.4262, R(C-G) = 0.8436, R(C-T) = 4.7709, R(G-T) = 1.0. The base frequencies were freqA = 0.2809, freqC = 0.2342, freqG = 0.2576, freqT = 0.2273. The rate heterogeneity was taken into account using a gamma distribution with a shape parameter (α) of the distribution estimated from the data via maximum likelihood (235). The Gamma distribution shape parameter was $\alpha = 1.0292$. Similarly, a model of evolution was optimized for the data set from each pair. These models of evolution were used to estimate corrected pairwise nucleotide distances for the data sets from each pair using PAUP*.

2.8.3 Distance Calculations of *gag* p17MA Sequences

Pairwise distances, defined as the percentages of mismatches between two aligned nucleotide and amino acid sequences, were used to study the extent of genetic variability within an individual. For intra-individual variability (within mothers'

sequence sets), pairwise distances were calculated for all possible comparisons of pair of sequences within the set using DNADIST of the PHYLIP package. Then the minimum, maximum, and median nucleotide and amino acid distances were calculated.

2.8.4 Distance Calculations of *nef* Sequences

A model of evolution was optimized for the data set from each pair as described above using Modeltest and PAUP*. These models of evolution were used to estimate pairwise nucleotide distances for the data sets from each pair using PAUP* to correct for multiple mutations at a given site, differences in nucleotide frequencies, etc.. Then the minimum, maximum, and median nucleotide and amino acid distances were calculated.

2.8.5 Population Genetics

There has been a great effort in the development of theoretical frameworks for estimating genetic diversity of populations based on the coalescent theory (55, 64, 119, 120). Genetic diversity is defined in population genetics as $\theta = 2N_{ei}\mu$, where N_{ei} is the inbreeding effective population size and μ is the per nucleotide mutation rate (130). We have used these techniques to explore the dynamics of HIV evolution between the transmitting and non-transmitting mother HIV populations. Specifically, we used the Watterson estimate (221) of genetic diversity based on segregating sites and the Kuhner estimates assuming constant population size (119). These were estimated using the programs Coalesce, which is part of the Lamarc software package: http://evolution.genetics.washington.edu/lamarc.html. At the same time, we can

estimate the population growth using a coalescent approach, such that $N_e(t) = N_0 e^{-rt}$ where $N_e(t)$ is the effective population size at time t, N_0 is the initial effective population size, -r is the growth rate (or decline rate when r > 0), and t is time since the initial generation (80, 120, 198). Using this approach, we can compare population growth within patients. We determined the Vasco estimates of genetic diversity and the population growth rate (g) assuming a variable population size (218), using the EVE estimator that may be obtained from Daniel Vasco at http://bioag.byu.edu/zoology/crandall_lab/Vasco/eve.htm.

2.8.6 Synonymous and Non-synonymous Substitutions

The selection pressure was calculated as the ratio of nonsynonymous to synonymous substitutions by comparing all possible pairs of sequences within a mother's, infant's and pair's set by using SNAP (may be obtained from ftp://ftp-t10.lanl.gov/pub/aids-db/PROGS/SynNonsyn/) based on the method of Nei and Gojobori (157).

2.8.7 Statistics

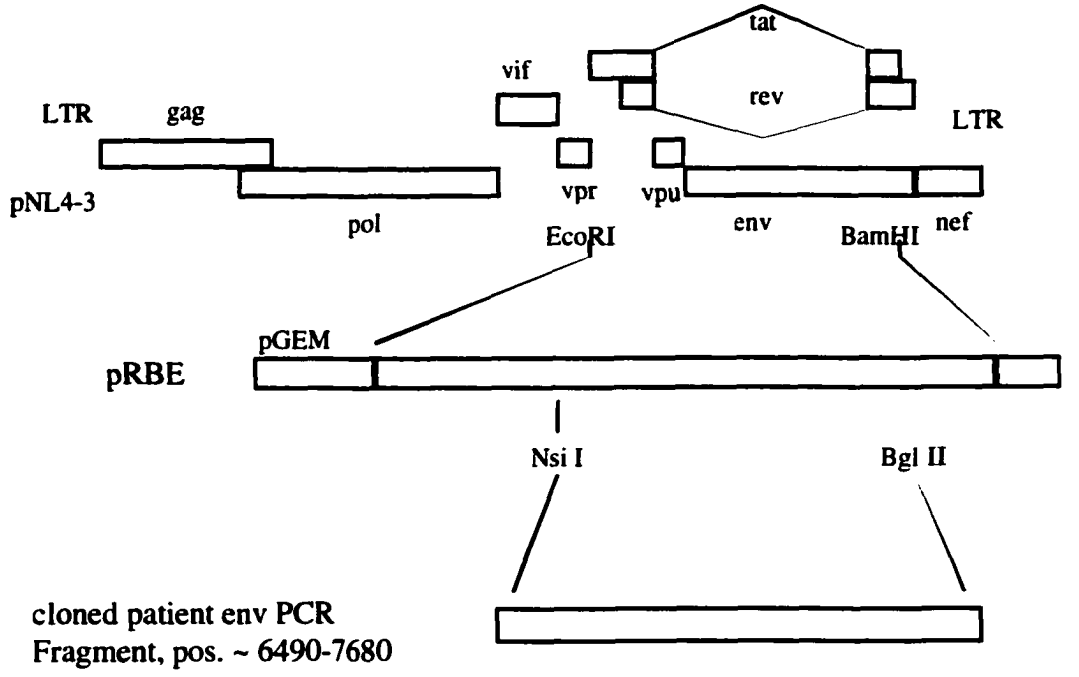
We used a chi-square test to determine the association of sequence motifs and features in the *gag* p17 matrix gene with transmitting mothers or non-transmitting mothers.

2.9 Construction of HIV-1 Chimeras

Sixteen *env gp120* region sequences from four mother-infant pairs, including mothers' minor and infants' major variants and twelve *env* region sequences from four

non-transmitting mothers were substituted into NL 4-3 using the BglII and NsiI sites flanking the *env* region. In order to clone the *env* region in NL4-3, the sub clone pRBE was created: pGEM (Promega) containing the EcoRI to BamHI fragment of NL4-3. The original PCR insert cloned into pGEM was partially digested with BglII and NsiI and the resulting fragment was reciprocally exchanged into pRBE digested with BglII and NsiI. Finally, the EcoRI-BamHI fragment containing the mother or infant *env* region was reciprocally substituted into pNL4-3. The reciprocally insertion of the sixteen mother-infant pair variants was performed by Mohammad Husain, Ph.D. Since the BglII sites in the *env* isolates of two of the non-transmitting mothers were missing, it was introduced into the *env* sequences by a PCR using the primers: ENV22 (as described above) and the primer ENV-BglII (5'-CAGGTCTAAAGATCTCG-GTTGTGTTATTAG, 7590 to 7620 antisense) and the same conditions as described above. All clones were sequenced following each transfer step.

Figure 9: Construction of gp120 Chimeras



2.10 Tissue Culture

2.10.1 Cell-Lines

The following ragents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: A3.01 from Dr. Thomas Folks (22). MT-2 from Dr Douglas Richman (87). The two T-lymphocyte cell lines, A3.01 and MT-2, were cultured in RPMI 1640 supplemented with 10% FBS, 1x penicillin/streptomycin.

2.10.2 Primary Cells

2.10.2.1 Isolation of Primary Cells

Primary monocytes-derived macrophages (MDM) and peripheral blood lymphocytes (PBL) were obtained from peripheral blood mononuclear cells (PBMC) on

a single-step Ficoll-Paque procedure (Pharmacia-LKB) gradient from the whole blood of normal donors. We collected blood from all donors with their consent and this study was approved by the Human Subjects Committee of the University of Arizona. The blood was spun for 45 min. at 450 RCF with the brake off. The PBMC were carefully aspirated from the interphase of the plasma and ficoll, washed three times with phosphate buffered saline (pH 7.4) containing no Ca^{++} or Mg^{++}, and resuspended in culture medium.

2.10.2.2 Primary Cell Culture

PBMC were cultured in RPMI 1640 supplemented with 10% FBS, 1x penicillin-streptomycin, 20 U of recombinant human interleukin-2 (Invitrogen, Carlsbad, CA) and 2.0 µg of phytohemagglutinin (PHA) (Sigma) per ml for 24 hours. PBL were removed, washed, and resuspended in the same medium without PHA. The remaining adherent cells were then cultured in RPMI 1640, 10% human serum, 1X penicillin-streptomycin, and 100 U of MCS-F (Sigma) before infection. The MDM were stained with esterase staining (Sigma) which resulted in ~90% pure monocyte population containing 15-20% of the original input cell numbers (155).

2.11 DNA Transfections for Virus Generation

HeLa cells were grown in DMEM and 10% FBS to 75% confluence, trypsinized and transfected with 10 µg of proviral DNA in 300 µl total volume of RPMI 1640, 10mM glucose via electroporation (276 V and 975 µF, BioRad). Cells were then plated onto 6-well tissue culture dishes in 2.5 ml DMEM 10% FBS, 1X penicillin-

streptomycin. Virus production was measured in culture media by reverse transcriptase (RT) assay 72 hours post-transfection.

2.12 Infections

Cell cultures using the lines A 3.01 ($1x10^6$/well), MT-2 ($1x10^4$/well), PBL ($2x10^6$/well), and MDM (~$0.5x10^4$/well) were infected with *env* gp120 region chimeras recombinant viruses derived directly from transfected HeLa cells. All infection experiments were performed in triplicates. In the infection charts error bars indicated the standard deviation. Due to the practical limitation to triplicates for each variant, combined with the expected variance in the RT assay, the standard deviation may reach fairly high levels. We used concentrations of HIV-1 of 10,000, 20,000, 50,000 cpm of RT activity of chimeric viruses. Viruses were adsorbed in target cells for 120 minutes in RPMI or DMEM without serum at 37°C in CO_2 incubator (Forma Scientific). After inoculation, 1 ml of culture media was added and the cells were maintained in the appropriate medium as described above for each cell and cell lines. Virus production was measured in 500 µl aliquots of supernatant removed every third day post-infection (PI) by reverse-transcriptase assay.

2.13 Cytopathic Effects

The syncytium inducing ability of the *env* region chimeras and a known syncytium inducing (SI) virus (NL 4-3) and a non-syncytium inducing (NSI) virus (Ada-M) was determined by infecting MT-2 cell lines with equal amounts of virus following the protocol in *ACTG Virology Manual for HIV Laboratories.* The

syncytium formation was monitored every three days in the cultures for 14 days post infection. Viruses that formed more than 4 syncytia per field of view in the culture were designated as SI viruses, whereas viruses that failed to induce syncytia in the culture were designated NSI.

2.14 Co-Receptor Usage

The following ragents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: U373-MAGI-CCR5E and U373-MAGI-CXCR4cem from Dr. Michael Emmerman (219). The human glio-blastoma cell line U373 was cultured in DMEM, 10% FBS, 0.2 mg/ml G418, 0.1 mg/ml hygromycin B and 1.0 μg/ml puromycin. The U373 cells expressing CD4 and either the CXCR4 or CCR5 co-receptor, contain the β-galacosidase indicator gene under the control of the HIV-1 LTR. Therefore, the HIV-1 transactivator tat, which is produced by the integrated provirus, upregulates the β-galacosidase reporter gene. After incubation with x-gal, infected cells stain blue. Briefly, 0.6 x 10^4 cells/well were seeded in 12 well plates. After overnight incubation, the culture medium was removed and 150 μl of virus dilution of culture supernatant and DEAE-Dextran to a final concentration of 20ug/ml were added. After a two hour adsorption period, fresh culture medium was added. 40-48 hours post infection the medium was removed and the cells were fixed with 1% formaldehyde and 0.2% glutaraldehyde for 5 min. Then. the cells were washed with PBS and stained for two hours (1ml stock: 950 ul PBS, 20ul 0.2 M potassium ferricyanide, 20 ul 0.2M potassium ferrocyanide, 1 ul 2M magnesium chloride, 10 ul of 40ul/ml x-gal in DMSO). To assess differences in infectivity of the

chimeras an equal amount of virus was used, as determined by reverse transcriptase assay (5000 cpm RT). The infected, blue-stained cells were counted and infectivity values for each of the chimeras are expressed as ratios of infected cells/well to infected cells/well of Adam (for U373-MAGI-CCR5) and NL4-3 (for U373-MAGI-CXCR4). Therefore, Adam and NL4-3 receive the infectivity score 1 and the chimeras a multiple or fraction of this score.

2.15 Reverse Transcriptase Assay

The presence of virus in a culture was determined by reverse transcriptase assay (RT) (4). Briefly, the RT assay consisted of 20 µl of culture supernatant mixed with 50 µl of RT cocktail (4 ml stock: 125 ul 3M KCl, 200 ul 0.1M DTT, 250 ul poly-A{100 ug/ml}, 250 ul 1M Tris pH 7.8, 25 ul 1M $MgCl_2$, 12.5 ul oligo dT{1mg/ml}, 12.5 ul 20% NP-40, 50 ul 10% Triton-X, 3.72 ml H_2O) which contains 1 ul ^{32}P (800 Ci/mmol) TTP/ml cocktail. The mix was incubated at 37° C for 2 hours, 5 ul was spotted onto DE81 paper (Whatman), and the paper was rinsed for 5 min in 2X SSC buffer (Life Technologies) three times. The discs were then placed into a scintillation vial with 4ml of scintillation cocktail and the amount of ^{32}P present on the discs was determined using a scintillation counter (Beckman). The amount of virus present in the culture is then expressed as counts per minute per ml of culture supernatant (cpm/ml).

3 RESULTS

3.1 Genetic Characterization of *gag* p17MA Sequences from Mothers and Infants following Perinatal Transmission

3.1.1 Phylogenetic Analysis of *gag* p17MA Sequences from Mother-Infant Pairs

Multiple PCRs were performed using PBMC DNA samples from seven mother-infant pairs to obtain 10 to 15 clones from each patient, resulting in 166 *gag* p17 clones, that were sequenced and used for analysis. The neighbor-joining (NJ) tree in Figure 10 was constructed by combining the 166 *gag* p17 sequences from the seven mother-infant pairs and the reference strain NL4-3 (subtype B)(153). The phylogenetic analysis revealed that the sequences from the seven mother-infant pairs were well discriminated and confined within subtrees, that were strongly supported by high bootstrap values (Fig 10). This indicates that the epidemiologically linked mother-infant pairs were closer to each other than to epidemiologically not linked sequences, and that there was no PCR product cross contamination (116, 127). Furthermore, the seven subtrees indicate that the p17 sequences are homogeneous within each mother and infant and that the mother and infant sequences were generally separated. This suggests that the infants were infected by one or few very similar variants present in the mother. We also constructed a NJ-tree combining the 166 *gag* p17 sequences from our seven mother-infant pairs and 181 other sequences from infected individuals present in HIV databases (Fig 11). These additional p17 sequences were from four independent studies of *gag* p17 sequences from infected individuals (100, 109, 156, 200) including 28 sequences belonging to the groups M (subtypes A, B, C, D, G and A-recombinants A/C, A/D, A/E,

A/G), O and N (196). In this global tree the 347 p17 sequences are grouped in a similar way as in trees based on the *env* (3), *vif* (239) and *vpr* (238) genes. The largest star-like cluster, contained all subtype B sequences. The subtype A, A recombinant and subtype D sequences diverged in multiple lineages. Subtype C (strains from Ethiopia and Brazil) diverged independently. The average distance among subtype B sequences, intra and inter-individual distances included, was 12%, ranging from 0 to 33%. The average pairwise distance between sequences of subtype A and A-recombinants was 18%, ranging from 5 to 33%. The average pairwise distance within the subtype C cluster was 21% ranging from 7 to 49%. The average distance between sequences from subtype B and the cluster of subtype A, C, and recombinants was 25%. All subtypes differ more than 40% from the group O sequences (MVP, ANT70). This analysis suggests that the *gag* p17 sequences from our seven mother-infant pairs are more closely related to subtype B than any other subtype. The *gag* p17 sequences of our seven mother-infant pairs are distinct from each other and from p17 sequences of other infected individuals present in HIV database, including HIV-1 NL4-3 that is used in our laboratory.

Figure 10: Phylogenetic Analysis of *gag* p17MA Sequences from Mother-Infant Pairs

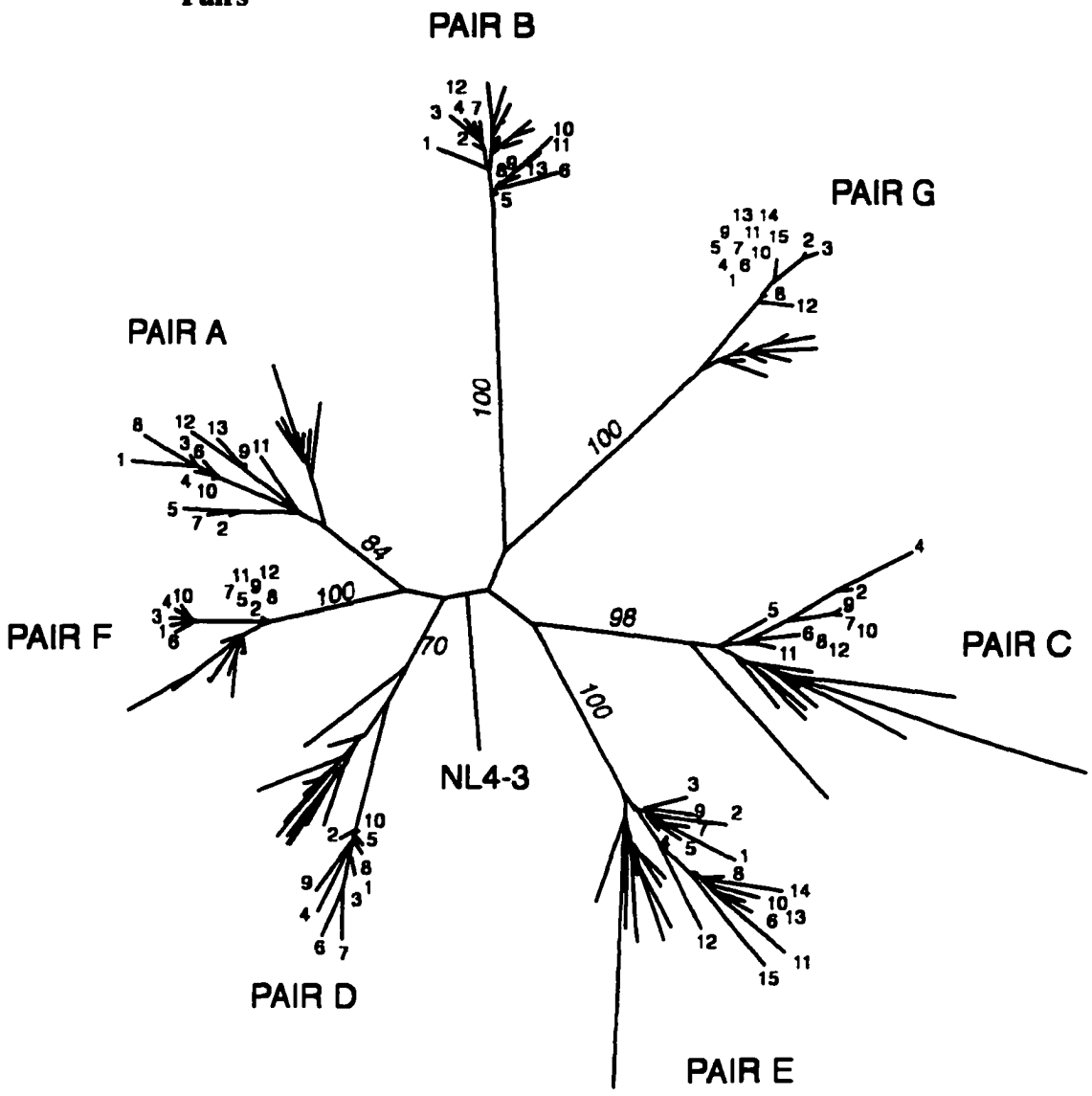


Fig. 10: Phylogenetic analysis of 166 *gag* p17 sequences from seven mother-infant pairs (A, B, C, D, E, F, and G). The NJ-tree is based on the distances calculated between the nucleotide sequences from the seven mother-infant pairs. Each terminal node represents one p17 sequence. The mother sequences in each pair are numbered, whereas the infant sequences are unlabelled. The numbers at branch points indicate the occurrence of branches over 100 bootstrap resamplings of the data set. The mother-infant pairs formed distinct clusters and are well discriminated and confined within subtrees, indicating that the epidemiologically linked sequences are closer to each other than epidemiologically unlinked sequences and that there was no PCR product cross contamination.

Figure 11: Global Phylogenetic Tree of *gag* p17MA Sequences

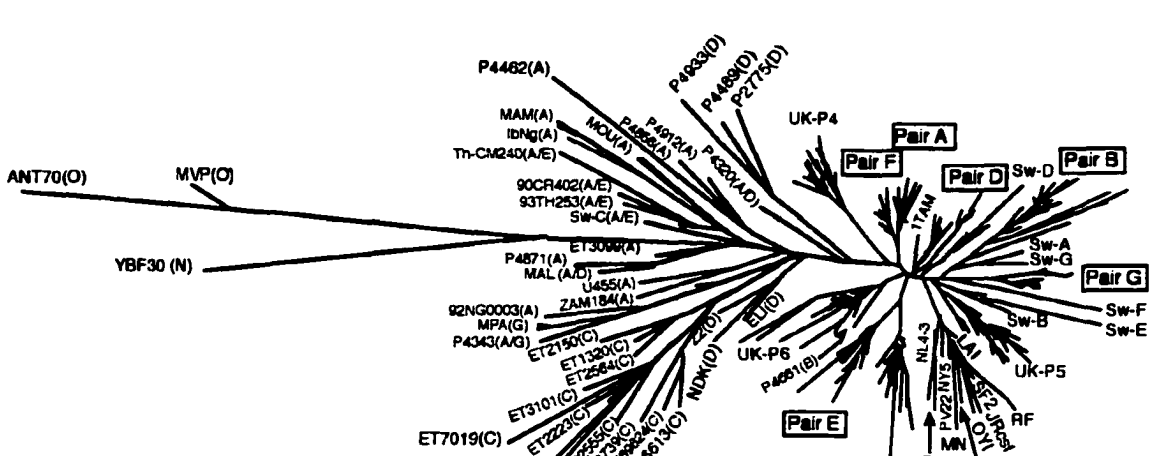


Fig. 11: Global phylogenetic tree of *gag* p17 sequences, including our 166 sequences from the seven mother-infant pairs and 181 other sequences from HIV-1 databases. The mother-infant pairs' sequences are labeled as pair A, B, C, D, E, F, and G. The sequences from Karlsson et al. (109) were obtained from seven patients labeled as Sw-A, B, D, E, F and G (subtype B) and Sw-C (subtype A). The sequences from Hughes et al. (100) were obtained from lymphoid and nonlymphoid tissues of 3 patients (UK-P4, P5, P6). Some of the sequences from Narwa et al. (156) were selected from mother and infant pairs, LEN and RYO (belonging to subtype B), MAM and MOU (subtype A), MPA (subtype G). The sequences from Sonnerborg et al., (200) are from a study of HIV-1 subtypes in Ethiopia (accession numbers U56336-U56366) are labeled ET for Ethiopian strains (all belonging to subtype C) and P for African isolates. The other sequences from the HIV-1 databases belong to the following subtypes (accession numbers shown in parenthesis): subtype B, MN (M17449), Jrcsf (M38429), NL4-3(U26942), NY5(M19921), PV22(K02083), SF2(K02007), CAM1(D10112), LAI(K02013), RF (M17451), OYI (M26727), 1TAM (pdb|1TAM); subtype D: ELI(K03454), NDK(M27323), Z2(M22639); subtype A and A/C, A/D, A/E, A/G recombinants: U455(M62320), Zam184(U86780), subtype C: 92BR025(U52953) and ET-C2220(U46016); subtype O: MVP(L20571), ANT70(L20587) and subtype N: YBF30 (AJ006022). Subtypes other than B are indicated in parenthesis after the isolate name. The largest star like cluster contains all subtype B sequences. Our seven mother-infant pair sequences clustered with subtype B sequences.

3.1.2 Analysis of Coding Potential of *gag* p17MA Isolates from Mothers and Infants

Figure 12 shows the multiple alignments of the deduced amino acid sequences (~132 amino acids) from our 166 p17MA sequences aligned in reference to the subtype B consensus sequence (consB)(153). The coding potential of the *gag* p17 open reading frame was maintained in most of the sequences. We analyzed 166 different p17 clones and 143 of them contained an intact p17MA open reading frame; this represents a 86.2% frequency of conservation of intact p17 open reading frames. The frequency of conservation was higher (90.2%) if mother E sequences are excluded that contained 60% defective p17 sequences. The frequency of defective p17 genes in our seven mother-infant pairs' sequences was 13.8% (9.8% excluding mother E sequences). We found that a total of 23 clones contained stop codons and/or lacked initiation codons. One striking observation was 60% defective p17 sequences in mother E. In addition, one clone in infant C (CI-8) and six clones in mother E (ME -6, 8, 10, 11, 12, and 14) contained stop codons or lacked initiation codons. However, the intact p17 sequences in mother E likely persisted in her infant. It is noteworthy that each mother-infant pair p17 amino acid sequences displayed a pattern that was not seen in epidemiologically unlinked pairs. For example, pair A shows 5 pair-specific substitutions, pairs B through G show 13, 12, 9, 11, 6, and 7 pair-specific substitutions respectively. Comparison shows that two pairs differ by at least 5 pair-specific substitutions. We also noticed patient-specific substitutions relative to the cons-B sequence. In addition, there were some common signature sequences seen in all mother-infant pairs' sequences, including

glutamic acid (E) or aspartic acid (D) at position 55, tyrosine (Y) or phenylalanine (F) at position 79, an aspartic acid (D) or glutamic acid (E) at position 93 and 102, and an alanine-glutamic acid (AD) at position 122-123 (Fig. 12).

Figure 12: Amino Acid Alignments of *gag* p17MA Sequences from Mother-Infant Pairs

PAIR A

Membrane binding, Residues 1-31

KIEEQN

QVSQNY

```
                    10        20        30        40        50        60        70        80        90       100       110       120       132
consB       MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEVKDTKEA LKKIEEQQNK SKKKAQQAAA A--TGNSS----QVSQNY
AM-1        .......... .......... S.......* .......... .......... .......... .E....A... E.......P. .I........ ..D....... .......... ........T. .--.....----.G....
AM-2        .......... .......... .......... .......... ......K... .......... .E....A... ........P. .I........ ..D....... .......... .......... .--.....----.G....
AM-3(4)     .......... .......... .......... .......... .......... .......... .E....A... ........P. .I........ ..D....... .......... ........T. .--.....----.G....
AM-5        .......... .........Q .......... .......... .......... .......... .E....A... ........P. .I........ ..D....... .......... .....P.... .--.....----.G....
AM-7        .......... .......... .......... .......... .......... .......... .E....A... ........P. .I........ ..D....... .......... .......... .--.....----.G....
AM-8        .......... .......... .......... .......... .......... .......... .E....A... ........P. .IR....... ..D....... .......... ........T. .--.....----.G....
AM-9        .......... .......... .......... .......... .......... .......... .E....A... ........P. .......... ..D....... .......... .......... .--.....----.G....
AM-11       .......... .......... .......... .......... .......... .......... .E....A... ........P. .......... ..D....... .......... .......... T--.....----.D....
AM-12       .......... .......... .......... .......... .....A.... .......... .E....AP.. ........P. .......... ..D....... .......... .N........ .--.....----.G....
AM-13       .......... .......... .......... .......... .......... .......... .E....A... ........P. .......... ..D.....P .......... .......... .--.....----.G....
AI-1(2)     .......... .....*.... .......... .......... .......... .......... .E....A... ........P. .......... ..D....... .......... .......... .--S....----.G....
AI-2        .......... .......... .......... .......... .......... .......... .ER...A... ........P. .......... ..D....... .......... .......... .--....N----.G....
AI-3        .......... .......... .......... ..D....I.. .......... ........R. .E....A... ........P. .......... ..D....... .......... .......... .--.....----.G....
AI-4(2)     .......... .......... .......... .......... .......... .......... .E....A... ........P. .......... ..D....... .......... .......... .--.....----.G....
AI-5        .......... .......... .......... .......... .......... .......... .E....A... ........P. .......... ..D....... .......... .........G .--.....----.G....
AI-6        .......... ....K..... .......... .......... .......... .......... .E....A... .....G..P. .......... ..D....... .......... .......... .--.....----.G....
AI-7        .......... .......... .......... .......... .......... .......... .E....A... ........P. .......... ..D....... .......... .......... .--....G----.G....
AI-9        .......... .......... .......... .......... ........D. .......... .E....A... ........P. .......... ..D....... .......... .......... .--.....----.G....
```

PAIR B

```
                    10        20        30        40        50        60        70        80        90       100       110       120       132
consB       MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEVKDTKEA LKKIEEQQNK SKKKAQQAAA A--TGNSS----QVSQNY
BM-1(3)     .......... .Q........ .......R.. ...L...... .......... ....D..... .......... .....K..P. .......... ...I...... .D........ ...T.P.... -AD.....----...KI-
BM-2        .......... .Q........ .......R.. .E.L...... .......... ....D..... ...I...... .....K..P. .......... ...I...... .D........ ...T.P.... -AD.....----...KI-
BM-3        .......... .Q........ .......R.. ...L...... .......... ....D..... .......... .....K..P. .......... G..I...... .D........ ...T.P.... -AD.....----...KI-
BM-4        .......... .Q........ .......R.. ...L...... .......... ....D..... .......... .....K..P. .......... ...I....K. .D........ ...T.P.... -AD.....----...KI-
BM-5        .......... .Q........ .......R.R ...L...... .......... ....D..... S......... .....K..P. .......... ...I...... .D........ R....P.... -AD.....----....I-
BM-6        .......... .Q........ .......R.R ...L...T.. .......... ....D..... .......... .....K..P. .......... G..I...... .D........ ....TP.... -AD.....----....I-
BM-8        .......... .Q........ .......R.R ...L...... .......... ....D..... .......... .....K..P. .......... ...I...... .D........ .....P.... -AD.....----...KI-
BM-9        .......... .Q........ .......R.R ...L...... .......... ....D..... .......... .....K..P. .......... ...I...... .D........ ...T.P.... -AD.....----...KI-
BM-10       .......... .Q........ .......R.R ...L...... .......... ....DS.... .......... .....K..P. .......... ...I...... .D........ ...T.P.... -AD.....----...KI-
BM-11       .......... .Q........ .......R.R ...L...... .......... ....DS.... .......... .....K..F. .......... ...I...... .D........ .....P.... -AD.....----...KI-
BM-13       .......... .Q........ .......R.R ...L...... .......... ....D..... .......... .....K..F. .......... ..GI...... .D........ .....P.... -AD.....----....I-
BI-1(2)     .......... .Q........ .......R.. ...L...... .......... ....D..... .E........ ...DIK..P. .......... ...I...... .D........ N....P.... -AD.....----...KI-
BI-2        .......... .Q........ .......R.. ...L...... .......... ....D..... .......... ....IK..P. ......Y... ...I...... .D........ .....P.... -AD.....----...KI-
BI-3(2)     .......... .Q........ .......R.. ...L...... .......... ....D..... .......... ....IK..P. .......... ...I...... .D........ .....P.... -AD.....----...KI-
BI-4(2)     .......... .Q........ .......R.. ...L...... .......... ....D..... .......... ....IK..P. .......... ...I...... .D........ ..E..P.... -AD.....----...KI-
BI-5(2)     ........C. .Q........ .......R.. ...L...... .......... ....D..... .......... ....I*..P. .......... ...I...... .D........ .....P.... -AD.....----...KI-
BI-6        .......... .Q........ .......R.. ...L...... .......... ....D..G.. .E.......R ....IK..P. .......... ...I...... .D........ .....P.... -AD.....----...KI-
```

PAIR C

Membrane binding, Residues 1-31

KIEEQN

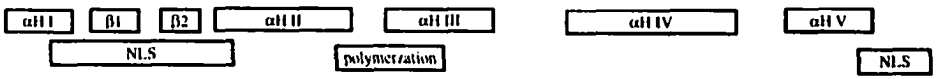
QVSQNY

	10	20	30	40	50	60	70	80	90	100	110	120	132
consB	MGARASVLSG	GELDRWEKIR	LRPGGKKKYK	LKHIVWASRE	LERFAVNPGL	LETSEGCRQI	LGQLQPSLQT	GSEELRSLYN	TVATLYCVHQ	RIEVKDTKEA	LEKIEEQNK	SKKKAQQAAA	A--TGNSS----QVSQNY
CM-2	.....	....K.....	....ER....	.....	.....S.	.....	.....	.....K....	.IW.....	..D.....	.D.V.....	.....-	-AD...K.S---.....
CM-4	.....	....K.....	....ER....	.....	.....S.	...P.....	.....	...K.K..F.	.IW.....	..D.....	.D.V.....	.....-	-AD...K.S---.....
CM-5	.....	.....	.....R....	.....	.....S.	.....	.....	.....N....	.IW.....	..D.....	.D.....	....GT...-	-AD...K.S----..N...
CM-6(4)	.....	.....	.....R....	.....	.....S.	.....	.....	.....K..F.	.I.V.....	..D.....	.D.V.....	.....-	-AD...N.S---.....
CM-7(2)	.....	.....	.....R....	.....	.....S.	.....	.....	.....K..F.	.IW.....	..D.....	.D.V.....	.....T.-	-AD...K.S---.....
CM-9	.....	.....	.....R....	.....	.....S.	.....	.....	.....K..F.	.IW.....	S.D.....	.D.V.....	.....T.-	-AD...K.S---.....
CI-1	.....	.....F.	.QQ....Q..	.....	.....	.....	.....	.....K..F.	.I.V.....	..N.....	.D.V.....	.....-	-AD...KNS---.....
CI-2(2)	.....	.....	.....Q..	.....	.....S.	.....	.....	.....K..F.	.I.V.....	..N.....	.D.V.....	.....-	-AD...K.S---.....
CI-3	...K.....	.....	.....Q..	.....	.....S.	.....	.....	.....K....	.I.V.....	..D.....	.D.V.....	.....-	-AD...K.S---.....
CI-5	.....	.....	.....R..	.....	....T...S.	.....	.....	.....K..F.	.IW.....	..D.....	.D.V.....	.....-	-AD...N.S---.....
CI-6	.....	.....	.....Q..	.....	.....S.	.....	.....	.....K....	.I.V.....	..N.....	.D.V.....	.....-	-AD...K.S---.....
CI-7	.....	.....	.....Q..	.....	.....I..S.	.....	.....	.....K..F.	...V.....	..N.....	.D.V.....	.....-	-AD...N.S---.....
CI-8	I.....	.....*...Q	.....Q..	....DM....	....T...S.	.....	.....	.....K..F.	.ISV.....	..D.....	.D.V.....	G.....-	-AD...K.S---.....
CI-9	.....	.....	.....Q..	.....	.....S.	.....	.....	.....K..F.	.I.V.....	..D.....	.D.V.....	.....L...-	-AD...KNS---.....
CI-10	.....	.....	.....Q..	.....	.....S.	.....K..	.....	.....K..F.	...V.....	..N.....	.D.V.....	.....-	-AD...KNS---.....
CI-11	.....	.....	.....Q..	.....	.....S.	.....	.....R	.....K....	.I.V.....	..D.....	.D.V.....	.....-	-AD...N.S---.....
CI-12	.....	.....	.....Q..	.....	.....S.	.....	.....	.....K....	.I.V.....	..D.....	.D.V*....	.....-	-AD...K.S---.....

PAIR D

	10	20	30	40	50	60	70	80	90	100	110	120	132
consB	MGARASVLSG	GELDRWEKIR	LRPGGKKKYK	LKHIVWASRE	LERFAVNPGL	LETSEGCRQI	LGQLQPSLQT	GSEELRSLYN	TVATLYCVHQ	RIEVKDTKEA	LEKIEEQNK	SKKKAQQAAA	A--TGNSS----QVSQNY
DM-1	.....	.....	.....R...R	...L.....	.....	.....	.....	.....K..F.	.I.....	..D.....	.....R..	.....--	-AD.....SQVS.....F
DM-2(3)	.....	.....	.....R...R	...L.....	.....	.....	.....	.....K..F.	.I.....	..D.....	.....	.....--	-AD.....SQVS.....F
DM-3(2)	.....	.....	.....R...R	...L..T...	.....	.....	.....	.....K..F.	.I.....	..D.....	.....	.....--	-AD.....SQVS.....F
DM-4	.....	.....	.....R...R	...L.....	.....	.....	.....	.....K..F.	.I.....	..D.....	.....	.....A..--	-AD..Y..SQVS.....F
DM-6	.....	.....	.....R...R	...L.....	.....	.....	V..I.....	.....K..F.	.IR.....	..D.....	.....	.....--	-AD.....SQVS.....F
DM-7	.....	.....	.....R...R	...L.....	.....	.....	...I.....	.....K..F.	.IT.....	..D.....	.....	.....--	-AD.....SQVS.....F
DM-9	.....	.....	.....R...R	...L.....	.....	.....	.....	.T...K..F.	.I.....	..D....*..	.....	.....--	-AD.....SQVS.....F
DI-1	.....	.....	.....R	.....R....	.....	.....	.....S....	.....K....	.....	..D.....	.D.....	.....--	-AD.....SQVS.....F
DI-2	.....	.....	.....Q	.....	.....	.....	.....S....	.....N....	.....	..D.....	.....	.....--	-AD.....SQVS.....F
DI-3	.....	.....	.....N...R	.....	.....	.....	.....S....	.....	.....	..D.....	.D.....	.....--	-AD.....SQVS.....F
DI-4	.....	.....	.....Q	.....	.....	.....	.....S....	.....	.....	.VD.....	.D.....	.....--	-AD.....SQVS.....F
DI-5	.....	.....	.....R	.....	.....	.....	.....S....	.....	.....	..D.....	.D.....	.....--	-AD.....SQVS.....F
DI-6	.....E.....	.....	.....R.....	.....	.....P.....	.....	.....	.....	.....	.....D.....D.....	.....	---AD....TSQVS.....F	
DI-7	.....	.....	.....R	.....	.....	..P.....	.....S....	.....K....	.....	..D.....	.....	.....--	-AD.....SQVS.....F
DI-8	.....	.....	.....Q	.....	.....	.....S...	.....S....	.....	.....	..D.....	.....	.....--	-AD.....SQVS.....F
DI-9	.....	.....	.....Q	.....	.....	.....	.....S....	.....K....	.....	..D.....	.D.....	.....--	-AD.....SQVS.....F
DI-10	.....	.....	.....Q	.....	.....	.....	.....	.....K....	.....	..D.....	.D.....	.....-	-AD...N.----.....
DI-11	.....	.....	.....Q	.....	.....	.....	.....	.....	.....	..D.....	.D.....	.....--	-AD.....SQVS.....F
DI-13	.....K.....	.....	.....R	.....	.....	.....	.....S....	.....T....	.....	..D.....	.D.....	.....*..--	-AD.....SQVS.....F
DI-12	.....	..*.....	.....R	...L.....	.....	.....	.....	.....	.....	..D.....	.D.....R..	.....--	-AD.....SQVS.....F

PAIR E



Membrane binding, Residues 1-31

KIEEQN

QVSQNY

	10	20	30	40	50	60	70	80	90	100	110	120	132
consB	MGARASVLSG	GELDRWEKIR	LRPGGKKKYK	LKHIVWASRE	LERFAVNPGL	LETSEGCRQI	LGQLQPSLQT	GSEELRSLYN	TVATLYCVHQ	RIEVKDTKEA	LEKIEEQQN	SKKKAQQAAA	A--TGNSS---QVSQNY
EM-1(2)	.....R	.....•....	.....Q.R	.....K.	.....	....D.....	.....A...	.....	.I.V.....	-.D.R.....	.D.....	.....	.AD...N.----.....
EM-3	.....	.....	•.....Q.R	.....	.....	....D.....	.....	.....	.I.V.....	..D.R.....	.D.....	.....	.AD...N.----.....
EM-4	.....	.....	.....Q.R	.....	.....	...LD.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....	-AD...N.----.....
EM-5	.....	.....	.....Q.R	.....	.....	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....	.AD...N.-----I....
EM-6	I.....R	.....•....	.....Q.R	.....	.....D.	....D.....	.....A...	.....	.I.V....E	.RD.R.....	.D.....	.....S.	.AD...N.----.....
EM-7	.....	.....	.....Q.R	.....V...	.....	....D.....	.....	.....	.I.V.....	..D.R.....	.D.....	...T.....	.AD...N.----.....
EM-8	I.....R	.....•....	.....Q.R	.....K.	.....	....D....V	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....-	-AD...N.----.....
EM-9	.....	.....	.....Q.R	.....'..	.....	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....	.AD...N.----.....
EM-10	I.....R	.....•....	.....Q.R	.....K.	.....	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	...V.....	.AD...N.----.....
EM-11	I.....R	.....•....	.....Q.R	.....K.	....T.....	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....-	-AD...N.----.....
EM-12	I.....R	.....•....	.....Q.R	.....K.	.....	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	V	.AD.R.N.----.....
EM-13	.....	....	.....Q.R	.....	.....H	....DS....	.....	.....G....	.I.V.....	..D.R.....	.D.....	.....	.AD...N.----.....
EM-14	I.....R	.....•....	.....Q.R	.....	.....	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....A....	.AD...N.----.....
EM-15	.....	.....	.....E.Q.R	.....K.	.....	....D.....	.....A...	.....	.I.V.....	..DAR.....	.D.....	.....	.AD...N.----.....
EI-1	.....	.....	.....Q.R	.....	.....	....D.....	.....	.....	.I.V.....	..DAR.....	.D.....	.....	.AD...N.----.....
EI-2(2)	.....	.....	.....Q.R	.....	.....	....D.....	.....	.....	.I.V.....	..D.R.....	.D.....	N.....	.AD...N.----.....
EI-3	.....	.....	.....Q.R	.....	.....	....D.....	.....	.....	...V.....	..D.R.....	.D.....	.....	.AD...N.----.....
EI-4	.....	.....	.....Q.R	.....	.....	....D.....	.....G...	.....	.I.V.....	..D.RN....	SD.....	.....	.AD...N.----.....
EI-5	.....	.....	.....Q.R	.....	.....	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....	.AD...N.----.....
EI-6	.....	.....	.....Q.R	.....	.....	....D.....	.....	.....	...V.....	..D.R.....	.D.....	.....	.AD..HN.----.....
EI-7(2)	.....	.....	.....Q.R	.....	.....	....D.....	.....	.....	.I.V.....	..D.R.....	.D.....	.....	.AD...N.----.....
EI-8	I.....	.K.N.....	.....Q..	.....	.....I..D.	.K..N..K..	.R.....	.....	.I.V.....	..D.R.....	.D.....	.....	.AD...N.----.....
EI-9	.....	.K.....	.....Q.R	.....	.....	....D.....	.....	.....	.I.V....L.	..D.R.....	.D.....	.....	.AD...N.----.....
EI-10	.....	.K.....	.....Q.R	.....	..Q.....	....D.....	.....	.....	.I.V.....	..D.R...G.	.D.....	.....	-AD...N.----.....
EI-11	.....	.....	.....Q.R	.....	.....H...	....D.....	.....A...	.....	.I.V.....	..D.R.....	.D.....	.....	.AG...N.----.....
EI-14	.....	.....	.....Q.R	.....	.....	....V.....	.....	.....	.I.V.....	..D.R.....	.D.....	.....•....	.AD...N.----.....

PAIR F

	10	20	30	40	50	60	70	80	90	100	110	120	132
consB	MGARASVLSG	GELDRWEKIR	LRPGGKKKYK	LKHIVWASRE	LERFAVNPGL	LETSEGCRQI	LGQLQPSLQT	GSEELRSLYN	TVATLYCVHQ	RIEVKDTKEA	LEKIEEQQK	SKKKAQQAAA	A--TGNSS---QVSQNY
FM-1	.....	.....	.....Q..	.....	.....	.....	.E....A...	.....	.I.....	..D.....	.D...Q....	.....	--N.....-----A....
FM-2(8)	.....	.....	.....Q..	.....	.....	.....	.E....A...	.....	.I.....	..D.....	.D.....	.....-	-AN.....-----A....
FM-4	.....	.....	.....Q..	.....	.....	.....	.E....A...	.....	.I.....	..D.....	.D.....E	.....-	-AN.....-----A....
FM-6	.....	.....	.....Q..	....A.....	.....	.....	.E....A...	.....	.I.....	..D.....	.D.....	.....-	-AN.....-----A....
FM-10	.....	.....	.....Q..	.....	.....	.....	.E....A...	.....	.I.....	..D.....	.D...G...	.....-	-AN.....-----A....
FI-1(2)	.....	.....	.....R..	.....	.....	.....	.E....A...	.....	.....	..D.....	.D.....	.....-	-AN.....----.....
FI-3	.....	.....	.....Q..	.....	.....	.....	.E....A...	.....	.....	..D.....	.D.....	.....-	-AN.....----.....
FI-4(2)	I.....	.....	.....R..	.....	.....	.....	.E....A...	.....	.....	..D.....	.D.....	.....-	-AN.....----.....
FI-5(2)	.....	.....	.....R..	.....	.....	.....	.E....A...	.....	.....	..N.....	.D.....	.....-	-AN.....----.....
FI-7(2)	.....	.....	.....R..	.....	.....P	.....	.E....A...	.....	.....	..D.....	.D.....T.	.....T-	-AN.....----.....
FI-10	.....G..	.....	.....R..	.....	.....	.....	.E.H..A...	.....	.....	..D.....	.D.....	.....-	-AN.....----.....
FI-11	.....G..	.....	.....R..	.....	.....	.....•...	.E.H..A...	.....	.....	..D.....	.D.....	.....-	-AN.....----.....

PAIR G

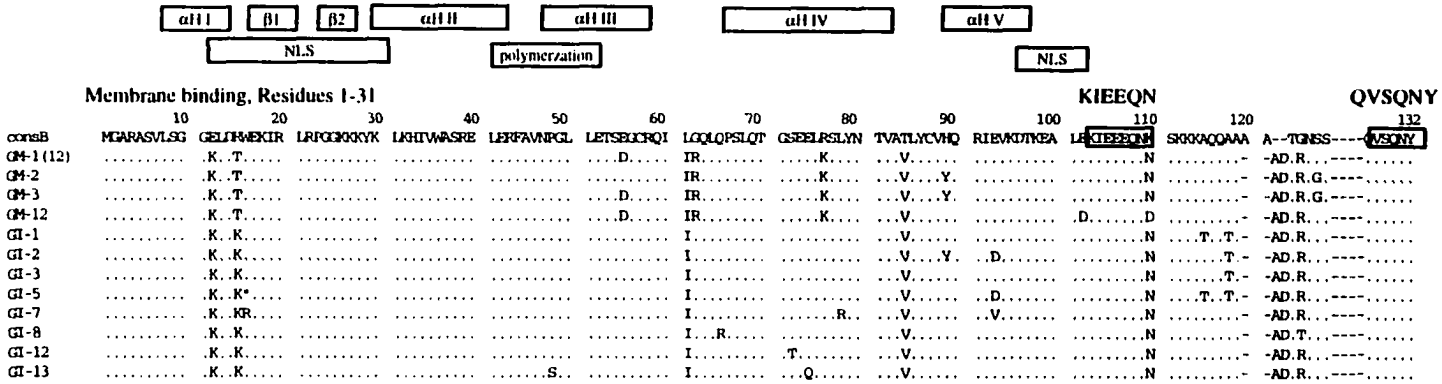


Fig 12: Multiple alignments of the deduced amino acid sequences of *gag* p17 sequence isolates from seven mother-infant pairs following perinatal transmission. In the alignment, the top sequence is the consensus sequence of subtype B (consB) and pairs A, B, C, D, E, F and G represent the seven mother-infant pairs. M indicates mothersequences and I indicates infant sequences. Identical clones are indicated in parentheses. Dots replace amino acids identical to the consB sequence, dashes represent gaps, and asterisks represent stop codons. Above the alignment, structural features, comprised of 5 α-helices (αH-I through αH-V) and 2 β-sheets (β1, β2), are shown. In addition, the functional domains, including membrane binding domain, nuclear localization signal, and polymerization site as well as two motifs implicated in transmission are shown.

3.1.3 Variability of *gag* p17 Sequences of Epidemiologically linked Mother-Infant Pairs

To determine the degree of variability of the *gag* p17 sequences from our seven mother-infant pairs, we analyzed variation in nucleotide and amino acid sequences as shown in Table 3. The nucleotide sequences of the p17 within mothers (mothers A, B, C, D, E, F and G) differed by 2.3, 1.0, 1.8, 1.2, 2.2. 0.3 and 0.3% (median values), respectively, ranging from 0 to 4.0%. The variability in the infant sets (infants A, B, C, D, E, F and G) was very similar to the variability of the mothers sets and differed 0.8, 0.8, 2.0, 2.0, 1.7, 1.0, and 1.1% (median values), respectively, ranging from 0 to 6.0%. Interestingly, the variability between mother and infant sets (epidemiologically linked pairs A, B, C, D, E, F and G) was also on the same order of 2.5, 1.3, 2.5, 3.5, 2.7, 1.8 and 2.0% (median values), respectively, ranging from 0.2 to 6.5%. The median values of amino acid sequence variability of p17 within mothers A, B, C, D, E, F and G were 2.3, 2.3, 3.0, 1.5, 3.8, 0.8, and 0.1%, within infants were 2.3, 2.3, 3.8, 3.0, 2.2, 2.3, and 3.0%, and between epidemiologically linked mother-infant were 3.0, 3.8, 5.2, 6.7, 4.1, 3.0 and 4.5% respectively. Interestingly, the younger infants (A, B and F) sequences were more homogeneous than the older infants' sequences. Moreover, the variability in general was lower between mother-infant pairs, suggesting that the epidemiologically linked mother-infant pairs' sequences are closer to each other. We also determined how much of the variability of the p17 sequences may be due to errors made by UlTma DNA polymerase. We rarely found any errors made by UlTma when using a known sequence of HIV-1 NL4-3 for PCR amplifications and DNA sequencing of the *gag* p17 gene.

Table 3: Nucleotide and Amino Acid Distances between *gag* p17MA Sequences from Mother-Infant Pairs

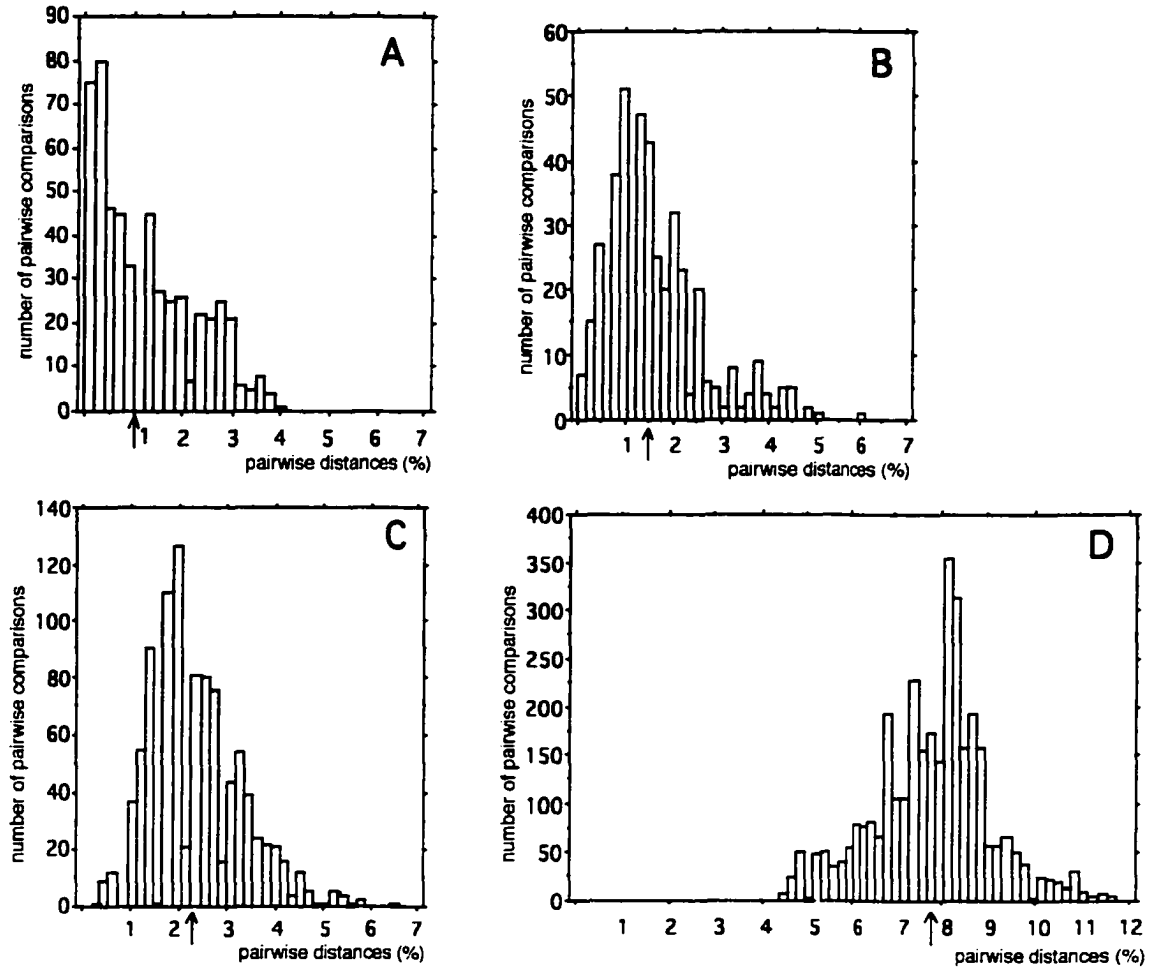
		% Distances[a]								
		Within mother set			Within infant set			Between mother and infant set		
Sequence	Pair	min	median	max	min	median	max	min	median	max
Nucleotide	A	0.3	2.3	3.8	0.0	0.8	2.0	1.5	2.5	3.5
	B	0.0	1.0	2.0	0.0	0.8	1.3	0.5	1.3	2.8
	C	0.0	1.8	3.8	0.8	2.0	6.0	1.0	2.5	6.5
	D	0.0	1.2	2.2	0.5	2.0	4.1	2.2	3.5	4.6
	E	0.0	2.2	4.0	0.2	1.7	4.5	0.2	2.7	5.7
	F	0.0	0.3	0.5	0.0	1.0	2.3	1.0	1.8	3.1
	G	0.0	0.3	1.5	0.3	1.1	2.0	1.5	2.0	3.3
	Total[b]	0.0	1.0	4	0.0	1.5	6.0	0.2	2.3	6.5
Amino Acid	A	0.0	2.3	6.1	0.0	2.3	3.8	0.0	3.0	6.1
	B	0.0	2.3	6.1	0.0	2.3	3.8	1.5	3.8	6.8
	C	0.0	3.0	7.5	0.8	3.8	10.5	2.3	5.2	12.0
	D	0.0	1.5	3.7	0.7	3.0	5.3	4.4	6.7	8.9
	E	0.0	3.8	8.2	0.0	2.2	9.7	0.0	4.1	11.2
	F	0.0	0.8	1.5	0.0	2.3	4.5	1.5	3.0	6.0
	G	0.0	01	3.8	0.8	3.0	4.5	3.8	4.5	7.6
	Total[b]	0.0	1.5	8.2	0.0	2.3	10.5	0.0	4.5	12.0

[a]Expressed as percent nucleotides (for nucleotide sequence) or percent amino acids (for amino acid sequence).
[b]Calculated for all the pairs taken together.

3.1.4 Variability of *gag* p17 Sequences of Epidemiologically unlinked Mother-Infant Isolates

Figure 13 shows histograms of the distributions of the *gag* p17 nucleotide sequence distances within the same mother (panel A) , within the same infant (panel B) , between epidemiologically linked mother-infant pairs (panel C), and between epidemiologically unlinked mothers (panel D); the medians of distributions were 1.0, 1.5, 2.3 and 7.7%, respectively. The data suggest that the p17 sequences of epidemiologically linked mother-infant pairs were closer than those of epidemiologically unlinked individuals, keeping in mind that a low degree of variability of the p17 sequences was observed in our seven mother-infant pairs. By using the sequence distances of a conserved region such as *gag* p17, we were able to easily differentiate the epidemiologically unlinked individuals from epidemiologically linked mother-infant pairs (Fig. 13D). Interestingly, the *gag* p17 sequences from older infants (14, 24, 28 and 34 months age), that were either as, or more, heterogeneous than their mothers, were still closer to their mothers' sequences (pair C, D, E and G) than epidemiologically unlinked sequences (Table 3), suggesting that the epidemiologically linked viral sequences can be identified even in older infants.

Figure 13: Distribution of Nucleotide Distances between *gag* p17MA Sequences from Mother-Infant Pairs



A Distribution of nucleotide distances within mothers sequences
B Distribution of nucleotide distances within infant sequences
C Distribution of nucleotide distances between epidemiologically linked mother-infant pair sequences
D Distribution of nucleotide distances between epidemiologically unlinked mothers sequences

3.1.5 Rates of Accumulation of Non-synonymous and Synonymous Substitutions

Several studies suggested that a ratio of nonsynonymous (d_n) to synonymous (d_s) substitutions (d_n/d_s) of more than 1 indicates positive selection pressure by the immune

system for escape variants (131, 230, 245, 247). Others have argued that the value of the ratio is less important than a shift from the background ratio (193). Comparisons of infant's sequences with mother's sequences from pairs A, B, C, D, E, F and G gave ratios of nonsynonymous to synonymous substitution, d_n/d_s , as determined by the method of Nei and Gojobori (157), of 0.2, 0.5, 0.6, 0.6, 0.3, 0.5 and 0.9, respectively. Thus, there was little selection pressure ($d_n/d_s < 1$) on *gag* p17 sequences to change. These values are comparable to values determined for the *vif* gene (239), but higher than that determined for the *vpr* gene (238) and much lower than the values determined for the *env* gene (3) of the same mother-infant pairs.

3.1.6 Analysis of Functional Domains of Gag p17MA Isolates from Mother-Infant Pairs

We next examined the functional domains essential for p17 matrix function in the deduced amino acid sequences from our seven mother-infant pairs (Fig. 12). A number of functions have been proposed for HIV-1 p17 matrix, including targeting of the Gag precursor protein to the plasma membrane (20, 75, 243, 246), virus assembly and release (63, 75), envelope glycoprotein incorporation into virus particles (60, 242), virus entry (241) and localization of the virus preintegration complex to the nucleus of nondividing cells (21, 67, 220). Examination of the membrane binding region including the myristoylation site, that is essential for targeting of Gag to the plasma membrane, showed a high degree of conservation of this site in mother-infant pairs' p17 sequences (Fig 12). Based on mutational analysis, several studies (21, 220) have identified a nuclear localization signal (NLS) in p17 that is required to import the preintegration

complex into the nucleus of nondividing cells. This may be critical in the establishment of HIV-1 infection in quiescent T-cells and macrophages. We examined the NLS (position 25 to 33) in our p17 mother-infant sequences and found that the NLS was conserved in most of the sequences. In addition, the two lysines (K) at position 26 and 27 (21, 220), a tyrosine (Y) at 29 (220) and a lysine (K) at 30 (220) in the NLS were critical for HIV-1 infection in macrophages. We found that the two lysines at positions 26 and 27 were conserved in most of the mother-infant sequences or substituted with compatible arginine, the tyrosine at 29 was highly conserved and the lysine at 30 was present in 5 pairs and substituted with arginine in 2 pairs. The virus assembly domains that require amino acids at positions 55-59 and 84-99 (63) were conserved in most of the sequences. We then examined the amino acid at position 35 (valine) that is essential for incorporation of gp120 into virus particles (63) and found that it was present in almost all of the sequences. Some other sites, including the polymerization site (position 47-59), phosphorylation sites (positions 110-114) and a tyrosine phosphorylation site (position 132) were conserved in most of the sequences. The data on the conservation of the functional domains required for p17 matrix function suggests that a functional matrix is essential for HIV-1 infection and replication in mother-infant isolates.

3.1.7 Analysis of 3-Dimensional Structure of Gag p17MA Isolates from Mother-Infant Pairs

To determine the structure-function relationship within the p17 from seven mother-infant isolates, we examined the 3D structure of Gag p17 matrix as shown in

(Fig 14). The p17 protein forms a compact fold comprised of five α-helices and two β-sheets (Fig 14). A highly basic platform is formed by the two β-sheets and fulfills essential functions including membrane targeting and nuclear localization. We then examined the three classes of mutants in our 166 p17 sequences as described before (23), including proteins containing substitutions at internal residues that grossly distorted the structure and particle formation, proteins with mutations at the trimer interface that allowed correct folding but were defective in virus assembly, and substitutions of basic residues in helix I that produced structurally normal but noninfectious virions. The most severe defects in viral replication were found due to mutations in internal sites at C57 and Y86, C87 and L50, L51 (23). In most of our 166 p17 sequences, the same amino acids at these positions were highly conserved (Fig 12). The next group of amino acids at the p17 trimer interface (positions V46 or T70, E74) which fold like wild type but produce defects in virus assembly (23) were found to be highly conserved in all mother-infant pairs' sequences. The third group of mutants in helix I that caused some relocation of virus assembly and produced structurally normal but noninfectious virions, including outside helix I at R39, R43 were highly conserved in most of the sequences. However, in pair E (mother-E) seven sequences were substituted with lysine (K) at position 39. The analysis of 3D structure-function relationship of p17 sequences reveals that the seven mother-infant pairs' sequences contained the amino acids required for correct folding and wild type activity.

Figure 14: 3D structure p17MA

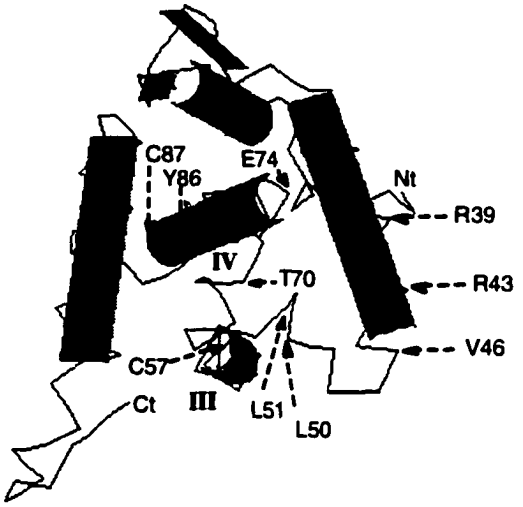


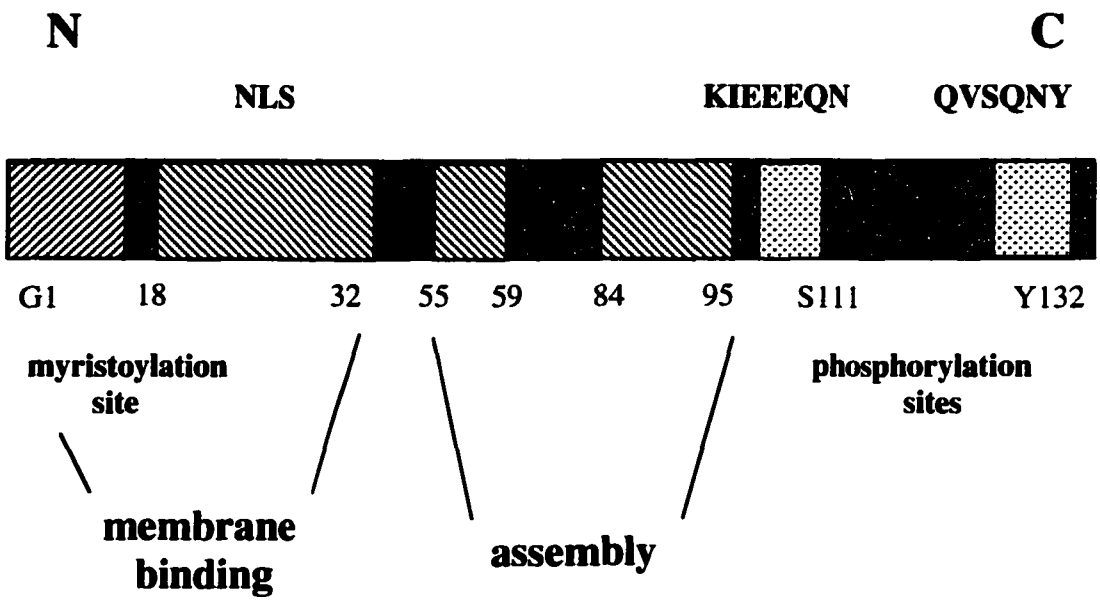
Fig 14: Three dimensional structure of Gag p17 (61) showing five α-helices (I, II, III, IV and V), two β-sheets (above and left of helix I). Nt is NH2-terminus and Ct the COOH-terminus. The amino acid sites shown are required for correct folding, normal p17-p17 interaction for trimer formation and basic residues important for normal and infectious virion production. The amino acids required for correct folding and biological activity were compared with the 166 Gag p17 sequences from our seven mother-infant pairs and were conserved in most of the mother-infant p17 sequences.

3.1.8 Identification of Gag p17 Matrix Protein Motifs Associated with Maternal-Fetal Transmission

Several motifs in *gag* p17 sequences of seven mother-infant pairs that could be associated with transmission were analyzed. Despite pair specific amino acid signature sequences, we found that glutamic acid (E) or aspartic acid (D) at position 55, tyrosine (Y) or phenylalanine (F) at position 79, an aspartic acid (D) or glutamic acid (E) at

position 93 and 102, and an alanine-glutamic acid (AD) at position 122-123 were present in most of the mother-infant pairs' sequences (Fig.12). The glutamic acid (E) at position 93 described by Narwa et al., (156) to be associated with transmitting status was only found in 2 of 7 mother-infant pairs analyzed and was substituted with aspartic acid (D) in 5 pairs. The KIEEEQN motif in the major antibody binding site that was also shown to be associated with transmission (156) was found to be generally conserved in 6 of the 7 pairs with some variability within the motif (Fig 12). However, a valine at position 103 that was previously shown to be associated with non-transmitting status (156) was found in 21 of the 22 mother-infant pair C sequences (Fig. 12). Moreover, the C-terminal 6-mer QVSQNY that is conserved in consB and was shown to be significantly associated with non-transmitting status (156), was conserved in 3 of the transmitting mothers and their infants (pairs C, E, and G) and in infant F sequences. In other mother-infant pairs (A, B, and D) and in the mother of pair F, there were substitutions at one or two positions.

Figure 15: p17MA Functional Domains



3.2 Genetic Characterization of *gag* p17MA Sequences from Infected Mothers Lacking Perinatal Transmission

3.2.1 Phylogenetic Analysis of *gag* p17MA Sequences from Non-transmitting Mothers

Multiple PCRs were performed using PBMC DNA samples from three nontransmitting mothers to obtain 10 to 15 clones from each patient, resulting in 82 *gag* p17 clones, that were sequenced and used for analysis. The neighbor-joining tree in Figure 16 was constructed by combining the 82 *gag* p17 sequences from the three nontransmitting mothers, including the three different sample sets in case of MNT-3, and the reference strain NL4-3 (subtype B). The phylogenetic analysis revealed that the three nontransmitting mothers formed separate well-discriminated clusters that were strongly supported by high bootstrap values. These data indicate that both the sequences from the same nontransmitting mother are closer to each other than other

nontransmitting mothers and the absence of PCR cross-contamination. In case of MNT-3, including the three different sample sets (MNT-3.2, MNT-3.5, and MNT-3.5F), the short branch lengths within the subtree illustrate the observed high level of homogeneity of the sequences at the different sample time points. We also traced a second NJ tree combining the 82 sequences from the nontransmitting mothers, our 154 sequences from mother-infant pairs (85) and other *gag* p17 sequences present in the HIV databases and found that our nontransmitting mothers' sequences were separated from other sequences and grouped with the subtype B sequences (not shown).

Figure 16: Phylogenetic Analysis of *gag* p17MA Sequences from Non-transmitting Mothers

Fig 16: Phylogenetic tree of the 82 *gag* p17 sequences from three infected nontransmitting mothers (MNT-1, MNT-2, MNT-3), including mother-3 with two deliveries (MNT-3.2 and MNT-3.5) and a follow-up sample obtained 1 year after the fifth delivery. The NJ-tree is based on distances calculated between nucleotide sequences from three nontransmitting mothers. Each terminal node of the tree represents one p17 sequence. The numbers on the branch points indicate the percent occurrence of branches over 1000 bootstrap resamplings of the data set. The sequences from each mother formed distinct clusters and are well discriminated and in confined subtrees, indicating that the variants from the same mother are closer to each other than to other mothers' sequences and that there was no PCR cross contamination.

3.2.2 Analysis of Coding Potential of *gag* p17MA Sequence Isolates from Non-transmitting Mothers and Comparison with Transmitting Mothers

Figure 17 shows the multiple alignments of the deduced amino acid sequences of *gag* p17MA from three nontransmitting mothers, including three different sample sets from MNT-3 representing second (MNT-3.2) and fifth (MNT-3.5) deliveries and a follow-up sample 1 year after the fifth delivery (MNT-3.5F) aligned with the subtype B consensus sequence (consB). We found that 69 out of 82 sequenced clones (84%) contained intact open reading frames. We also found that 11 out of 19 clones from MNT-2 contained no initiation codons. In addition, some of these clones carried stop codons. All isolates from MNT-3, representing three different time points, displayed intact open reading frames. In comparison, the previously characterized p17 sequences from seven transmitting mothers displayed a conservation of p17 open reading frames of 88.6%. However, mother-E harbors almost all of the defective variants in the transmitting mother data set (9 out of a total of 10). Therefore, one out of seven transmitting mothers displayed defective sequences, while the remaining six mothers displayed a conservation of p17 open reading frames of 98.9%. The amino acid alignment in Fig. 17 also shows that each nontransmitting mother displayed unique amino acid patterns that were not seen in other mothers. Several amino acids, including lysines (K) at positions 12 and 76, glutamine (Q) or lysine (K) at 15, alanine (A) at 54, and aspartic acid (D) at 102 and 121 were present in most of the nontransmitting mothers' sequences and absent in the previously analyzed transmitting mothers' sequences ($P<0.001$).

Figure 17: Amino Acid Alignments of *gag* p17MA Sequences from Non-transmitting Mothers

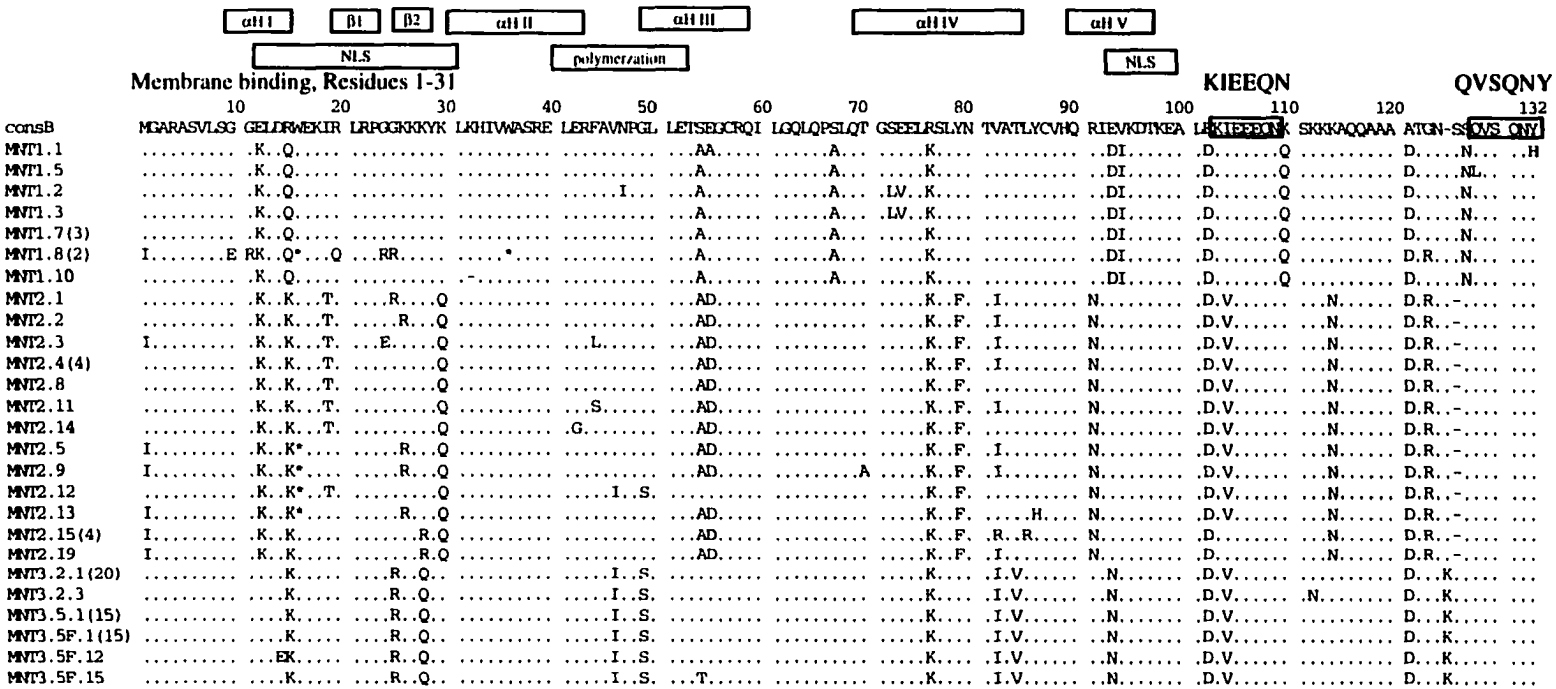


Fig. 17: Multiple alignments of the deduced amino acid sequences of *gag* p17 sequences isolated from three nontransmitting mothers. In the alignment, the top sequence (consB) is the consensus sequence of subtype B as defined elsewhere and MNT-1 and MNT-2 correspond to nontransmitting mother 1 and 2; MNT-3.2, MNT-3.5 and MNT-3.5F correspond to nontransmitting mother-3 after her second and fifth deliveries as well as a follow-up sample 1 year after the fifth delivery. Identical clones are indicated in parentheses. Dots replace amino acids identical to the consB sequence, dashes represent gaps, asterisks represent stop codons. Above the alignment, structural features, comprised of 5 α-helices (αH-I through αH-V) and 2 β-sheets (β1, β2), are shown. In addition, the functional domains, including membrane binding domain, nuclear localization signal, and polymerization site as well as two motifs implicated in transmission are shown.

3.2.3 Variability of *gag* p17 Sequences from Non-transmitting Mothers and Comparison with Transmitting Mothers

We determined the degree of genetic variability of *gag* p17 sequences of nontransmitting mothers' isolates by two different methods. The first method determines the degree of genetic diversity on the basis of pairwise comparison of the nucleotide sequences as well as amino acid sequences. The second method calculates the genetic diversity in each nontransmitting mother by estimating the nucleotide substitutions per generation. As determined by pairwise comparison, the nucleotide sequences of the *gag* p17 gene differed within each nontransmitting mother (MNT-1, MNT-2, MNT-3.2, MNT-3.5, MNT-3.5F) by 1.0, 2.0, 0, 0, 0 % (median values), respectively, and ranged overall from 0% to 5.6% (Table 4). The pairwise comparison of the amino acid sequences revealed variability within each nontransmitting mother (MNT-1, MNT-2, MNT-3.2, MNT-3.5, MNT-3.5F) of 2.3, 3.8, 0, 0, 0% (median values), respectively, and ranged overall from 0% to 7.7% (Table 4). This is in comparison to the seven transmitting mothers' nucleotide distances that range from 0 to 4.0% (1.0% median value) and amino acid from 0 to 8.2% (1.5% median values)(Table 3). Thus the nucleotide and amino acid distances of *gag* p17 sequences were lower in nontransmitting mothers compared to our previously analyzed mother-infant pairs' sequences. The population genetics method that determines the mutation rate per site per generation was employed by using phylogenetic maximum likelihood estimates to compare levels of genetic diversity (θ) within the five sample sets of nontransmitting mothers as shown in Table 5. While MNT-2 shows the greatest estimate of genetic

diversity, the three different sample sets of MNT-3 had the lowest estimates of genetic diversity.

Table 4: Nucleotide and Amino Acid Distances between *gag* p17MA Sequences from Non-transmitting Mothers

	Distances (% nucleotides)			Distances (% aminoacids)		
Mother	**Min**	**Median**	**Max**	**Min**	**Median**	**Max**
Mnt1	0.0	1.0	4.3	0.0	2.3	7.7
Mnt2	0.0	2.0	5.6	0.0	3.8	7.7
Mnt3.2	0.0	0.0	0.5	0.0	0.0	0.8
Mnt3.5	0.0	0.0	0.0	0.0	0.0	0.0
Mnt3.5F	0.0	0.0	0.5	0.0	0.0	1.5
Total[a]	0.0	0.0	5.6	0.0	0.0	7.7

Abbreviations: Mnt nontransmitting mother
[a] Determined for all mothers taken together.

3.2.4 Rates of Accumulation of Non-synonymous and Synonymous Substitutions in Non-transmitting Mothers and Comparison with Transmitting Mothers

The selective pressure on the *gag* p17 sequences from nontransmitting mothers was determined by calculating the ratios of nonsynonymous(d_n) and synonymous(d_s) substitutions using the SNAP algorithm. Analysis showed a d_n/d_s ratio for the five sample sets of nontransmitting mothers (MNT-1, MNT-2, MNT-3.2, MNT-3.5 and MNT-3.5F) were of 0.52, 0.29, 0, 0, 0, respectively (Table 5 A). These values are slightly lower than *gag* p17 sequences form transmitting mothers and their infants (Table 5 B) and comparable to some of the nontransmitting mothers' *vif* and *vpr* sequences (237).

Table 5: Genetic Diversity Estimates and Selection Intensity in the *gag* p17 MA Sequences within the Non-transmitting Mothers and Transmitting Mothers

A	Genetic Diversity			Ratio of syn. to nonsyn. substitutions
Mother	**N[a]**	**θ W[b]**	**θ C[c]**	**dn/ds[d]**
Mnt1	10	0.0205	0.0196	0.52
Mnt2	19	0.0244	0.0474	0.29
Mnt3.2	21	0.0014	0.0012	0
Mnt3.5	15	0.0000	0.0000	0
Mnt3.5F	17	0.0030	0.0034	0
Total	82	0.0098	0.0143	0.162

B	Genetic Diversity			Ratio of syn. to nonsyn. substitutions
Mother	**N[a]**	**θ W[b]**	**θ C[c]**	**dn/ds[d]**
MA	13	0.0244	0.0464	0.12
MB	13	0.0129	0.0239	0.58
MC	12	0.0428	0.0419	0.46
MD	10	0.0148	0.0268	0.38
ME	15	0.0374	0.0819	0.33
MF	12	0.0050	0.0057	0.27
MG	15	0.0054	0.0066	0.56
Total	88	0.0204	0.0333	0.39

[a] Number of p17 clones used in the estimate for each mother
[b] Genetic diversity estimated by the model of Watterson
[c] Genetic diversity estimated by Coalesce
[d] Nonsynonymous-to-Synonymous ratios estimated by SNAP

3.2.5 Analysis of Functional Domains of Gag p17MA Isolates from Non-transmitting Mother and Comparison with Transmitting Mothers

Next, we analyzed the functional domains essential for Gag p17 matrix function in nontransmitting mothers' sequences (Fig. 17). A number of functions have been

proposed for HIV-1 p17 matrix, including targeting of Gag to the plasma membrane, virus assembly and release, envelope glycoprotein incorporation into the virion, and localization of the virus preintegration complex to the nucleus of nondividing cells. Examination of the membrane-binding region, including the N-terminal myristoylation site that is essential for targeting of Gag to the plasma membrane (246), showed a high degree of conservation of this site in the three mothers' p17 matrix sequences (Fig. 17). The nuclear localization signal (NLS) in p17 matrix (position 25 to 33), including three lysines (K) at positions 26, 27 and 30 and a tyrosine (Y) at 29 that is required for import of the preintegration complex into the nucleus of the nondividing cells (21, 220), was fairly conserved in the nontransmitting mothers' sequences as well as mother-infant pairs' sequences. In the MNT-2 sequences, the basic lysine (K) at position 30 was substituted by glutamine (Q). The two domains required for virus assembly positions 55-59 and 84-99 (63) as shown for consensus-B sequences were generally conserved in most of the sequences. However, MNT-2 displayed a substitution of a glutamic acid (E) at position 55 by a compatible aspartic acid (D) and a substitution of an arginine (R) at position 91 by an asparagine (N). We then examined the valine at position 35 that is essential for incorporation of gp120 into virus particles (63) and found that it was present in all of the MNT sequences. In the polymerization site (position 47-59), sequences of MNT-1 and MNT-2 carried a substitution of serine (S) at position 54 by alanine (A) and MNT-3 a substitution of glycine (G) at 49 by a serine (S). This serine (S) at position 54 was highly conserved in our analyzed mother-infant sequences , subtype B sequences and p17 sequences analyzed by others (156). Other sites,

including the phosphorylation (positions 110-114) and a tyrosine phosphorylation (position 132), were conserved in most of the sequences. The functional domains required for Gag p17 matrix function, including targeting of Gag to the plasma membrane (243), virus assembly and release (31, 63, 201), envelope glycoprotein incorporation into the virion (60), and localization of the virus preintegration complex to the nucleus of nondividing cells (21, 220) were generally conserved in MNT and mother-infant pairs' sequences. However, other functional domains, including the polymerization site (positions 47-59) was less conserved in nontransmitting mothers' sequences compared to transmitting mothers' sequences.

3.2.6 Comparison of Motifs in Gag p17MA Isolates from Mother-Infant Pairs and Non-transmitting Mothers

We compared the motifs in the *gag* p17 gene that may be associated with lack of vertical transmission based on our analysis and previous analysis of p17 matrix sequences by others Narwa et. al. (156). The glutamic acid (E) at position 93 that was previously shown to be associated with the transmitting status (156) was not found to be associated with transmitting (Fig. 12) or nontransmitting status (Fig. 17). The KIEEEQN motif (positions 103-109) in the major antibody binding site that was also shown to be associated with transmitting status (156) was generally found to be conserved in our transmitting mothers (Fig. 12) and less conserved in our nontransmitting mothers' sequences ($P<0.001$, chi square test) (Fig. 17). Moreover, a valine (V) at position 104 was found to be associated with nontransmitting status ($P<0.001$) (156) (85). The C-terminal 6-mer QVSQNY that is conserved in cons-B and

shown to be associated with nontransmitting status (156) was conserved in 3 of our 7 transmitting mothers (Fig. 12) but highly conserved in all our nontransmitting mothers' sequences (P<0.001) (Fig. 17). In summary, the KIEEEQN motif was generally conserved in our transmitting mothers' but variable at position 104 due to a valine (V) substitution in our MNT sequences. Moreover, the C-terminal 6 mer QVSQNY that is conserved was significantly associated with our MNT sequences (Fig. 17).

3.3 Genetic Characterization of *nef* Sequences from Mothers and Infants following Perinatal Transmission

3.3.1 Phylogenetic Analysis of *nef* Sequences from Mother-Infant Pairs

Multiple PCRs were performed using PBMC DNA samples from seven mother-infant pairs to obtain 6 to 18 clones from each patient, resulting in 196 *nef* clones that were sequenced and used for analysis. The neighbor-joining tree in Figure 18 was constructed using the 196 *nef* sequence from the seven mother-infant pairs including the set of twins of Mother-H (a total of 15 individuals) and the reference strain NL4-3 (subtype B). The phylogenetic analysis showed that the sequences isolated from the seven mother-infant pairs formed well-discriminated clades that were strongly supported by high bootstrap values (branch labels denote the percentage of 1000 bootstrap replicates) and the absence of PCR cross-contamination. This indicates that isolates form epidemiologically linked individuals are closer related to each other than to isolates from epidemiologically unlinked individuals. The seven mother-infant pair subtrees show that mother and infant sequences were generally separated in distinct

subclusters. However, this was not true for pair-B and pair-C, where the mother and infant sequences were intermingled. The separation of mother and infant sequences in most pairs indicates that the infant was most likely infected by one or few variants found in the mother. Of note, the sequences isolated from the twins of Mother-H (TH1 and TH2) intermingle, therefore both infants were most likely infected by the same or a very similar variant. We also traced a second tree combining our 196 mother-infant pair sequences, with other *nef* sequences present in the HIV-1 databases and found that our mother-infant pair sequences were separated from the other sequences and grouped with the subtype B clade (not shown).

Figure 18: Phylogenetic Analysis of *nef* Sequences from Mother-Infant Pairs

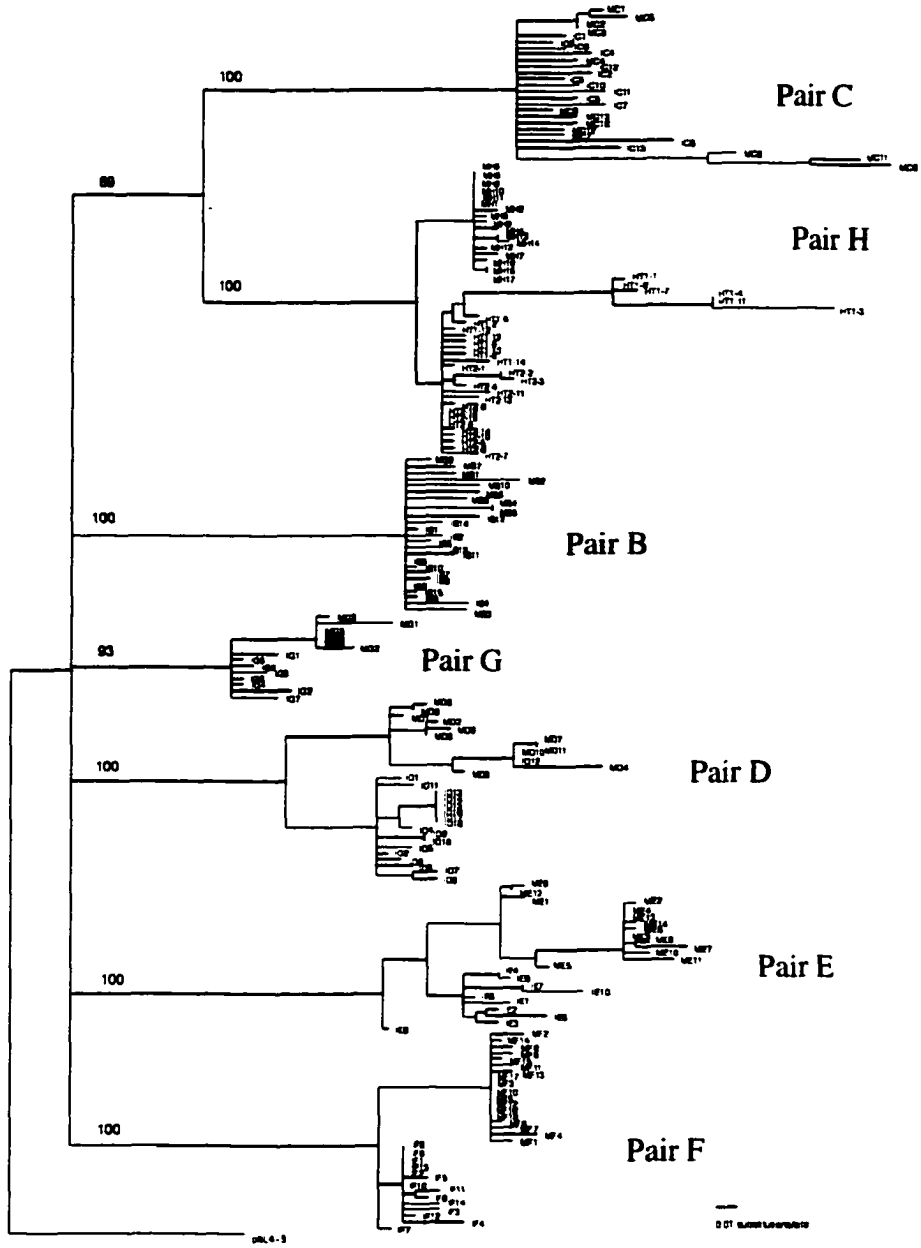


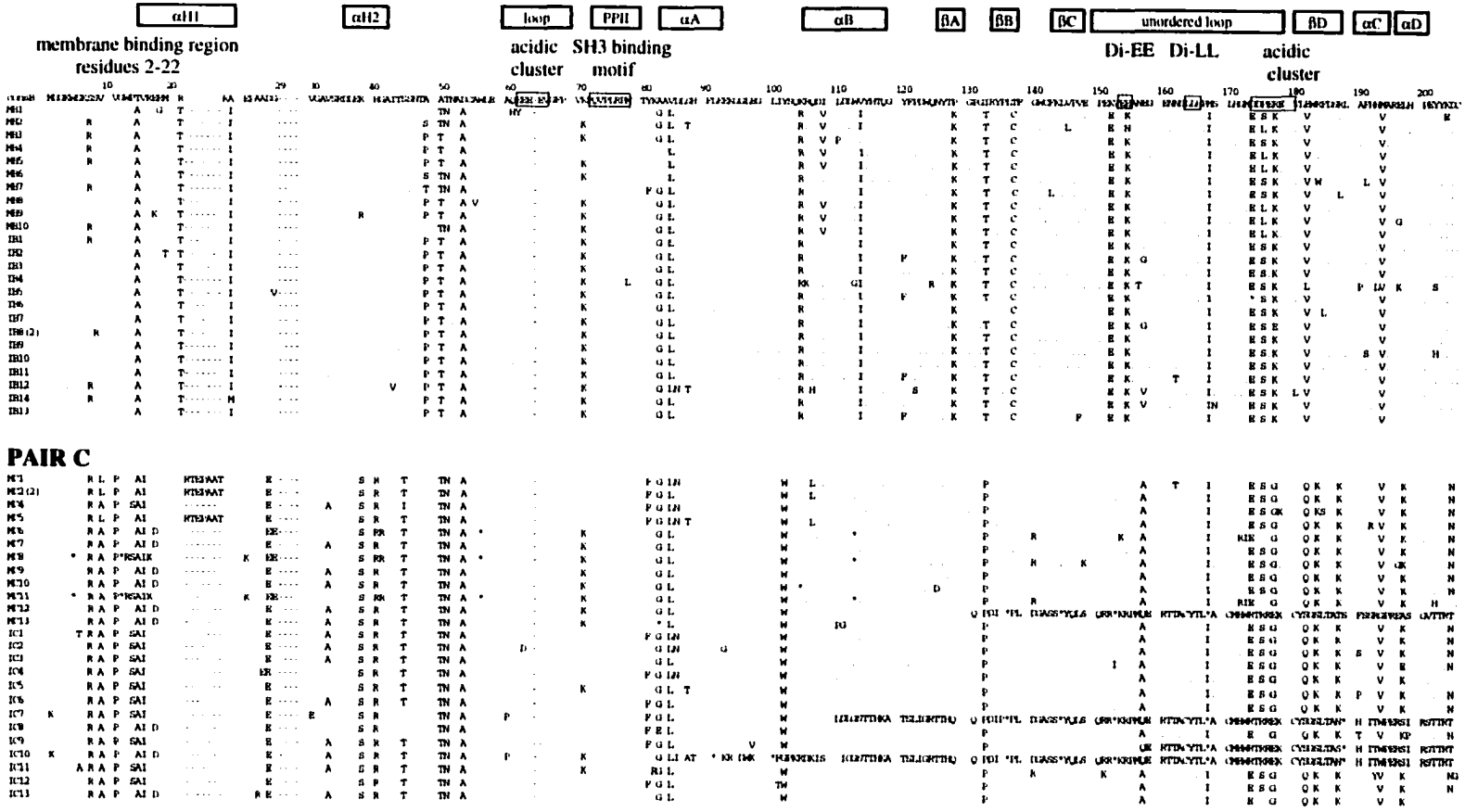
Fig. 18: Phylogenetic analysis of 196 *nef* sequences from seven mother-infant pairs (B, C, D, E, F, G, and H). The NJ-tree is based on the distances calculated between the nucleotide sequences from the seven mother-infant pairs. Each terminal node represents one *nef* sequence. The numbers on the branch points indicate the percent occurrence of branches over 1000 bootstrap resamplings of the data set. The sequences from each mother formed distinct clusters and are well discriminated and in confined subtrees, indicating that the variants from the same mother are closer to each other than to other mothers' sequences and that there was no PCR cross contamination.

3.3.2 Analysis of Coding Potential of *nef* Sequence Isolates from Mothers and Infants

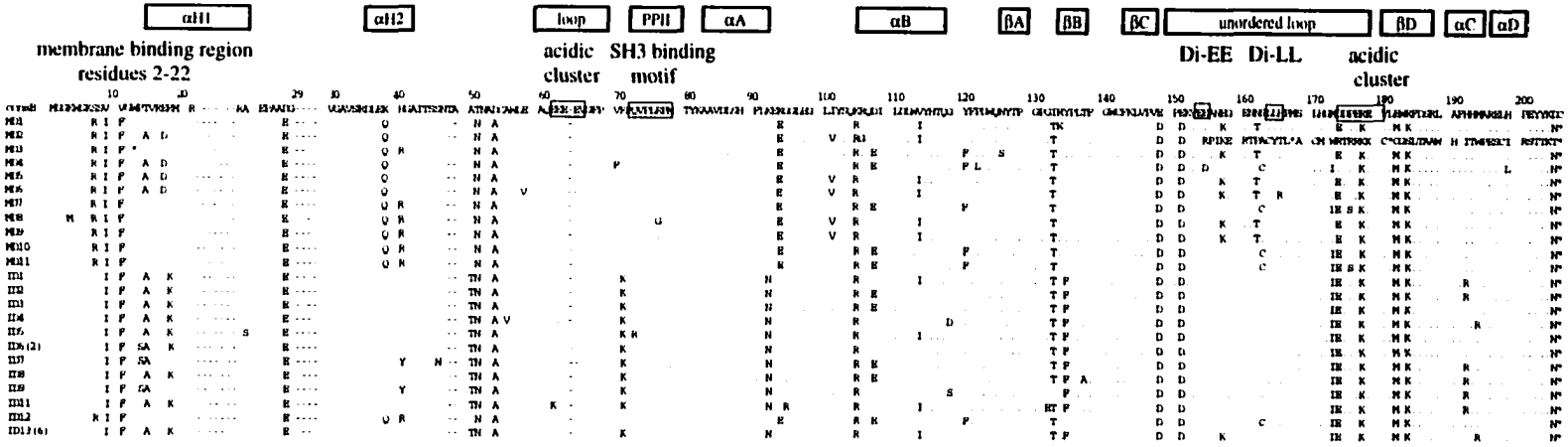
The multiple alignments of the deduced amino acid sequences of the HIV-1 *nef* gene isolated from PBMC DNA of the seven mother-infant pairs are shown in Figure 19. The 196 amino acid sequences were aligned in reference to the subtype B consensus sequence (consB). The coding potential of the *nef* open reading frame was maintained in most of the sequences. Of the analyzed 196 *nef* clones, 170 contained intact *nef* open reading frames that equals to a 86.7% frequency of conservation of intact *nef* open reading frames. Accordingly, the frequency of defective *nef* genes in our seven mother-infant pair sequences was 13.3%. We found that a total of 23 clones contained one or more stop codons or single nucleotide frame shift mutations that resulted in multiple stop codons. All clones from mother-infant Pair-D contained stop codons at the end of the *nef* reading frame. Therefore, these stop codons were not counted as internal stops, which in contrast would render the *nef* gene non-functional. Nine sequences lacked initiation codons. However, there were only 3 sequences that were defective solely on the basis of a missing initiation codon. Most sequences that lacked an initiation codon also had multiple internal stop codons or a frame shift mutation. One striking observation was that the 14 isolates from twin TH1, contained only 7 fully functional clones. The clones either missed initiation codons, contained stop codons or had single frame shift deletions that led to non-functional proteins. This leaves the frequency of intact *nef* open reading frames for this particular infant at 50%. This is especially surprising since both the mother (MH) and the sibling twin (HT2)

have respective frequencies of intact open reading frames of 94% and 93% . We could also document mother-infant pair specific amino acid patterns, that were not seen in epidemiologically unlinked mother-infant isolates. For instance, in the Pair-B sequences, we found the pair specific amino acid substitutions R->T^{20} , A->S or P^{49}, P->K^{129}, and K->E at position 152. In the Pair-C sequences we found the signature sequence SAI or AI (residues 14-16) and a Y->W^{102} substitution. The Pair-D sequences displayed a pair specific signature sequence of IF^{11} instead of the two valines found in the consB sequence. Also, the Pair-H sequences had a pair specific substitution of SS->IN^{46}. The pair specific amino acid patterns enabled us to readily distinguish the mother-infant isolates. Also we found substitutions that were common to all mother infant-pairs' sequences (T->A^{15}) or residues that were variable in all mother-infant pairs' sequences (E^{181} -> V, Q or M).

Figure 19: Amino Acid Alignments of *nef* Sequences from Mother-Infant Pairs

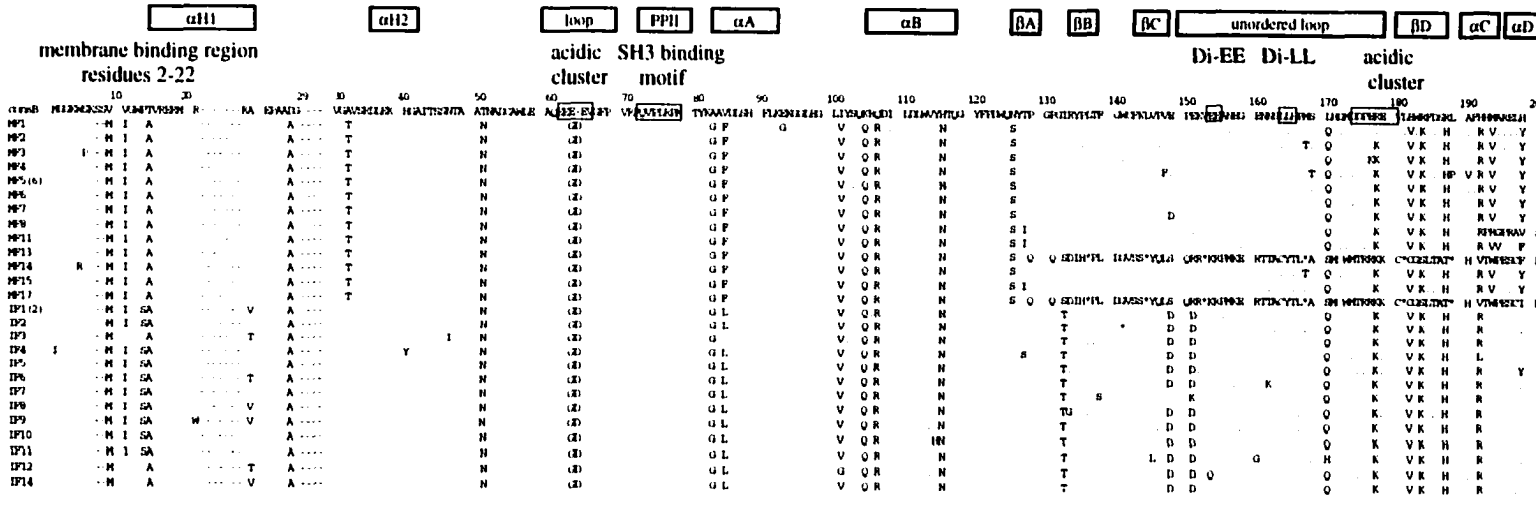


PAIR D


αH1
αH2
loop
PPII
αA
αB
βA
βB
βC
unordered loop
βD
αC
αD
membrane binding region
residues 2-22
acidic
cluster
SH3 binding
motif
Di-EE
Di-LL
acidic
cluster

PAIR E

PAIR F


αH1
αH2
loop
PPII
αA
αB
βA
βB
βC
unordered loop
βD
αC
αD
membrane binding region
residues 2-22
acidic
cluster
SH3 binding
motif
Di-EE
Di-LL
acidic
cluster

PAIR G

PAIR H

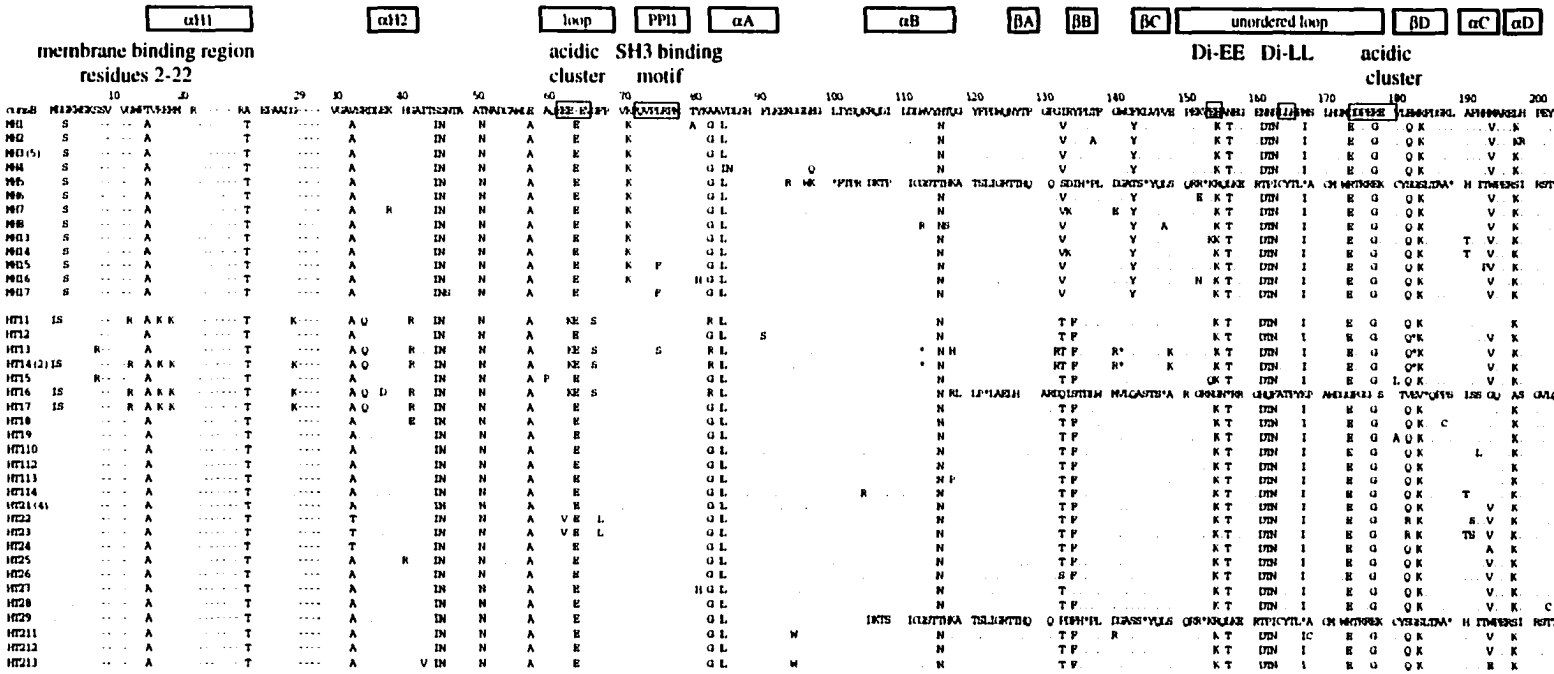


FIG. 19 Multiple alignments of the deduced amino acid sequences encoded by the *nef* gene of HIV-1 from seven mother-infant pairs after perinatal transmission. In the seven mother-infant pairs sequences B, C, D, E, F, G, and H correspond to pair B, C, D, E, F, G, and H respectively. The letter M indicates mother sequences and the letter I indicates infant sequences, and HT1 and HT2 represent the two twins of mother H. The number of identical sequences is indicated in parenthesis. The sequences are aligned in reference to the subtype-B consensus sequence (consB) at the top. Dots replace amino acid agreement with consB, dashes represent gaps, and asterisks represent stop codons. The structural elements of Nef are indicated above the alignment. Boxes around amino acids indicate motifs important for Nef function.

3.3.3 Variability of *nef* Sequences of Epidemiologically linked Mother-Infant Pairs

We determined the degree of genetic variability of the *nef* sequences isolated from the seven mother-infant pairs by two different methods. The first methods determines the degree of genetic diversity on the basis of pairwise comparison of the nucleotide as well as amino acid sequences. The second method calculates the genetic diversity of each mother set, infant set and mother-infant pair by estimating the nucleotide substitution rate per site per generation.

As determined by pairwise distance comparison the nucleotide sequences, the *nef* gene differed within each mother (MB, MC, MD, ME, MF, MG, MH) by 1.86, 2.51, 2.03, 1.27, 0.5, 0.66, 0.59% (median values), respectively, and ranged overall from 0% to 9.24% with a median of 0.85% (Table 6). The nucleotide variability in the infant sets (IB, IC, ID, IE, IF, IG, TH1, TH2) was similar and the sequences differed by 0.99, 2.13, 1.21, 1.65, 1.18, 1.26, 2.43, 0.84% (median values), respectively, ranging from 0 to 4.93% with a median of 1.21%. The variability between mother and infant sets (epidemiologically linked pairs B, C, D, E, F, G, and H) was also of the same order of magnitude at 1.85, 2.39, 4.03, 3.54, 2.63, 2.07, 2.83%, 2.25% (median values), respectively, ranging from 0 to 8.44% with a median of 2.63%. The pairwise comparison of the amino acid sequences revealed a variability within each mother (MB, MC, MD, ME, MF, MG, MH) of 3.4, 4.94, 4.48, 2.11, 0.98, 1.47, 1.51% (median values), respectively, and ranged overall from 0% to 10.89% with a median of 2.01% (Table 6). The amino acid sequence variability in the infant sets (IB, IC, ID, IE, IF, IG,

TH1, TH2) was 1.94, 3.94, 2.49, 3.16, 1.96, 1.96, 4.55, 1.51% (median values), respectively, ranging from 0 to 10.53% with a median of 2.45%. The amino acid sequence variability between mother and infant sets (epidemiologically linked pairs B, C, D, E, F, G and H) was 3.4, 4.43, 8.23, 6.32, 5.41, 3.43, 5.03, 3.52% (median values), respectively, ranging from 0 to 15.79% with a median of 5.79%. Surprisingly, the infants nucleotide sequences, as well as the amino acid sequences, were either as variable as the mothers sequences or displayed higher variability. Only three infants sequence sets (IB, IC and ID) were less variable than the sequences found in their respective mothers. The other infant's sequence sets displayed slightly higher variability (IE and TH2) or much higher variability than the sequence sets from their corresponding mothers. Interestingly, there was a considerable difference in the variability of the sequences isolated from the set of twins of Mother-H. The nucleotide sequences from TH1 differed by 2.43% (median value) and the amino acid sequences differed by 4.55% (median value). This is in contrast to TH2 whose nucleotide sequences differed by 0.84% (median value) and amino acid sequences differed by 1.51% (median value). We also determined the nucleotide distances of *nef* sequences between epidemiologically unlinked individuals and found that epidemiologically unlinked mothers differed from 7.88 to 17.44% (median 11.37%) and infants from 6.33 to 16.32% (median 10.52%). These distances are significantly higher compared with epidemiologically linked mother-infant pairs (Table 6).

3.3.4 Dynamics of HIV-1 *nef* Sequences Evolution in Mother-Infant Pairs

The dynamics of genetic diversity were examined for the *nef* sequences isolated from the seven mother-infant pairs by using the Watterson model, the program Coalesce assuming a constant population size, and a maximum likelihood estimator (EVE) that assumes variable population size. The genetic diversity parameter θ, estimated as nucleotide substitutions per site per generation, and estimates of the growth rate for each patient's HIV-1 population are shown in Table 7. The levels of genetic diversity among mother-infant pairs, as estimated by Watterson and Coalesce methods, ranged from 0.013 to 0.034 and 0.016 to 0.089, respectively. The genetic diversity estimated by the Vasco method (EVE) that also calculated the growth rates ranged from 10.5 to 318.5 for the mother's sequence sets and from 20.5 to 329.0 for the infant's sequence sets. The viral populations found in the infants displayed higher genetic diversity in comparisons to the viral populations found in the mothers. The population growth rate ranged in the mothers sets from almost no growth (g= -0.25) to a rapidly expanding population (g= -94.25). However, only Mother-B displayed such a rapidly expanding population. In comparison, the viral populations harbored by Mother-H had a growth rate of -18.25, MF a growth rate of -10.25 and the remaining mothers viral populations showed almost no expansion. The infants populations displayed a much higher rate of growth of -11.25, -96.75, -3.25, - 5.25, -60.25, -56.75, -3.0, -5,25 for IB, IC, ID, IE, IF, IG, HT1 and HT2, respectively. The growth rates with negative values indicate virus population expansion (218). The total level of genetic diversity and growth rate for the mothers were 68.7 and -18.57, respectively compared with the infants of 93.8 and -

30.22, respectively (Table 7). This is also reflected in the Coalesce total estimates of 0.034 for the mothers and 0.045 for the infants. Therefore, the viral populations in the infants displayed slightly higher genetic diversity as well as higher growth rates than the populations in the mothers. Interestingly, the HIV-1 population in Infant-C had the highest growth rate (g=-96.75) whereas its mother had only a growth rate of −1.0. Also, Infant-G whose mother had the lowest growth rate (g= -0.25), displayed a very high growth rate of -56.75.

Table 6: Nucleotide and Amino Acid Distances between *nef* Sequences from Mother-Infant Pairs

		% Distances[a]								
		Within mother set			Within infant set			Between mother and infant set		
Sequence	Pair	min	median	max	min	median	max	min	median	max
Nucleotide	B	0.16	1.86	3.12	0.16	0.99	3.55	0.82	1.85	3.93
	C	0.00	2.51	9.24	0.68	2.13	4.07	1.04	2.39	8.44
	D	0.00	2.03	4.52	0.00	1.21	4.93	0.00	4.03	6.66
	E	0.00	1.27	3.66	0.18	1.65	3.44	2.01	3.54	5.93
	F	0.00	0.50	1.72	0.16	1.18	2.10	1.88	2.63	4.25
	G	0.00	0.66	3.03	0.33	1.26	1.89	1.34	2.07	4.82
	H	0.00	0.59	1.37	-	-	-	-	-	-
	HT1[c]	-	-	-	0.00	2.43	4.68	1.71	2.83	6.50
	HT2[c]	-	-	-	0.00	0.84	2.28	1.54	2.25	3.79
	Total[b]	0.00	0.85	9.24	0.00	1.21	4.93	0.00	2.63	8.44
Amino Acid	B	0.49	3.40	5.83	0.00	1.94	8.74	1.46	3.4	8.74
	C	0.00	4.94	10.89	0.99	3.94	7.88	1.48	4.43	10.34
	D	0.00	4.48	8.96	0.00	2.94	8.96	0.00	8.23	10.95
	E	0.00	2.11	6.84	0.53	3.16	10.53	3.16	6.32	15.79
	F	0.00	0.98	9.31	0.00	1.96	4.41	3.43	5.41	12.25
	G	0.00	1.47	6.86	0.00	1.96	3.43	2.54	3.43	8.82
	H	0.00	1.51	3.02	-	-	-	-	-	-
	HT1[c]	-	-	-	0.00	4.55	9.60	3.02	5.03	11.62
	HT2[c]	-	-	-	0.00	1.51	4.52	2.51	3.52	7.04
	Total[b]	0.00	2.01	10.89	0.00	2.45	10.53	0.00	5.79	15.79

[a]Expressed as percent nucleotides (for nucleotide sequence) or percent amino acids (for amino acid sequence).
[b]Calculated for all the pairs taken together. [c]Mother H twins

3.3.5 Rates of Accumulation of Non-synonymous and Synonymous Substitutions

Several studies suggested that a ratio of nonsynonymous (d_n) to synonymous (d_s) substitutions (d_n/d_s) of more than 1 indicates positive selection pressure by the immune system for escape variants (131, 245, 247). Others have argued that the value of the ratio is less important than a shift from the background ratio (193). The d_n/d_s ratios, calculated by the method of Nei and Gojobori (157), for each patient set (mothers sets, infants sets) are shown in Table 7. Analysis showed that the d_n/d_s ratios for the seven mother sample sets (MB, MC, MD, ME, MF, MG and MH) were 0.48, 0.39, 0.87, 0.7, 0.22, 1.27, 0.63 respectively. The d_n/d_s ratios for the eight sample sets from the corresponding infants (IB, IC, ID, IE, IF, IG, TH1, TH2) were 0.43, 0.41, 0.56, 0.7, 0.53, 0.53, 0.69, 0.32 respectively. The d_n/d_s ratios between mother and infant sets (epidemiologically linked pairs B, C, D, E, F, G and H) were 0.53, 0.38, 0.83, 0.7, 0.49, 0.5, 0.56 respectively. Therefore, there was little selection pressure ($d_n/d_s<1$) for MB, MC, ME, MD and MH as well as all the corresponding infants. A value >1, indicating strong directional selection pressure, was only associated with the sample set of Mother-G. However, due to the restriction in sample size of this mother (6 sequences) one has to be cautious since this may be an artefact.

Table 7: Genetic Diversity Estimates and Selection Intensity in the *nef* Sequences within Mother-Infant Pairs

	Mother sets						Infant sets					
Pair	**N[a]**	**θ_W[b]**	**θ_C[c]**	**θ_V[d]**	**g[e]**	**dn/ds[f]**	**N[a]**	**θ_W[b]**	**θ_C[c]**	**θ_V[d]**	**g[e]**	**dn/ds[f]**
B	10	0.024	0.06	318.5	-94.25	0.48	15	0.021	0.042	52.5	-11.25	0.43
C	13	0.034	0.052	34.5	-1.00	0.39	13	0.027	0.089	329.0	-96.75	0.41
D	11	0.019	0.029	37.0	-4.75	0.87	18	0.021	0.035	33.0	-3.25	0.56
E	14	0.019	0.036	10.5	-1.25	0.70	10	0.022	0.046	20.5	-5.25	0.70
F	18	0.013	0.022	28.5	-10.25	0.22	14	0.017	0.042	132.5	-60.25	0.53
G	6	0.015	0.016	12.0	-0.25	1.27	8	0.014	0.031	123.5	-56.75	0.53
H	17	0.013	0.026	40.0	-18.25	0.63	-	-	-	-	-	-
HT1	-	-	-	-	-	-	14	0.023	0.045	37.5	-3.0	0.69
HT2	-	-	-	-	-	-	15	0.014	0.027	22.0	-5.25	0.32
Total[g]	89	0.02	0.034	68.7	-18.57	-	107	0.02	0.045	93.8	-30.22	-

Abbreviations and variables: θ = genetic diversity; dn/ds = selection intensity

[a]Number of *nef* clones sequenced

[b]Genetic diversity estimated by the method of Watterson

[c]Genetic diversity estimated by Coalesce

[d]Genetic diversity estimated by the method of Vasco (EVE)

[e]Growth rate estimated by the method of Vasco (EVE)

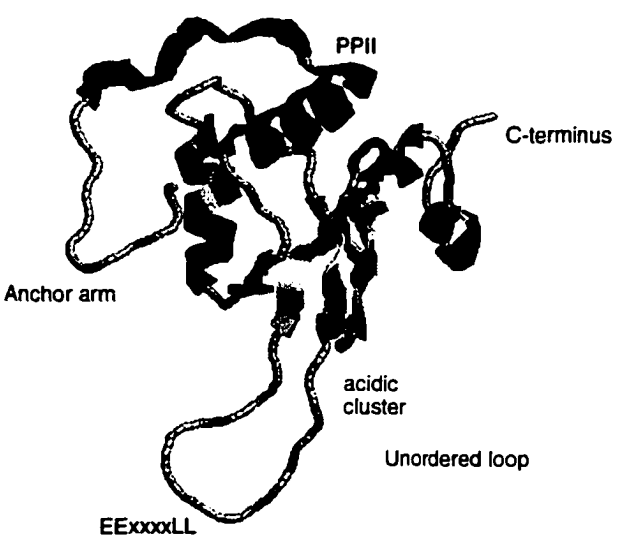
[f]Nonsynonymous to synonymous substitution ratios estimated by SNAP

[g]Totals are calculated as the average of all values

3.3.6 Analysis of Functional Domains of *nef* Sequence Isolates from Mother-Infant Pairs

Although it has been difficult to solve the full-length Nef structure using NMR and x-ray crystallography, fragments of the protein have been structurally characterized (Figure 20). Nef is a N-terminally myristoylated 206 aminoacid protein (~27 kDa), that is phosphorylated at serine and threonine residues (18, 234) and is associated with cellular membranes (158). In addition, Nef forms homomeric oligomers and intramolecular disulfide bonds (83, 244). The HIV-1 Nef protein can be divided into three general regions: 1.) a structurally flexible N-terminal arm of ~70 residues, 2.) a conserved, highly structured and folded core domain of ~120 residues (residues ~70-190) and 3.) a short C-terminal arm of ~ 10 residues.

Figure 20: 3-D Structure Nef



3.3.6.1 Analysis of nef Motifs involved in Membrane Anchoring

Since Nef acts as a connector between target molecules on the cell surface (CD4 and MHC-I) and the cellular protein trafficking machinery, localization of Nef at the cell membrane is essential for the main functions of Nef (8) (158) . This is also true for the interactions of Nef with signaling receptor bound kinases. It was shown that mutations in Nef that affect the regions necessary for membrane localization affect HIV-1 pathogenicity (8). Furthermore, cell membrane association of Nef is required for adequate incorporation of Nef into maturing virions and may therefore be important for virion infectivity (222). First we examined the N-terminal membrane anchor arm that consists of a myristoyl group that is postranslationally added to the first glycine (G^2), a patch of basic amino acids (residues 4-22), and an amphiphatic helix (αH1, residues 17-22) with several arginine residues on the outside (12, 14) that interact with the negatively charged inner side of the cellular membrane. Examination of our 7 mother-infant pairs' sequences revealed that the myristoyl acceptor site (G^2) was highly conserved in all mother-infant sequences, except for five sequences in the sample set from infant HT1 (Fig. 19). These five sequences also lacked initiation codons and two of the five sequences contained several stop codons. The patch of basic amino acids (residues 4-22) was mostly conserved in the mother-infant sequences and the few observed substitutions are likely not to grossly change the basic character of this region. Only the sequences from pairs C and D displayed considerable variation where the serine at position eight (S^8) was substituted with the basic arginine (R), the valine at position 10 (V^{10}) with either an alanine (A), leucine (L)(pair-C) or an isoleucine

(I)(pair-D), and the V^{11} was substituted with a proline (P) (pair-C) or phenylalanine (F) (pair-D) Overall Pair-C displayed the most variation in the basic membrane binding region (residues 4-22) with multiple substitutions and a seven amino acid insertion in several Mother-C sequences. In addition, several isolates form Mother-G had a five amino acid insertion. However, it is not unusual to find such length polymorphism in the N-terminal region of Nef. The observed substitutions in this region did not grossly change the overall charge of this region, the opposite some even increased the number of positive charges, which would increase the affinity of this part of Nef to the cell membrane. The residues K^4, K^7, R^{17}, R^{19}, R^{22} that are believed to be required for incorporation of Nef into new virions (222) were all well conserved in most mother-infant pair isolates. Several studies indicate that one or more of the serine residues (S^6, S^8, S^9) and the threonine residues (T^{15}, T^{80}) may be phosphorylated by cellular serine/threonine kinases (18, 35, 72, 234). We found that the T^{80}, that is a potential phospholipase C phosphorylation site, was highly conserved in all mother-infant isolates. This is in contrast to the T^{15} that was substituted with an alanine (A) in all mother-infant sequences except for several pair-D sequences. In addition, the S^8 in several sequences from pairs-B and E and all sequences from pair-C and mother-D, was substituted with an arginine (R). Although the biological role of Nef phosphorylation remains to be determined, it was speculated that phosphorylation of the Nef N-terminus may counteract membrane association and may therefore be important for internalization of Nef and its associated proteins. The N-terminal region of the first α-helix (αH1) was well conserved in five out of the seven mother-Infant pairs. It was

shown that deletion of the residues 17-22 resulted in decrease of Nef phosphorylation (14). In this region only some Pair-C isolates, Pair-D isolates and several isolates from HT1 displayed substitutions at positions 18 and 19. The region in between αH1 and αH2 was fairly conserved in most pairs. Either there were no changes or substitutions with neutral or charged amino acids that would only enhance Nef's ability to interact with the cell membrane.

3.3.6.2 Analysis of nef Motifs involved in CD4 Downmodulation

Nef mediated down modulation of CD4 is accomplished by rerouting of expressed CD4 at several levels. As mentioned earlier, Nef is targeted to the cell membrane as its primary site of action. Here it binds to the cytoplasmic tail of CD4 and recruits cellular adaptor complexes that lead to the internalization and ultimately to degradation of CD4 (10, 77, 81, 125, 129, 165). The CD4 receptor, similar to many other membrane bound receptors, contains a dileucine motif at the membrane proximal region of the cytoplasmic tail (5). Cellular adaptor proteins (APs) may bind to this LL-based endocytosis signal. After binding, the APs recruit cytosolic clathrin to the cell membrane to form clathrin coated pits (CCP) that are internalized and form endosomes. Nef can alter CD4 internalization control by displacing a regulatory, CD4 associated kinase and therefore Nef makes the CD4 endocytosis signal available for APs. Therefore, Nef may override the internal control of CD4 internalization. In addition, Nef itself contains a dileucine motif (LL^{165}) (19) that can recruit APs to the cell membrane and therefore enhance internalization of CD4 as well as MHC-I. A direct CD4 binding site in Nef is comprised of the residues W^{57}, L^{58} and E^{59} (79, 81). This

motif was highly conserved in five out of the seven mother-infant pairs (pairs B, C, D, F, and G) (Fig. 19). However, in mother-infant pair-H the glutamic acid (E^{59}) was substituted with alanine (A) and in pair-E the leucine (L^{58}) was substituted with the compatible amino acid valine (V). The substitution of the acidic E^{59} with a basic alanine in pair-H may render the CD4 binding motif in the sequences of Pair-H non-functional. Nef contains a dileucine based endocytosis signal ($E/D^{160}xxxLL^{165}$) (19) that can recruit APs to the cell membrane and therefore enhance internalization of CD4 as well as MHC-I. We found that this motif was highly conserved in all our seven mother-infant pair sequences (Fig. 19). However, Pair-H had a pair specific signature sequences of DTN^{163} in the variable part of this motif just upstream of the LL^{165}, thus it should not interfere with the functionality of this motif. Additional AP binding domains are comprised of the acidic cluster (D^{174}-E^{179}) as well as a DD^{175} motif that interacts with the regulatory subunit of the v-ATPase and may facilitate AP-2 recruitment (133). The acidic cluster was variable in most of the mother-infant pair isolates however the observed variations did not grossly change the character of the motif and therefore should not alter its functionality. For instance, in most sequences of mother-infant pairs B, C, D, E, and H, the D^{174} was replaced by the acidic amino acid E. Also, the R^{178} was substituted by either the basic K (pairs B, D, E, F) or the neutral G (pairs C, H). Although we observed variation in this motif, the substitutions did not grossly change the character of the motif and therefore should not alter its functionality. Interestingly, the glutamic acid residue at 177 (E^{177}) was highly conserved in all isolates from the seven mother-infant pairs. It was demonstrated that a $E->G^{177}$ mutation resulted in a

dominant-negative Nef protein that decreases HIV-1 production as well as infectivity (53). Nef also interacts with the β-subunit of the COP I coatomer (16, 76) that is involved in the ER-Golgi and endosomal sorting of proteins. This interaction may facilitate the degradation of target proteins bound to Nef by routing the complex to lysosomes. Mutating of a di-acidic EE^{155} motif disrupts the COP I binding and prevents Nef induced CD4 lysosomal targeting (165). When examined we found that the EE^{155} motif was highly conserved in pairs C, D, E, F, and G. In contrast, in all of the isolates from pair-B and pair-H the E^{154} was conserved but the E^{155} was substituted with the basic amino acid K.

3.3.6.3 Analysis of nef Motifs involved in MHC-I Downmodulation

The down modulation of MHC class I follows a similar principle as the down modulation of CD4. Nef interacts with surface expressed MHC-I molecules and governs the internalization of MHC-I and targeting of MHC-I to the trans-golgi network and ultimately to degradation (78, 137, 187). Of note, the motifs and residues in Nef responsible for MHC-I down modulation are largely independent of the motifs responsible for CD4 down modulation. The motifs in Nef implicated to be important for MHC-I downregulation are in the N-terminal anchor arm of Nef (helix αH1) as well as in the N-terminal part of the folded core. A proximal flexible loop (residues 58-69) contains an acidic cluster (residues 62-65) and is required for MHC-I down modulation. In addition the proline-rich sequence (P^{69}xxPxxPxxP^{78}) that forms the polyproline helix (PPII, residues 70-77) is essential for MHC-I down modulation. This motif is also

important in altering cell signaling pathways which will be described in more detail below. First we examined the residues of αH1 (residues 15-22), that overall were mostly conserved, which was described earlier. As mentioned before, deletion of the residues 17-22 resulted in decrease of Nef phosphorylation (14) but it was also reported that αH1 may interact with Src kinases without activating them (137). A recent report showed that a methionine in the helix (M^{20}) is indispensable for MHC-I down modulation by Nef (7). We found that the M^{20} was conserved in all mother-infant pair sequences. The four E^{62}-E^{65} , in the acid cluster, were highly conserved in the isolates from five of the seven mother-infant pairs (pairs B, C, D, G and H). Pair-H had a pair specific insertion of an additional glutamic acid (E), adding to the acidic character of the region, however infant HT1 had 5 isolates with a substitution of the E with the basic amino acid K. We also found variability in Pairs E and F where the E^{64} was substituted with a compatible D or a neutral G and E^{65} was substituted again with an acidic D. Therefore, the loop encompassing the acidic cluster was functionally well conserved in all mother-infant pair isolates.

3.3.6.4 Analysis of nef Motifs involved in T-cell Activation

Nef interacts with a multitude of host factors that are part of the cellular signaling machinery. Analysis of a large number of mutant Nef proteins demonstrated that the effects of Nef on CD4 expression and on TCR/CD3 signaling are separable. The ability of Nef to block CD3 signaling was selectively abolished by mutations in the central part of the Nef protein (104). The proline rich region P^{69}xxPxxPxxP^{78} forms a left-handed polyproline type II helix and is the central point in these interactions. In

addition, the core α-helices αA and αB form a cavern surrounding the PxxP motif to position this motif correctly for the interaction with the SH3 domains of the kinases. The PxxP motif lets Nef directly bind to SH3 domains of cellular kinases, such as Hck and Lyn (128, 129, 177). Nef binds to the SH3 domains of Hck and Lyn kinases with high affinity, and with lower affinity to Lck, Fyn, and Src. Nef activates Hck, does not affect Lyn or Src and decreases the activity of Lck and Fyn. Therefore, the tissues distribution of the individual kinases may be the critical. The kinase Hck is restricted to macrophages, in contrast to T-cells where Nef may interact with the lower affinity targets Fyn, or Lck. Nef may also interact with the SH3 domain of the guanine nucleotide exchange factor Vav (54). Moreover, it was reported that Nef may interact with members of other major signaling pathways such as PAKs (182), and protein kinase C (PKC) (199). In addition, Nef may stimulates TCRζ (93, 232, 233) that results in upregulation of Fas ligand. This would trigger apoptosis in CTLs which are attacking HIV-1 infected T-cells. The central point in these interactions is the proline rich region P^{69}-P^{78} , we found that the four prolines were highly conserved and variability was restricted to a substitution of R^{71} with the compatible K in sequences from pairs B, C, infant-D and mother-H and a substitution with a T in isolates from infant-E. The two alpha-helices (αA and αB), which form a cavern that positions the PxxP motif correctly for the interaction with cellular kinases, were maintained in all mother-infant isolates as determined by the program SSpro (Protein secondary structure prediction based on Bidirectional Recurrent Neural Networks) (11).

3.4 Biological Characterization of *env* gp120 Isolates Associated with and without Vertical Transmission

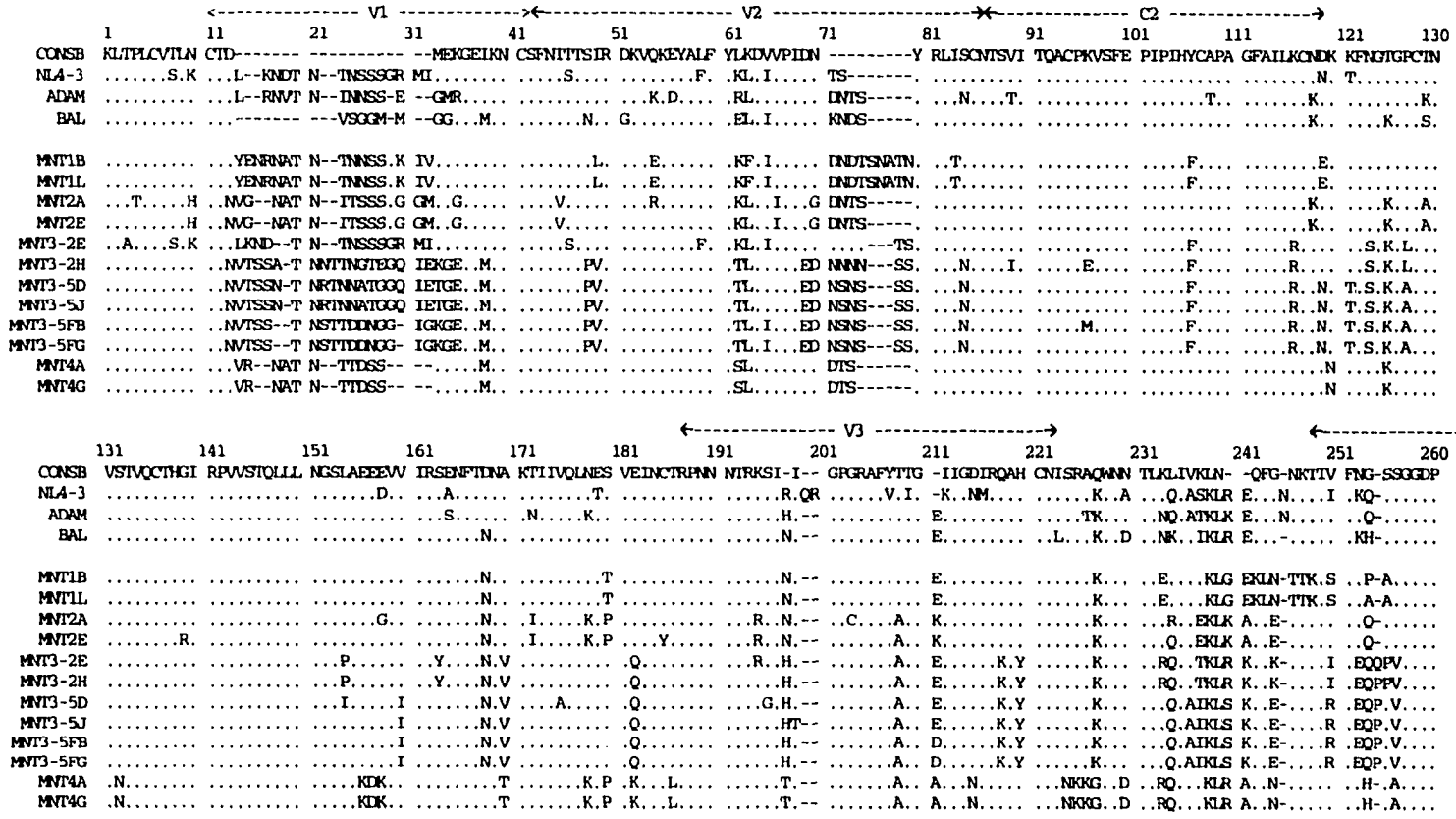
It was previously demonstrated that an HIV-1 minor V3 region genotype of an infected mother is selectively transmitted to her infant, predominating initially as a homogeneous population and then becoming diverse as the infant grows older (3). In addition, the biological characterization of V3 region isolates from mothers and infants associated with perinatal transmission indicated that these chimeras could not replicate in T-cell lines but replicated in PBL and MDM at a lower level compared with the laboratory adapted macrophage-tropic (R5) isolates (141). Furthermore, Matala and Husain et al. (144) characterized the *env* gp120 sequences from non-transmitting mothers and five mother-infant pairs, respectively. While distinct domains within gp120 are involved in specific inteactions, it becomes increasningly clear that complex interactions between these domains are essential for the full range of biological properties required for productive infection. Here we compare the replication kinetics, cellular tropism and co-receptor utilization of gp120 isolates from transmitting mothers and their infants with gp120 isolates from non-transmitting mothers. Two gp120 variants from each transmitting mother (M) and their infants (I) as well as two gp120 variants from each non-transmitting mother (Mnt) were selected to be reciprocally inserted into the infectious molecular clone NL4-3. The selected mother-infant and non-transmitting mothers' clones are shown as an amino acid alignment in Figure 21 and Figure 22, respectively. The top sequences represent the consensus sequence for subtype B followed by the lymphotropic variant NL4-3, and the macrophage-tropic

variants BAL and Ada-M. Several aminoacids in the V1/V2 regions have been associated with replication efficiency and spread in macrophage culture. In particular, a leucine (L) at position 59 (71 for mother-infant pairs) and a valine-valine (VV) or isoleucine-valine (IV) motif found at positions 65 and 66 is present in the selected Mnt sequences, except in the isolate Mnt3.2e that displayed a phenylalanine (F) at position 59. The lymphotropic variant NL4-3 also displays a F at this position in contrast to the macrophage-tropic variants Ada-M and BAL. The mother-infant gp120 isolates have the leucine conserved at this position. The presence of a threonine (T) at position 46 (position 58 for mother-infant isolates) and a serine (S) at position 79 (position 90 for mother-infant isolates) that were associated with the ability to spread in macrophage culture and were highly conserved among all our mother-infant and non-transmitting mother clones. An aspartic acid (D) at position 16 was absent in most of our clones, which was reported to confer lymphotropism. An earlier study identified specific amino acid residues in the C4/V5 region, which were able to confer lymphotropism, including the lysine (K) at position 324 (position 331 for mother-infant isolates) that is not conserved in most of our clones except the isolates from IE-2. The serine (S) at position 335 (position 342 for mother-infant isolates) is present only in the Mnt2 isolates as well as Mnt3.2e in the non-transmitters. Similarly, the serine is only present in the isolates from IC.

The cloning of the patients' gp120 into the infectious molecular clone NL4-3 involved one sub-cloning step. First, the patients *env* gp120 fragment (NsiI-BglII) was transferred into the EcoRI-BamHI sub-clone. Then, the EcoRI-BamHI fragment was

transferred into pNL4-3 (Fig. 9). All transfers were confirmed by DNA sequencing to verify the correct orientation and maintenance of the open reading frame. The full length gp120 chimeras were transfected into HeLa cells by electroporation, and virus production was measured by reverse transcriptase assay. We found that all our gp120 chimeras replicated in HeLa cells (Fig. 23 A and B). To avoid tissue culture adaptations, the generated virus was not passaged or co-cultured. However, in the following infection experiments we were unable to clearly show replication for the following chimeras IG12, Mnt1-l, Mnt2-h, Mnt3.2-e, Mnt3.2h and Mnt3.5-j. Since these chimeras did not show defects as determined by nucleotide sequencing and did replicate in HELA cells, the possibility of defects that may have been introduced during the reciprorcal transfer process is unlikely. Thus, these chimeras may represent variants that replicate at a very low level, and/ or use co-receptor other than CXCR4 and CCR5. Which is under further investigation. The chimera ME17 was excluded from the further experiments due to concerns of possible contamination. Further, experiments will be conducted to verify the validity of the results. Virus stocks were stored at -70° and used for infection experiments.

Figure 21: Multiple Alignment of the deduced *env* gp120 Amino Acid Sequences from Non-transmitting Mothers



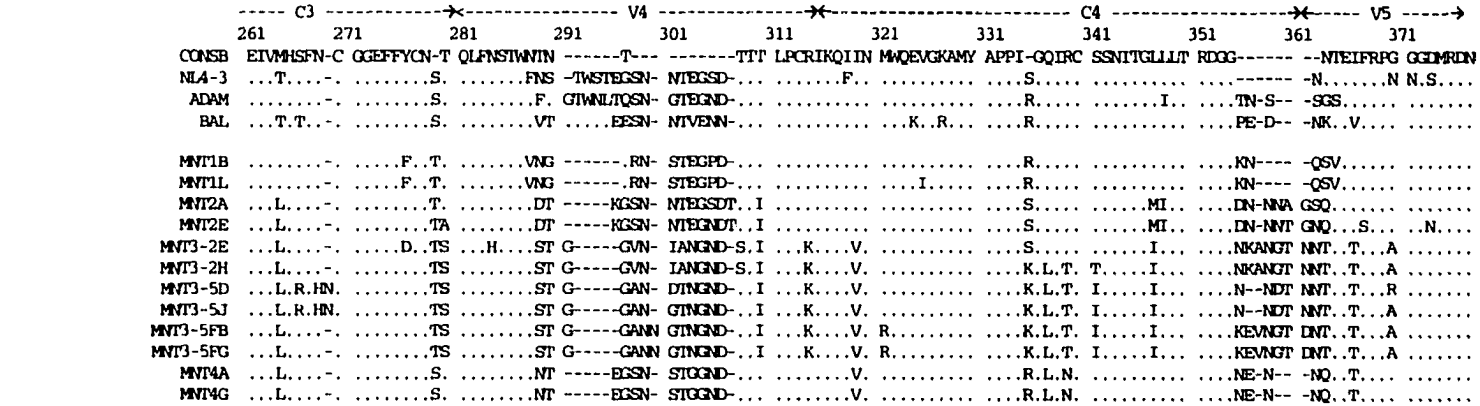
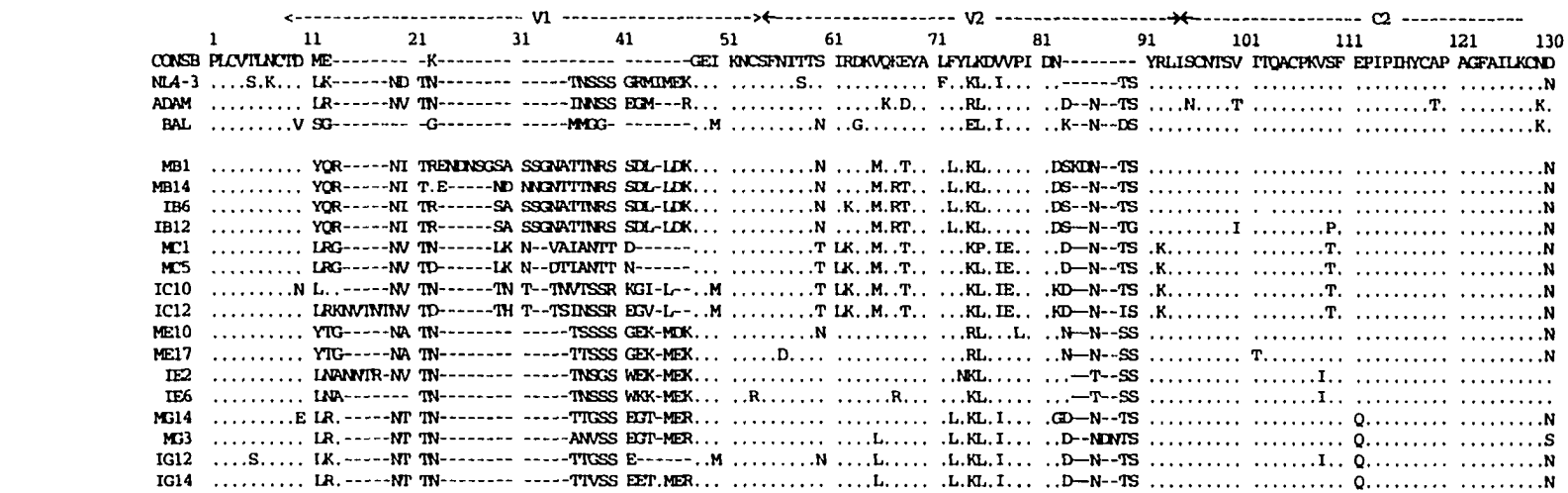


Figure 22: Multiple Alignment of the deduced *env* gp120 Amino Acid Sequences from Mother-Infant Pairs



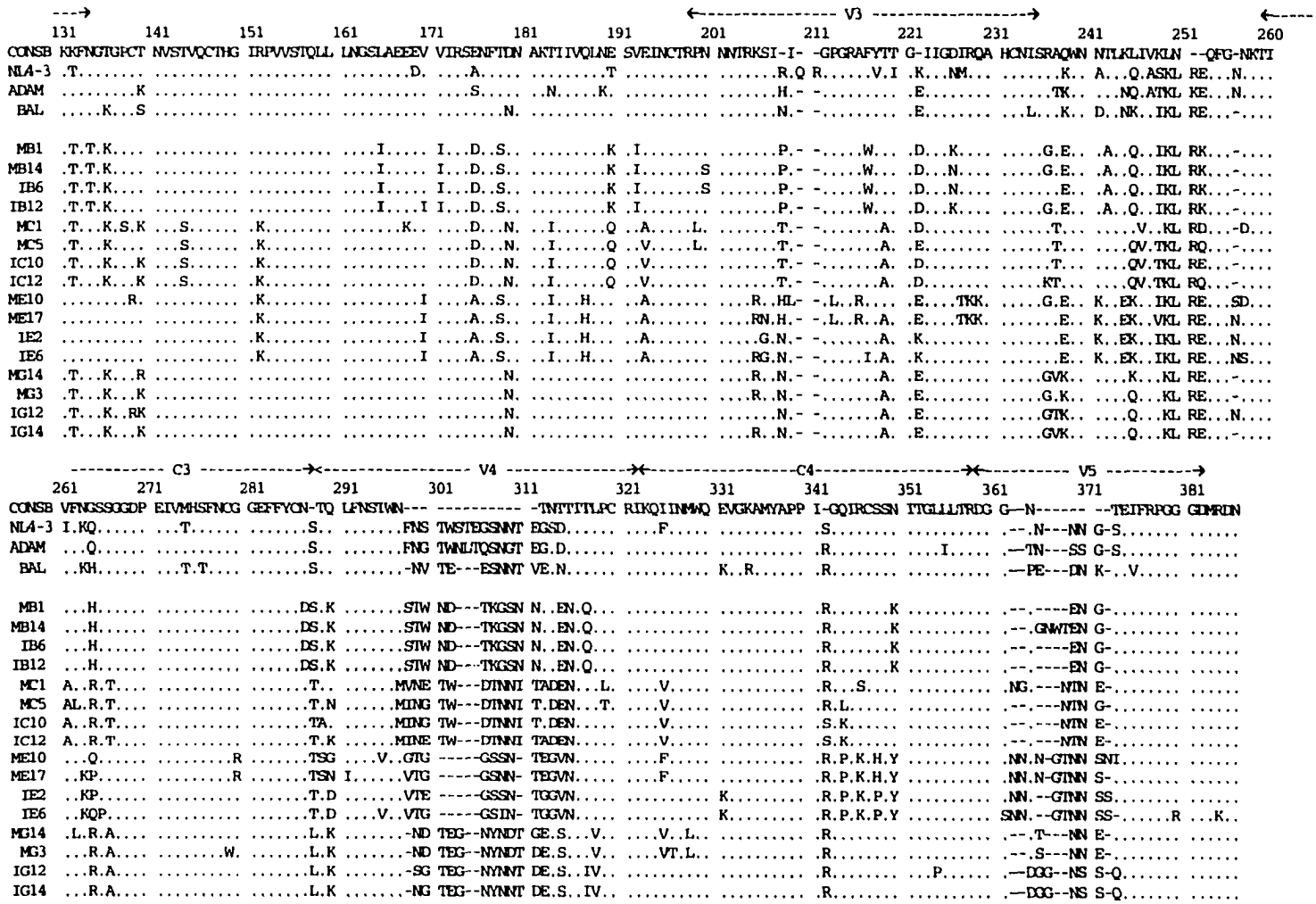
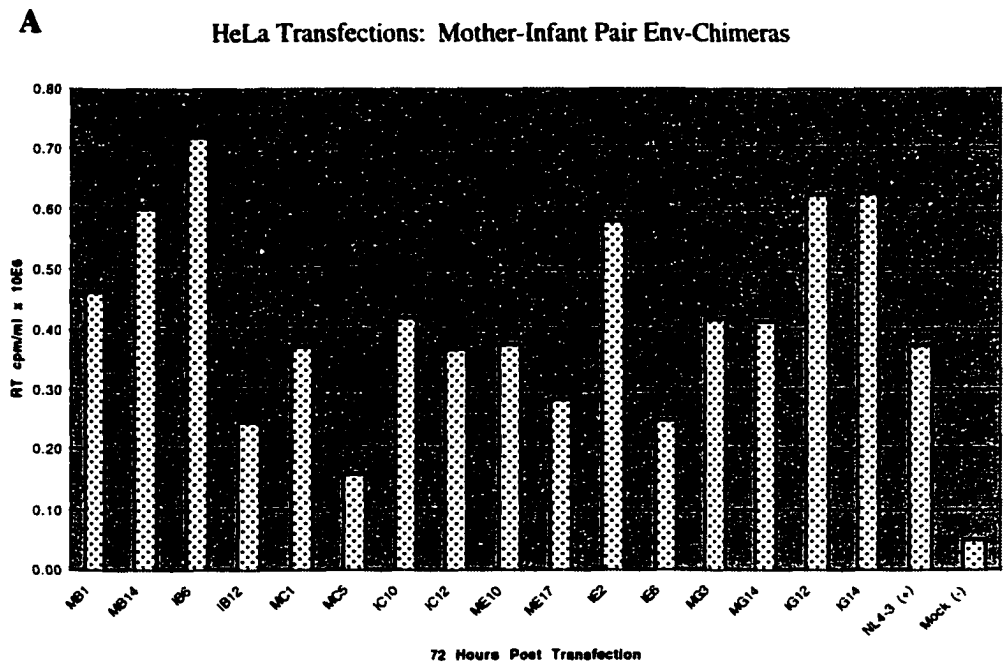
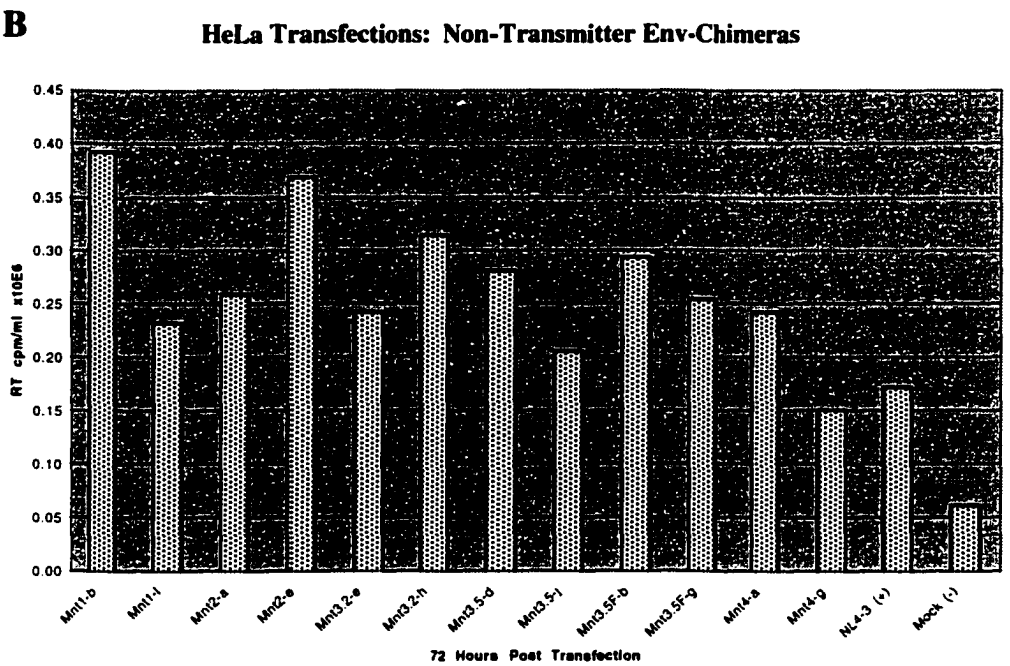

V3
C3
V4
C4
V5

Figure 23: Transfection of proviral gp120 chimeras into HeLa cells





3.4.1 Effect of Mother-Infant Pairs' and Non-transmitting Mothers' gp120 on HIV-1 Replication in T-lymphocyte Cell Lines

We first determined whether the mother-infant and non-transmitting mothers' gp120 chimeras retained the lymphotropic properties of the parental clone NL4-3 by infecting the T-cell line A3.01 with our chimeric viruses, NL4-3, and the macrophage-tropic isolate Ada-M.

Our results show that the lymphotropic clone NL4-3 productively infected and replicated in the T-cell line achieving high levels of replication over the 21 day (Fig. 24) and 24 day periods (Fig. 25 all infections were performed in triplicates, error bars indicate standard deviation). In contrast, the macrophage-tropic isolate Ada-M did not replicate in the A3.01 cell line. Similarly, the mother-infant gp120 chimeras (Fig. 24) as well as the non-transmitting mother gp120 chimeras (Fig. 25) did not replicate in the A3.01 cell line, suggesting that our chimeras had lost the lymphotropism of the parent clone NL4-3 due to the reciprocal insertion of the patients' gp120.

Figure 24: Infection of A3.01 T-Cell line with Mother-Infant gp120 Chimeras

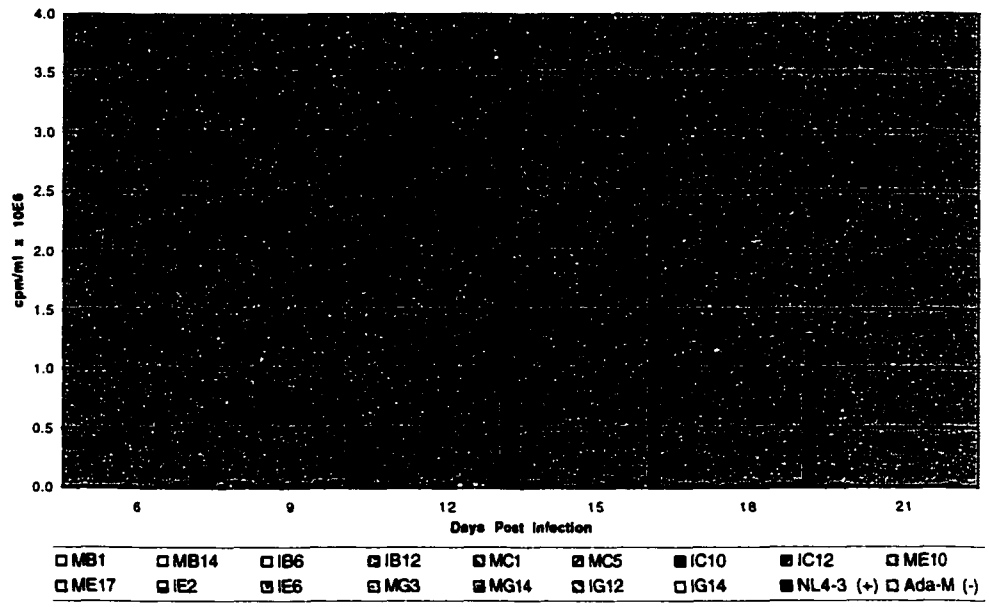
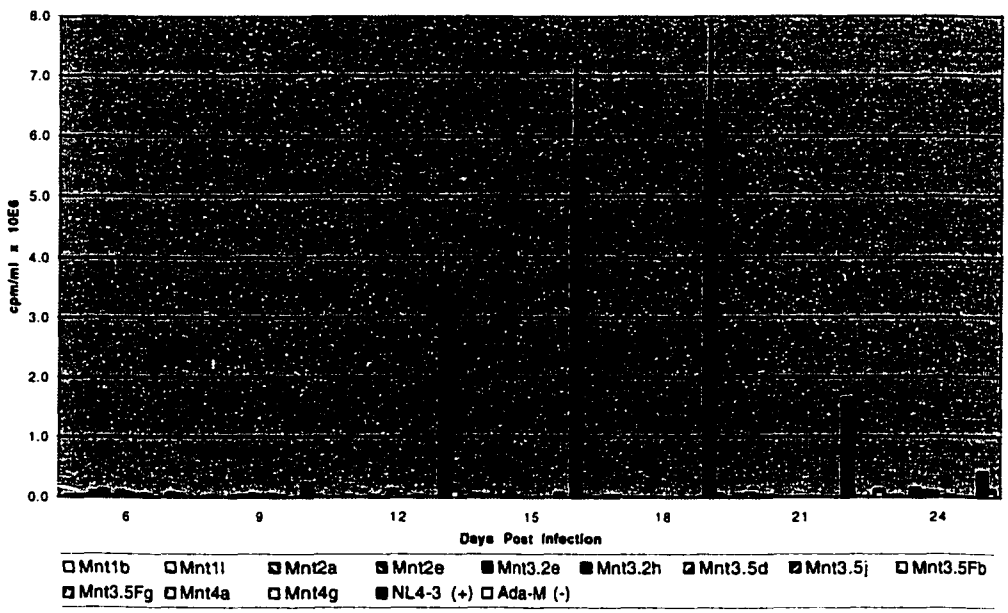


Figure 25: Infection of A3.01 T-Cell line with Non-transmitting Mother gp120 chimeras



3.4.2 Phenotype of Mother-Infant Pairs' and Non-transmitting Mothers' gp120 isolates

3.4.2.1 Cytopathic Effect of Mother-Infant Pairs' and Non-transmitting Mothers' gp120 isolates

We examined the cytopathic effect (ability to induce syncytia) of our gp120 chimeras by infecting the T-cell line MT-2 that is used to determine syncytia formation. NL4-3 produced high levels of syncytia (more than 10 per field of view) and is designated as syncytium-inducing (SI) phenotype, whereas the macrophage-tropic virus Ada-M produced no syncytia in culture, therefore referred to as non-syncytium-inducing (NSI) phenotype. As shown in Table 8 and Table 9, all of our gp120 chimeras failed to produce syncytia in MT-2 cells, denoting NSI phenotypes. This data suggests that the gp120 from mother-infant pairs and non-transmitting mothers changed the SI parental virus NL4-3 to NSI phenotype.

3.4.2.2 Co-Receptor Utilization of Mother-Infant Pairs' and Non-transmitting Mothers' gp120 isolates

We determined the co-receptor utilization of our chimeras using the U373-MAGI indicator cell lines. These cell lines express the CD4 molecule in conjunction with either the CCR5 or the CXCR4 co-receptor and are infectable with either macrophage-tropic (R5) or lympho-tropic (X4) viruses, respectively. To assess differences in infectivity of the chimeras an equal amount of virus was used and infectivity values for each of the chimeras are expressed as ratios of infected cells/well to infected cells/well of AdaM (for U373-MAGI-CCR5) and NL4-3 (for U373-MAGI-CXCR4). Therefore, AdaM and NL4-3 receive the infectivity score 1 and the chimeras

a multiple or fraction of this score. As shown in Table 8 and Table 9, the gp120 chimeras, including the (R5) viruses Ada-M and BaL, were able to infect the U373-MAGI CCR5$^+$ cell line but were unable to infect the U373-MAGI CXCR4$^+$ cell line. In contrast, the X4 viruses NL4-3 and LAV were able to infect the U373-MAGI CXCR4$^+$ cell line but unable to infect the U373-MAGI CCR5$^+$ cell line. There was a considerable difference in the infectivity of the different chimeras. The infectivity of the transmitting mother's chimeras was similar to the infectivity of the R5 virus AdaM but much lower than the infectivity of the laboratory adapted R5 virus BaL (Table 8). This is in contrast to the infectivity of the non-transmitting mother's chimeras. Only the infectivity of Mnt4g was similar to AdaM and the infectivity of Mnt2a was much below that of AdaM. The infectivity of the remaining chimeras was two (Mnt3.5Fb) to five fold (Mnt3.5d, Mnt3.5Fg, Mnt4a) higher than that of AdaM. We were unable to show the infectivity of the chimeras Mnt1l, Mnt2e, Mnt3.2e, Mnt3.2h, Mnt3.5j, and IG12. These chimeras may represent variants that display very low levels of infectivity below the detection level of this assay or use a co-receptor other than CXCR4 and CCR5. While U373-MAGI cells do not represent primary lymphocytes and macrophages, the results show that the gp120 of the mother-infant pairs and non-transmitting mothers confer the ability to use the CCR5 co-receptor, suggesting R5 tropism and document differences in infectivity.

Table 8: MT2 cell line phenotype and co-receptor utilization of HIV-1 Mother-Infant pair gp120 chimeras

Clone	Syncytia	Phenotype[A]	Infectivity[b] on MAGI-CCR5	Infectivity[b] on MAGI-CXCR4
NL4-3	+	SI	-	1
LAV	N/D	N/D	-	0.2
BaL	-	NSI	2.3	-
AdaM	-	NSI	1	-
MB-1	-	NSI	0.4	-
MB-14	-	NSI	0.7	-
IB-6	-	NSI	0.3	-
IB-12	-	NSI	0.8	-
MC-1	-	NSI	0.6	-
MC-5	-	NSI	0.3	-
IC-10	-	NSI	1.3	-
IC-12	-	NSI	0.7	-
ME-10	-	NSI	0.03	-
IE-2	-	NSI	1.2	-
IE-6	-	NSI	0.9	-
MG3	-	NSI	0.1	-
MG-14	-	NSI	1.8	-
IG12	- ?	- ?	- ?	- ?
IG-14	-	NSI	0.9	-

Table 9: MT2 cell line phenotype and co-receptor utilization of HIV-1 Non-transmitting Mother gp120 chimeras

Clone	Syncytia	Phenotype[a]	Infectivity[b] on MAGI-CCR5	Infectivity[b] on MAGI-CXCR4
NL4-3	+	SI	-	1
LAV	N/D	N/D	-	0.2
BaL	-	NSI	2.3	-
AdaM	-	NSI	1	-
Mnt1-b	-	NSI	1.7	-
Mnt1-l	- ?	- ?	- ?	- ?
Mnt2-a	-	NSI	0.1	-
Mnt2-e	- ?	- ?	- ?	- ?
Mnt3.2-e	- ?	- ?	- ?	- ?
Mnt3.2-h	- ?	- ?	- ?	- ?
Mnt3.5-d	-	NSI	5.2	-
Mnt3.5-j	- ?	- ?	- ?	- ?
Mnt3.5F-b	-	NSI	2.3	-
Mnt3.5F-g	-	NSI	5.5	-
Mnt4-a	-	NSI	5.0	-
Mnt4-g	-	NSI	1.1	-

[a]SI, syncytium inducing; NSI, non-syncytium inducing; phenotypes of HIV-1 chimeric clones were determined in MT-2 cells (*ACTG Virology Manual for HIV Laboratories*).
[b]Equal amounts of virus was used. Infectivity values are expressed as ratios of infected cells/well to infected cells/well of AdaM (for CCR5) and NL4-3 (for CXCR4).
NL4-3 and LAV are X4 and BaL and AdaM are R5 HIV-1 strains;
M stands for mother, I for infant; N/D stands for not determined;

3.4.3 Effect of Mother-Infant Pairs' and Non-transmitting Mothers' gp120 on HIV-1 Replication in Primary Cells

Since macrophage-tropic isolates are known to replicate in primary peripheral blood lymphocytes (PBL) because of the presence of CCR5 co-receptors, we used our chimeras to infect PBLs to determine the replication kinetics of the chimeric viruses.

Figures 26 and 27 (A and B) show the replication kinetics of the transmitting mothers' and the non-transmitting mothers' gp120 chimeras in PBL over a 30 day period (all infections were performed in triplicates, error bars indicate standard deviation). Most of our chimeras replicated in PBL but at a lower level compared with the laboratory adapted macrophage-tropic isolate Ada-M. Moreover, our chimeras replicated at the same or higher levels compared with the lymphotropic strain LAV and the parental strain NL4-3. The chimeric viruses containing the gp120 from mothers B and G achieved high levels of replication with a peak at day 15. However, the gp120 chimeras derived from mothers C and E replicated at a lower level (Fig. 26).

The chimeric viruses containing the gp120 from the non-transmitting mothers Mnt1 and Mnt2 replicated at a low level comparable to transmitting mother E chimeras. The chimeras Mnt4a and g replicated to high levels similar to the mothers B and G chimeras (Fig. 27 A). The gp120 chimeras from Mnt3, which represent three time points in this mother's infection history, replicated at levels similar to the other non-transmitting mothers' chimeras. Chimeras Mnt3.5d and Mnt3.5Fg replicated at high levels comparable to Mnt4. However, chimera Mnt3.5Fb replicated to a much lower level (Fig. 27 B). Whereas the replication of most of the non-transmitting mothers' and

transmitting mothers' chimeras peaked at day 15, the replication of chimera Mnt2a peaked at a low level at day 30. Both the mother-infant pairs' and the non-transmitting mothers' gp120 chimeras replicated in primary peripheral blood lymphocytes (PBL) with mother-infant pair chimeras replicating at a slightly higher level than the non-transmitting mothers' chimeras. However, were unable to document a significant difference in replication kinetics of the transmitting and non-transmitting mothers gp120 chimeras.

Figure 26: Infection of PBL with Transmitting Mother gp120 chimeras

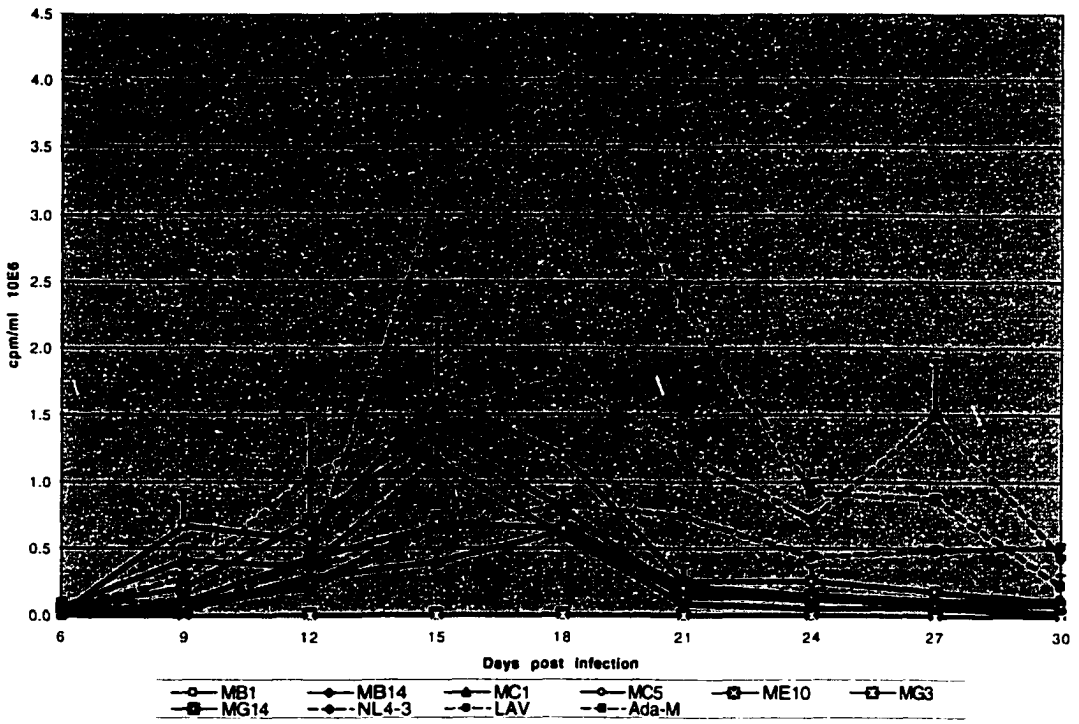
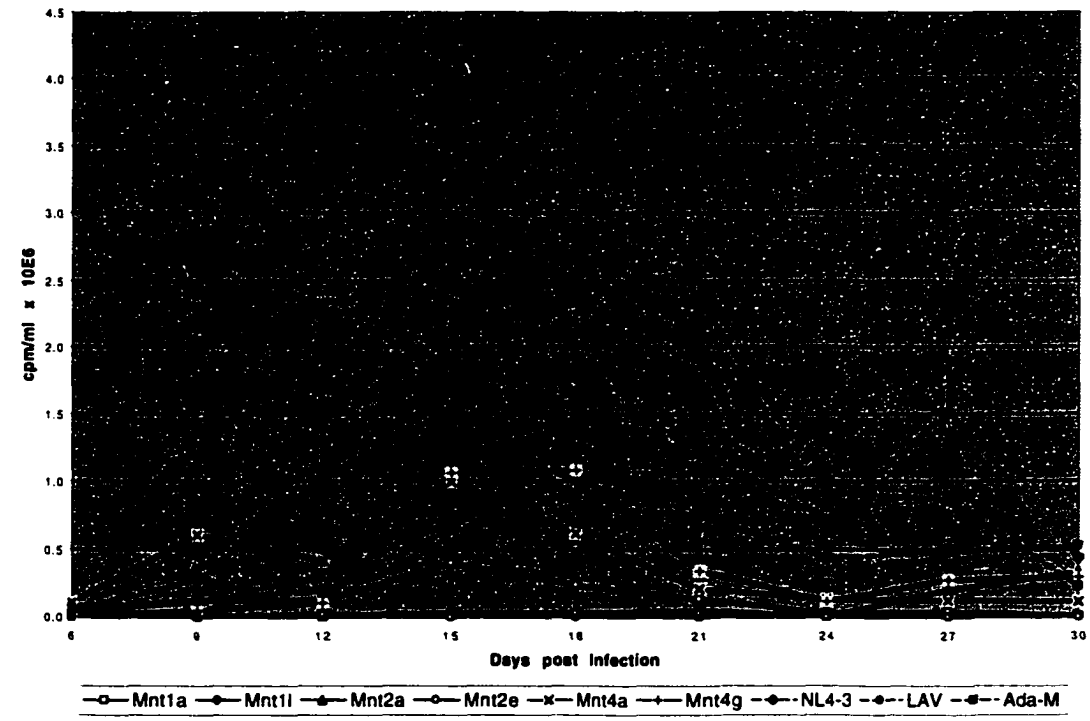


Figure 27: Infection of PBL with Non-transmitting Mother gp120 chimeras

A



B

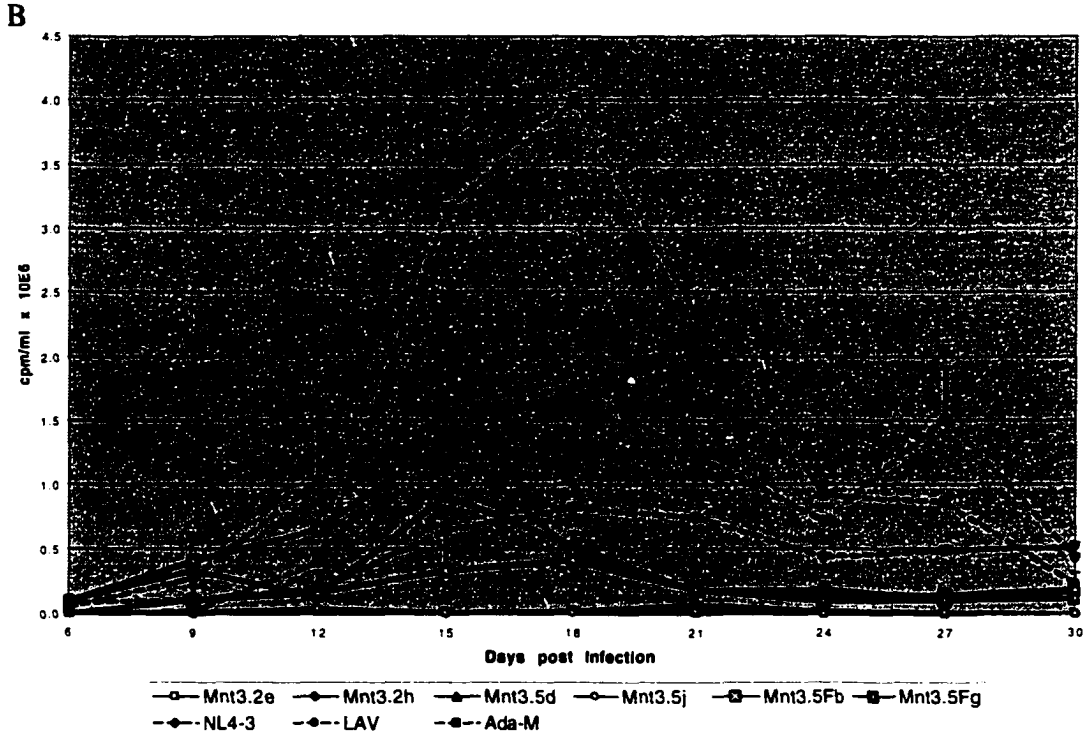
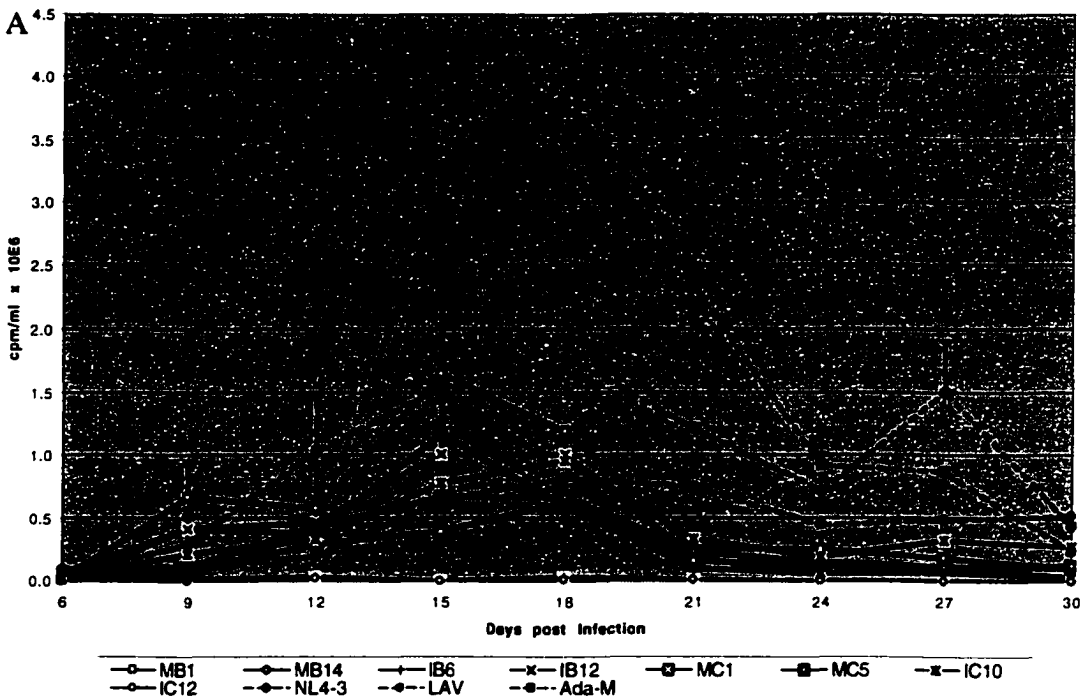
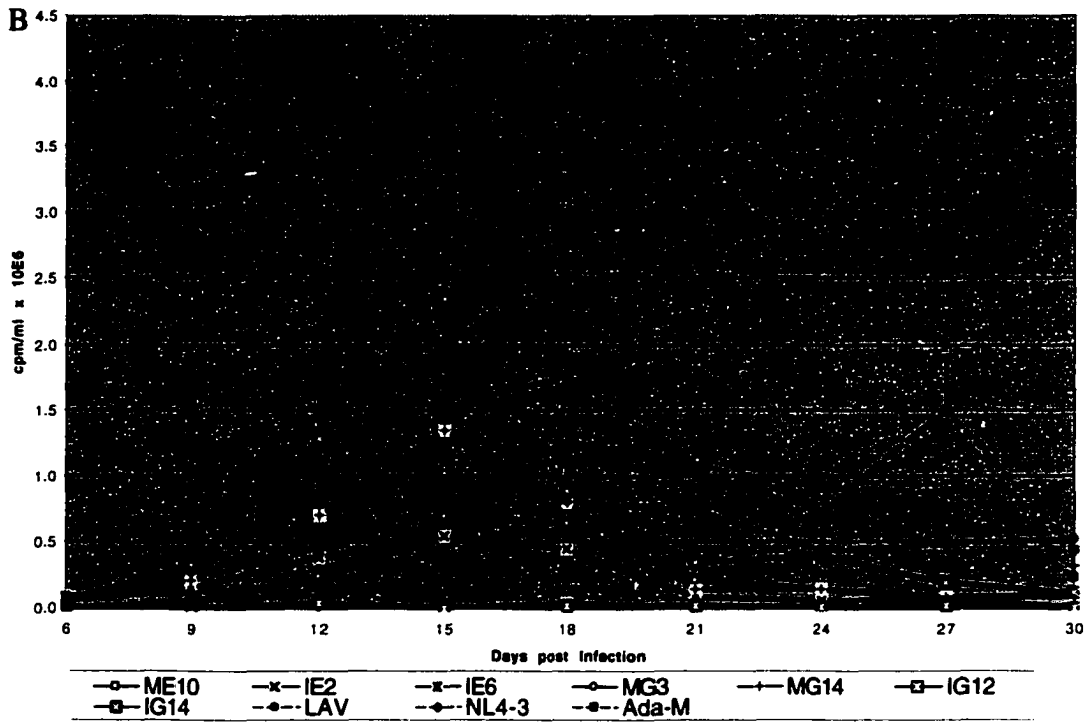


Figure 28 A and B show the replication of the infants' gp120 chimeras in comparison with their respective mothers' chimeras. The gp120 chimeras derived from infants B, C and E replicated to similar levels as their respective mothers. This is in contrast to the chimera derived from infant G (IG14) that achieved much higher levels of replication than its mother chimera (MG14) and the parental strain NL4-3 (Fig. 28 B).

Figure 28: Infection of PBL with Mother-Infant pair gp120 chimeras




B
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
cpm/ml x 10E6
6
9
12
15
18
21
24
27
30
Days post infection
ME10
IE2
IE6
MG3
MG14
IG12
IG14
LAV
NL4-3
Ada-M

4 DISCUSSION

While perinatal transmission may be multifactorial in nature, characterization of HIV-1 transmitted from mother to infant may provide relevant information towards the development of strategies for prevention and treatment of HIV-1 infection in children. Here we provide evidence that several regions of the HIV-1 genome may be involved in mother-infant transmission. The two structural genes *gag* p17 and gp120 as well as the accessory gene *nef* are important factors for viral replication. The results of the characterization of the HIV-1 *gag* p17, *nef* and the *env* gp120 from mother-infant pairs as well as non-transmitting mothers might be helpful in understanding the pathogenesis of HIV-1 infection in mothers and infants including the molecular mechanisms involved in perinatal transmission, which may aid in the development of better strategies for prevention and treatment.

4.1 Role of *gag* p17MA in Mother-Infant Transmission

In this study, we have performed a complete analysis of HIV-1 *gag* p17MA sequences directly derived from the DNA of uncultured PBMC from both mother-infant pairs and non-transmitting mothers (MNT). The study subjects were seven HIV-1 infected mothers and their infected infants, and three HIV-1 infected mothers who failed to transmit to their infants. This included MNT3 who gave birth to a total of five uninfected children. The *gag* p17MA sequence isolates, from the seven mother-infant pairs provide evidence for the maintenance of an intact and functional HIV-1 *gag* p17 matrix open reading frame following mother-to-infant transmission with a 86.2%

frequency (90.7% excluding mother E who harbored 60% defective p17 sequences). The seven transmitting mothers' sequences alone (excluding the infants' sequences) displayed a frequency of p17MA open reading frames of 88.6%. Similarly, the *gag* p17MA sequences isolated form the non-transmitting mothers displayed a frequency of intact open reading frames of 84%. However, among the transmitting mothers' data sets, mother-E harbors almost all of the defective sequences (9 out of the total of 10 for all sequences combined). Therefore, one out of the seven transmitting mothers harbored defective sequences, while the remaining six mothers displayed a 98.9% conservation of p17MA open reading frames. Although we were unable to document a clear correlation between maintenance of p17MA open reading frames and transmission, most of the non-transmitting mothers displayed a lower frequency of intact p17MA open reading frames. The genetic variability, measured as nucleotide and amino acid distances, of *gag* p17 matrix sequences from the non-transmitting mothers were more homogenous compared to the sequences from the mother-infant pairs. For the seven transmitting mothers' data sets the nucleotide distances ranged from 0.2 to 6.5% (2.3% median values) and amino acid distances from 0 to 12.0% (4.5% median values)(Table 3), compared to the five MNT sample sets' nucleotide distances that ranged from 0 to 5.6% (0% median values) and amino acid from 0 to 7.7% (0% median values)(Table 4). Similar results have been described for the *vif* and *vpr* (237) genes as well as *env* V3 region isolates (142) from the same mothers. We also compared the levels of genetic diversity, estimated as nucleotide substitutions per site per generation (θ), of the seven transmitting mothers with the five sample sets from the non-

transmitting mothers. The non-transmitting mothers' sample sets displayed slightly lower estimates of genetic diversity in comparison to the transmitting mothers similar to the observed levels of genetic diversity for the V3 region (142). The functional domains required for Gag p17 matrix, including targeting of the Gag precursor protein to the plasma membrane (20, 75, 243, 246), virus assembly and release (63, 75), envelope glycoprotein incorporation into virus particles (60, 242), virus entry (241) and localization of the virus preintegration complex to the nucleus of nondividing cells (21, 67, 220) were generally conserved in both the mother-infant pairs' sequences as well as the MNTs' sequences. However, other functional domains, including the polymerization site were less conserved in non-transmitting mothers' sequences compared to transmitting mothers' sequences.

For the mother-infant pairs there was a low degree of heterogeneity of *gag* p17 nucleotide and amino acid sequences within mothers, within infants and between epidemiologically linked mother-infant pairs. This is in agreement with earlier published reports for *gag* (97, 156) and other conserved genes, such as *vif* (239), *vpr* (238) and *nef* from HIV-1-infected individuals. The epidemiologically linked mother-infant pairs' sequences were closer than the sequences from epidemiologically unlinked individuals, suggesting that mother-infant pairs' sequences can be identified by comparison of a conserved gene such as *gag*. In addition, there was very little selection pressure for change on *gag* p17 sequences. This finding is in contrast to the *env* V3 region sequences from the same seven mother-infant pairs for which a high selection pressure for change was observed (3). However, the little selection pressure observed

on p17MA sequences is consistent with those observed for *vif* (239), *vpr* (238), *vpu* (236) and *nef* (Table 7) sequences for the same mother-infant pairs. Analysis of the 3-D structure-function relationship of the p17 matrix proteins deduced from the 166 p17 sequences obtained from the seven mother-infant pairs showed that, consistent with a previous report (23), the amino acids required for correct folding or proper MA-MA interaction in order for p17MA to establish HIV-1 infection and replication, were highly conserved. These data suggest that the amino acids essential to confer a native 3-D structure to p17MA were present in most of mother-infant sequences.

The nucleotide and amino acid distances of *gag* p17 sequences were lower in the non-transmitting mothers compared to the mother-infant pairs' sequences. An interesting finding was that the sequences isolated from the from the different sample time points from MNT3 (MNT-3.2, MNT-3.5, and MNT-3.5F) were highly homogenous. The genetic distance data for sequences from MNT-3 (Table 4), in conjunction with the phylogenetic analysis (Fig. 16), demonstrate this high level of homogeneity. We also compared levels of genetic diversity (θ) within the five sample sets of nontransmitting mothers. While MNT-2 shows the greatest estimates of genetic diversity, the three different sample sets of MNT-3 had the lowest estimates of genetic diversity. Analysis of nonsynonymous (d_n) and synonymous (d_s) substitutions showed that the d_n/d_s ratios for the five sample sets of non-transmitting mothers were slightly lower than the d_n/d_s ratios for *gag* p17 sequences form transmitting mothers and their infants and are comparable to d_n/d_s ratios for some of the non-transmitting mothers' *vif* and *vpr* sequences (237). This analysis also shows that there was no selective pressure

for change in either the transmitting mothers' sequences or the non-transmitting mothers' sequences.

A number of functions have been proposed for HIV-1 p17 matrix, including the targeting of Gag to the plasma membrane (20, 75, 243, 246), virus assembly and release (63, 75), envelope glycoprotein incorporation into virus particle (60, 242), virus entry (241) and localization of the virus preintegration complex to the nucleus of nondividing cells (21, 67, 220). Examination of the membrane-binding region, including the N-terminal myristoylation site that is essential for targeting of Gag to the plasma membrane, showed a high degree of conservation of this site in the mother-infant pairs' p17 matrix sequences as well as the three non-transmitting mothers' sequences. The nuclear localization signal (NLS) in p17 matrix (position 25 to 33), including three lysines (K) at positions 26, 27 and 30 and a tyrosine (Y) at 29 that is required for import of the preintegration complex into the nucleus of the nondividing cells, was fairly conserved in the mother-infant pairs' sequences as well as nontransmitting mothers' sequences. In the MNT-2 sequences, the basic lysine (K) at position 30 was substituted by glutamine (Q). The presence of a nuclear localization signal which supports HIV-1 replication in non-dividing cells including macrophages and quiescent T-cells (21, 66, 89, 220) has been controversial (59). The two domains required for virus assembly positions 55-59 and 84-99 (63) were generally conserved in the mother-infant pairs' sequences as well as in most of the nontransmitting mothers' sequences. The valine at position 35 that is essential for incorporation of gp120 into virus particles (63) was found that it was present in both all the mother-infant pairs' sequences and the MNT

sequences. In the polymerization site (position 47-59), sequences of MNT-1 and MNT-2 carried a substitution of serine (S) at position 54 by alanine (A) and MNT-3 a substitution of glycine (G) at 49 by a serine (S). This serine (S) at position 54 was highly conserved in our previously analyzed mother-infant sequences, subtype B sequences and p17 sequences analyzed by others (156).

Important humoral and cellular immune responses to HIV-1 p17 matrix protein have been demonstrated in infected patients (30, 149). In addition, the antibodies against p17 have been shown to provide a correlation between HIV-1 infection and clinical status of HIV disease in infected adults and children, suggesting a protective role for the immunological response against p17 (121, 132, 216). Several motifs in the p17 matrix protein were described by Narwa et al (156) to be associated with transmission. A KIEEEQN motif in the major antibody binding site was shown to be associated with transmission (156). Moreover, a C-terminal 6-mer QVSQNY that is conserved in CON-B and was shown to be significantly associated with non-transmitting status (156). We observed several motifs in the mother infant pairs' p17MA sequences that could be associated with transmission, including glutamic acid (E) or aspartic acid (D) at position 55, tyrosine (Y) or phenylalanine (F) at position 79, an aspartic acid (D) or glutamic acid (E) at positions 93 and 102, and a dipeptide alanine-glutamic acid (AD) at positions 122-123 were present in most of the mother-infant pairs' sequences. Narwa et al., (156) have shown that a glutamic acid (E) at position 93 and KIEEEQN at 103-109 to be significantly associated with transmission. Examination of our sequences revealed that the glutamic acid at 93 was present only in

one of seven pairs and KIEEEQN at 103-109 in six of seven pairs. Therefore, the KIEEEQN motif was generally found to be conserved in our transmitting mothers and less conserved in our non-transmitting mothers' sequences ($P<0.001$, chi square test). Also, the glutamic acid (E) at position 93 was not found to be associated with transmitting or non-transmitting status. To our surprise, a valine at position 104 that was shown to be significantly associated with non-transmitting status (156) was found in 21 of the 22 clones of mother-infant pair C. Nonetheless, the valine (V) at position 104 was found to be associated with non-transmitting status ($P<0.001$). The C-terminal 6-mer QVSQNY that is conserved in consB and shown to be associated with nontransmitting status was conserved in 3 of our 7 transmitting mothers but highly conserved in all our nontransmitting mothers' sequences ($P<0.001$). Several other amino acid motifs, including a lysine (K) or glutamine (Q) at position 15, an alanine (A) at 54, a lysine (K) at 76, a valine (V) at 104 and an aspartic acid (D) at 102 and 121 were conserved in most of the non-transmitting mothers' sequences ($P<0.001$) compared to our previously analyzed transmitting mothers' sequences and consB sequence. Some of the differences seen in our study from Narwa et al., (156) could be due to their direct sequencing of the PCR products and our sequencing of several clones or variants (8-21) from each patient after cloning of the PCR products as well as the time of sampling and clinical stage of the patients. Some motifs in *gag* p17 matrix that differ between transmitting and non-transmitting mothers could be used as new targets for preventive strategies for HIV-1 perinatal transmission.

4.2 Role of *nef* in Mother-Infant Transmission

We analyzed *nef* sequences form seven mother-infant pairs following perinatal transmission, including a set of twins of mother H. We describe the maintenance of intact *nef* open reading frames with high frequency (86.7%) following mother to infant transmission. The phylogenetic analysis demonstrated that each mother-infant pair's sequences were separated from each other and most of the infants were likely infected with one or few variants found in the mothers (Fig. 18). Analysis of the genetic variability, measured as nucleotide and amino acid distances, showed low degree of variability in most mother-infant pairs and are comparable to the sequence distances described for other conserved HIV-1 genes such as *gag* (84, 85), *vif* (239) and *vpr* (238). In addition, the estimation of the mutation rate per generation per site (θ) and the population growth rate (g) (Table 7) showed that the mother-infant pairs' sequences displayed a lower level of estimates of genetic diversity and higher rates of growth, including a slightly higher level of estimates of genetic diversity and growth rates in some infants' sequences compared with their mothers' sequences. We also found that there was little selection pressure on the *nef* sequences from the mothers and the infants for change as determined by the ratio of nonsynonymous (d_n) to synonymous (d_s) substitutions (Table 7). We also analyzed the functional domains and motifs required for Nef activity of the deduced amino acid sequences and describe evidence of conservation of the motifs that are important membrane association, down modulation of the CD4 receptor and MHC-I, and the alteration of the activation state of cells.

The analyzed 196 *nef* sequences from the seven mother-infant pairs displayed an 86.7% frequency of intact open reading frames that is comparable with other conserved HIV-1 genes in mother-infant transmission, including *gag* p17MA (86.2%), *vif* (89.8%) (239), *vpr* (92.1%) (238) and *tat* (90.9%) (101) from the same mother-infant pairs, suggesting that *nef* may play an important role in perinatal transmission. An interesting finding was that the sequences isolated from one of the twins of mother-H (HT1) contained only 50% intact open reading frames, which is consistent with earlier reports of presence of attenuated *nef* genes in long-term nonprogressors (42, 139). However, the sibling twin (HT2) and the mother (MH) had respective frequencies of intact open reading frames of 93% and 94%. Furthermore, we found a considerable difference between HT1 and HT2 in terms of genetic variability in both the nucleotide and amino acid differences. HT1 whose sequences displayed the higher frequency of defective reading frames was also more heterogeneous. Moreover, our phylogenetic analysis demonstrated that the sequence isolated from the twins HT1 and HT2 intermingle but are separated from the mother's sequences in a sub clusters within the pair-H sub tree. This suggests that both infants were most likely infected by the same or a very similar variant. Thus the differences in the maintenance of intact open reading frame as well as genetic variability are likely to be due to HT1 specific host factors. The observed differences in genetic variability may be due to the difference in selection pressure as the analysis of the accumulation of non-synonymous and synonymous substitutions might suggest. HT1 shows a d_n/d_s ratio of 0.69, though still smaller than 1, two fold higher than the d_n/d_s ratio of HT2. In addition, since it was shown that an intact *nef*

gene is not necessary for replication of HIV-1 but required for infection as well as enhanced replication in certain cells such as monocytes and macrophages and more importantly required for pathogenicity, it is interesting to note that one twin harbors intact and functional *nef* sequences (HT2) whereas the sibling twin (HT1) does not. It would have been interesting to genetically characterize the viral variants found in the twins over the time of infection as well as to follow disease progression. Unfortunately, we were unable to obtain follow-up samples from this interesting history of infection.

Our data on a low degree of variability of mother-infant pairs' *nef* sequences are consistent with earlier published analysis of *nef* sequences from HIV-1 infected individuals (98) and other *nef* sequences described in HIV databases (153). The infants' sequences displayed either similar or higher variability compared with their mothers' sequences. While three infant's sequence sets (IB, IC and ID) were less variable compared with their respective mothers, other infant's sequence sets (IE, IF, IG, HT1 and HT2,) displayed higher variability than their mothers. We could not demonstrate a correlation between genetic variability and age of the infected infants as shown for the envelope V3 region sequences of the same mother-infant pairs (3). In addition, the genetic diversity (θ) estimation, as estimated by EVE, revealed that the viral populations in the infants displayed higher genetic diversity as well as higher growth rates than the populations in the mothers. Although we observed a considerable difference in sequence variability between the twins HT1 and HT2, the population growths rates did not differ significantly (-3.0 and -5.25 respectively). The observed levels of genetic diversity in the *nef* gene from the seven transmitting mothers as

estimated by the Watterson model and Coalesce are similar to the estimates based on the V3 region (142) as well as the *gag* p17 gene from the same mothers but higher than estimates for the V3 region (142) and *gag* p17 gene (Table 5) from nontransmitting mothers.

Finally, we analyzed the functional domains and motifs required for Nef activity of the deduced amino acid sequences isolated from the seven mother-infant pairs. Nef is one of the important virulence factors for the primate lentiviruses and, while *nef* is dispensable for HIV-1 replication, it is important in infectivity, pathogenesis (86) and disease progression (42, 112, 126, 139). Specifically, we analyzed the motifs that are important for Nef membrane association, down modulation of the CD4 receptor and MHC-I, and the alteration of the activation state of cells. We found that the residues and motifs that form the membrane anchor arm were functionally highly conserved in all mother-infant pair sequences. Specifically, the myristoyl acceptor site (G^2) was highly conserved in all mother-infant sequences. The patch of basic amino acids (residues 4-22) was mostly conserved in the mother-infant sequences and the few observed substitutions would not grossly change the basic character of this region. The opposite, some substitutions increased the number of positive charges that would increase the affinity of Nef to the cell membrane. The direct CD4 binding site in Nef, that is comprised of the residues W^{57}, L^{58} and E^{59} (79, 81) was highly conserved in five out of the seven mother-infant pairs (Fig. 19). We also found that the dileucine based endocytosis signal in Nef ($E/D^{160}xxxLL^{165}$) (19) that can recruit APs to the cell membrane and therefore enhance internalization of CD4 as well as MHC-I was highly

conserved in all our seven mother-infant pair sequences. The additional AP binding domains that are comprised of the acidic cluster (D^{174}-E^{179}) as well as a DD^{175} motif were mostly conserved. The acidic cluster was variable in most of the mother-infant pair isolates however the observed variations did not grossly change the character of the motif and therefore should not alter its functionality. Thus, our data suggest that membrane localization as well as the ability to downregulate surface CD4 is an important determinant for perinatal transmission. This is especially evident in the high conservation of the glutamic acid (E^{177}) in all seven mother-infant pairs' sequences, since it was shown that a E->G^{177} mutation resulted in a dominant-negative Nef protein that decreases HIV-1 production as well as infectivity (53). The motifs in Nef implicated to be important for MHC-I downregulation of surface expressed MHC-I (78, 137, 187) were well conserved in the sequence isolates from the seven mother-infant pairs. The central point of Nef interactions with host factors that are part of the cellular signaling machinery is the proline rich region P^{69}xxPxxPxxP^{78}. We found that the four prolines were highly conserved and variability was restricted to a substitution of R^{71} with the compatible K in sequences from pairs B, C, infant-D and mother-H and a substitution with a T in isolates from infant-E. Therefore, our data suggest that this central motif that lets Nef directly bind to SH3 domains of cellular kinases is important for mother-infant transmission. Nef binds to the SH3 domains of Hck and Lyn kinases with high affinity, and with lower affinity to Lck, Fyn, and Src. Nef activates Hck, does not affect Lyn or Src and decreases the activity of Lck and Fyn. Therefore, the tissues distribution of the individual kinases may be the critical, as the kinase Hck is restricted

to macrophages, in contrast to T-cells where Nef may interact with the lower affinity targets Fyn, or Lck. Binding of Nef to Hck is required for enhanced growth of nef$^+$ viruses in monocytes (177, 195) and since transmitted viruses (perinatal as well as sexual transmission) were shown to be of the macrophage tropic kind, the conservation of the motif interacting with Hck may be important and may provide further evidence that the macrophage tropic variants are important in mother-infant transmission.

4.3 Role of gp120 in Mother-Infant Transmission

In this study, we compared the biological properties of transmitting mothers' gp120 isolates with gp120 isolates from mothers who failed to transmit the virus by a reciprocal insertion of the patient derived gp120 into the infectious viral clone NL4-3. The results show that the gp120 from the mother-infant pairs and the non-transmitting mothers altered the cellular tropism of the lymphotropic parent clone NL4-3 to that of macrophage-tropism. The gp120 chimeras were unable to replicate in the T-lymphocyte cell line A3.01, but infected and replicated productively in peripheral blood lymphocytes. The gp120 chimeras displayed NSI phenotype, as determined by MT-2 cell line infection. In addition, both the transmitting mothers' and non-transmitting mothers' gp120 chimeras were able to utilize the CCR5 co-receptor but not the CXCR4 co-receptor for infection and were thus determined to be R5 or macrophage-tropic. We also documented a difference in infectivity between the chimeric virus. Some of non-transmitting mothers' chimeras displayed higher infectivity than the transmitting mothers' gp120 chimeras. However, for several non-transmitting mothers' chimeras (Mnt1-l, Mnt2-h, Mnt3.2-e, Mnt3.2h and Mnt3.5-j) we were unable to clearly show

replication competence or infectivity. Since these chimeras did not show defects as determined by nucleotide sequencing and did replicate in HeLa cells, the possibility of defects that may have been introduced during the reciprocal transfer process is unlikely. Thus, these chimeras may represent variants that replicate at a very low level, and/or use co-receptor other than CXCR4 and CCR5, including CCR2b, CCR3, CCR8, GPR15 (BOB), STRL33 (Bonzo), V28, and Apj. The infection of peripheral blood lymphocytes (PBLs) revealed that there was no clear distinction of replication efficiency between the transmitting mothers' gp120 chimeras and the non-transmitting mothers' chimeras. The higher infectivity of the non-transmitting mothers' chimeras on the MAGI-CCR5 cell line did not clearly correlate with higher replication levels in PBLs. A possible explanation is that these variants are able to attach and enter the target cell (thus the higher infectivity on the MAGI-CCR5 cell line) but are unable to establish infection in PBL or replicate efficiently in PBL, due to the complex interactions between the C1, C2, and V1/V2 with the V3 regions (26).

It was reported that macrophage-tropic HIV-1 variants are predominantly transmitted during vertical as well as sexual transmission (178, 183, 217, 247). The transmission of a minor macrophage-tropic variant was even demonstrated in the presence of major X4 variants (229), and clearly demonstrates the selective nature of transmission of HIV-1. To explain this biased nature of transmission in the maternal-fetal setting, several models were proposed. The first model suggests transmission of multiple variants but selective amplification of the macrophage-tropic variants in the recipient. The second model suggests that the properties of the virus determine that

only the macrophage-tropic variants reach the recipient. However, a combination of the models of selective amplification and selective transmission may be most likely. For instance, maternal R5 strains replicating in the genital mucosa may preferentially infect fetal Langerhans cells in mucocutaneous fetal tissue. In addition, these variants may be able to replicate more efficiently in these macrophage lineage cells and therefore be selectively amplified. The fact that we could not demonstrate a significant advantage of replication in PBLs for either the transmitting mothers' or non-transmitting mothers' gp120 chimeras, may argue against the selective amplification model. However, our results only demonstrate that any of the variants may replicate in PBLs equally well. Therefore, selective amplification in this cell type seems unlikely. However, selective amplification in the recipient may very well be the case if the predominant cell type for a transmitted variant to encounter are macrophage lineage cells.

REFERENCES

1. **Abrams, E. J., Weedon, J., and Steketee, R. W.** 1998. Association of HIV viral load early in life with disease progression among HIV-infected infants. J Infect Dis. **178:**101-108.

2. **Ahmad, N.** 2000. Molecular mechanisms of human immunodeficiency virus type 1 mother-infant Transmission. Adv Pharm. **49:**387-415.

3. **Ahmad, N., Baroudy, B. M., Baker, R. C., and Chappey, C.** 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.

4. **Ahmad, N., and Venkatessan, S.** 1988. Nef protein of HIV-1 is a transcriptional repressor of HIV-1 LTR. Science. **241:**1481-85.

5. **Aiken, C., Konner, J., Landau, N. R., Lenburg, M. E., and Trono, D.** 1994. Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. Cell. **76:**853-64.

6. **Akaike, H.** 1974. A new look at the statistical model indentification. IEEE Trans Autom Contr. **19:**716-723.

7. **Akari, H., Arold, S., Fukumori, T., Okazaki, T., Strebel, K., and Adachi, A.** 2000. Nef-induced major histocompatibility complex class I down-regulation is functionally dissociated from its virion incorporation, enhancement of viral infectivity, and CD4 down-regulation. J Virol. **74:**2907-12.

8. **Aldrovandi, G. M., Gao, L., Bristol, G., and Zack, J. A.** 1998. Regions of human immunodeficiency virus type 1 nef required for function in vivo. J Virol. **72:**7032-9.

9. **Ariyoshi, K., Weber, J., and Walters, S.** 1992. Contribution of maternal viral load to HIV-1 transmission. Lancet. **340:**435.

10. **Arold, S., Franken, P., Strub, M. P., Hoh, F., Benichou, S., Benarous, R., and Dumas, C.** 1997. The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling. Structure. **5:**1361-72.

11. **Baldi, P., Brunak, S., Frasconi, P., Soda, G., and Pollastri, G.** 1999. Exploiting the past and the future in protein secondary structure prediction. Bioinformatics. **15:**937-46.

12. **Barnham, K. J., Monks, S. A., Hinds, M. G., Azad, A. A., and Norton, R. S.** 1997. Solution structure of a polypeptide from the N terminus of the HIV protein Nef. Biochemistry. **36:**5970-80.

13. **Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science. **220:**868-71.

14. **Baur, A. S., Sass, G., Laffert, B., Willbold, D., Cheng-Mayer, C., and Peterlin, B. M.** 1997. The N-terminus of Nef from HIV-1/SIV associates with a protein complex containing Lck and a serine kinase. Immunity. **6:**283-91.

15. **Bawdon, R. E., Gravell, M., Roberts, S., Hamilton, R., Dax, J., and Sever, J.** 1995. Ex vivo human placental transfer of human immunodeficiency virus-1 p24 antigen. Am J Obstet Gynecol. **172:**530-2.

16. **Benichou, S., Bomsel, M., Bodeus, M., Durand, H., Doute, M., Letourneur, F., Camonis, J., and Benarous, R.** 1994. Physical interaction of the HIV-1 Nef protein with beta-COP, a component of non-clathrin-coated vesicles essential for membrane traffic. J Biol Chem. **269:**30073-6.

17. **Blanche, S., Mayaux, M.-J., Rouzious, C., Teglas, J. P., Firtion, G., Monopoux, F., Ciraru-Vigeneran, N., Meier, F., Tricoire, J., Courpotin, C., Vilmer, E., Griscelli, C., Delfraissy, J.-F., and French Pedriatic HIV Infection Study Group.** 1994. Relation of course of HIV-1 infection in children to the severity of the disease in their mothers at delivery. N Engl J Med. **330:**308-312.

18. **Bodeus, M., Marie-Cardine, A., Bougeret, C., Ramos-Morales, F., and Benarous, R.** 1995. In vitro binding and phosphorylation of human immunodeficiency virus type 1 Nef protein by serine/threonine protein kinase. J Gen Virol. **76:**1337-44.

19. **Bresnahan, P. A., Yonemoto, W., Ferrell, S., Williams-Herman, D., Geleziunas, R., and Greene, W. C.** 1998. A dileucine motif in HIV-1 Nef acts

as an internalization signal for CD4 downregulation and binds the AP-1 clathrin adaptor. Curr Biol. **8:**1235-8.

20. **Bryant, M., and Ratner, L.** 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus type 1. Proceedings of the National Academy of Science, USA. **87:**523-527.

21. **Bukrinsky, M. I., Haggerty, S., Dempsey, M. P., Sharova, N., Adzhubei, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M., and Stevenson, M.** 1993. A nuclear localization signal within HIV-1 matrix protein that governs infection of non-deviding cells. Nature. **365:**666-669.

22. **Buttke, T. M., and Folks, T. M.** 1992. Complete replacement of membrane cholesterol with 4,4',14-trimethyl sterols in a human T cell line defective in lanosterol demethylation. J Biol Chem. **267:**8819-26.

23. **Cannon, P. M., Matthews, S., Clark, N., Byles, E. D., Iourin, O., Hockley, D. I., Kingsman, S. M., and Kingsman, A. J.** 1997. Structure-function studies of the human immunodeficiency virus type 1 matrix protein p17. J Virol:3474-3483.

24. **Cao, Y., Krogstad, P., Korber, B., Koup, R., Muldoon, M., Macken, C., Song, J.-L., Jin, Z., Zhao, J.-Q., Chapp, S., Chen, I., Ho, D., Ammann, A., and The Ariel Project Investigators.** 1997. Maternal HIV-1 viral load and vertical transmission of infection: The Ariel Project for the prevention of HIV transmission from mother to infant. Nat Med. **3:**549-552.

25. **Carrillo, A., and Ratner, L.** 1996. Cooperative effects of human immunodeficiency virus type 1 envelope variable loops V1 and V3 in mediating infectivity for T-cells. J Virol. **70:**1310-1316.

26. **Carrillo, A., and Ratner, L.** 1996. Human immunodeficiency virus type 1 tropism for T-lymphoid cell lines: role of the V3 loop and C4 envelope deteminants. J Virol. **70:**1301-1309.

27. **Casella, C. R., Raffini, L. J., and Panganiban, A. T.** 1997. Pleiotropic mutations in the HIV-1 matrix protein that affect diverse steps in replication. Virology. **228:**294-306.

28. **Centers for Disease Control and Prevention.** 1987. Classification system for HIV in children under 13 years of age. Morbid Mortal Weekly Rep. **36:**225-236.

29. **Centers for Disease Control and Prevention.** 2000. HIV/AIDS Surveillance Report 2000 Vol. 12, No. 1. Centers for Disease Control and Prevention.

30. **Charguelegue, D., Colvin, B. T., and O'Toole, C. M.** 1993. A 7 year analysis of anti Gag (p17 and p24) antibodies in HIV-1 seropositive patients with hemophilia: immunoglobulin G titer and avidity are early predictors of clinical course. AIDS. **7:**S87-90.

31. **Chazal, N., Gay, B., Carriere, C., Tournir, J., and Boulanger, P.** 1995. Human immunodeficiency virus type I MA deletion mutants expressed in baculovirus-infected cells: cis and trans effects on Gag precursor assembly pathway. J Virol. **69:**365-375.

32. **Cheng-Mayer, C., Liu, R., Landau, N. R., and Stamatatos, L.** 1997. Macrophage tropism of human immunodeficiency virus type 1 and utilization of the CC-CKR5 coreceptor. J Virol. **71:**1657-61.

33. **Chesebro, B., Wehrly, K., Nishio, J., and Perryman, S.** 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.

34. **Chouquet, C., Burgard, M., Richardson, S., Rouzioux, C., and Costagliola, D.** 1997. Timing of mother-to-child HIV-1 transmission and diagnosis of infection based on polymerase chain reaction in the neonatal period by a non-parametric method. AIDS. **11:**1183-4.

35. **Coates, K., Cooke, S. J., Mann, D. A., and Harris, M. P.** 1997. Protein kinase C-mediated phosphorylation of HIV-I nef in human cell lines. J Biol Chem. **272:**12289-94.

36. **Coffin, J. M.** 1995. HIV population dynamics in vivo: Implications for genetic variation, pathogenesis and therapy. Science. **267:**483-489.

37. **Coffin, J. M.** 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. J Gen Virol. **42:**1-26.

38. **Cohen, G. B., Gandhi, R. T., Davis, D. M., Mandelboim, O., Chen, B. K., Strominger, J. L., and Baltimore, D.** 1999. The selective downregulation of

class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. Immunity. **10:**661-71.

39. **Colgrove, R. C., Pitt, J., Chung, P. H., Wells, S. L., and Japour, A. J.** 1998. Selective vertical transmission of HIV-1 antiretroviral resistance mutations. AIDS. **12:**2281-2288.

40. **Coll, O., Hernandez, M., Boucher, C. A., Fortuny, C., de Tjada, B. M., Canet, Y., Caragol, I., Tijnagel, J. M., Bertran, J. M., and Ezpanol, T.** 1997. Vertical HIV-1 transmission correlates with a high maternal viral load at delivery. J Acquir Immune Defic Syndr Hum Retrovirol. **14:**26-30.

41. **Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D., and Baltimore, D.** 1998. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature. **391:**397-401.

42. **Deacon, N. J., Tsykin, A., Solomon, A., Smith, K., Ludford, M., Hooker, M., McPhee, D. A., Greenway, A. L., Ellet, A., and Chatfiled, C.** 1995. Genomic structure of an attenuated quasi species of HIV-1 from blood transfusion donor recipients. Science. **270:**988-991.

43. **Delwart, E., Pan, H., Sheppard, H., Wolpert, D., Neumann, A., Korber, B., and Mullins, J.** 1997. Slower evolution of human immunodeficiency virus type 1 quasispecies during progression to AIDS. J Virol. **71:**7498-7508.

44. **Delwart, E. L., Shepard, H. W., Walker, B. D., Goudsmit, J., and Mullins, J. I.** 1994. Human immunodeficincy virus type 1 evoloution in vivo tracked by DNA heteroduplex mobility assay. J Virol. **68:**6672-6683.

45. **Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R., Hill, C. M., Davis, C. B., Peiper, S. C., Schall, T. J., Littman, D. R., and Landau, N. R.** 1996. Identification of a major co-receptor for primary isolates of HIV-1. Nature. **381:**661-666.

46. **Dickover, R. E., Garratty, E. M., Plaeger, S., and Bryson , Y. J.** 2001. Perinatal transmission of major, minor, and multiple maternal human immunodeficiency virus type 1 variants in utero and intrapartum. J Virol. **75:**2194-2203.

47. **Domingo, E.** 1992. Genetic variation and quasispecies. Curr Opin Genet Dev. **2:**61-63.

48. **Dorfman, T., Mammano, F., Haseltine, W. A., and Gottlinger, H. G.** 1994. Role of the matrix protein in virion association of the human immunodeficiency virus type 1 envelope glycoprotein. J Virol. **68:**1689-1696.

49. **Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayanan, C., Maddon, P. J., Koup, R. A., Moore, J. P., and Paxton, W. A.** 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature. **381:**667-673.

50. **Edelstein, R. E., Manns-Arcurino, L. A., and Hughes, J. P.** 1997. Risk of mother-to-infant transmission of HIV-1 is not reduced in CCR5/32ccr5 heterozygotes. J AIDS. **16:**243-246.

51. **Eigen, M.** 1993. Viral Quasispecies. Sci Am. **269:**42-49.

52. **European Collaborative Study Group.** 1988. Mother to child transmission of HIV-1. Lancet. **2:**1039-1042.

53. **Fackler, O. T., d'Aloja, P., Baur, A. S., Federico, M., and Peterlin, B. M.** 2001. Nef from human immunodeficiency virus type 1(F12) inhibits viral production and infectivity. J Virol. **75:**6601-8.

54. **Fackler, O. T., Luo, W., Geyer, M., Alberts, A. S., and Peterlin, B. M.** 1999. Activation of Vav by Nef induces cytoskeletal rearrangements and downstream effector functions. Mol Cell. **3:**729-39.

55. **Felsenstein, J.** 1992. Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phlogenetic estimates. Genet Res. **59:**139-147.

56. **Felsenstein, J.** 1989. Phylip phylogenetic inference package (version 3.2). Cladiatice. **5:**164-166.

57. **Foti, M., Mangasarian, A., Piguet, V., Lew, D. P., Krause, K. H., Trono, D., and Carpentier, J. L.** 1997. Nef-mediated clathrin-coated pit formation. J Cell Biol. **139:**37-47.

58. **Fouchier, R. A., Groenink, M., Kootstra, N. A., Tersmette, M., Huisman, H. G., Miedema, F., and Schuitemaker, H.** 1992. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. J Virol. **66:**3183-7.

59. **Freed, E. O., Englund, G., and martin, M. A.** 1995. Role of basic domain of human immunodeficiency virus type 1 matrix in macrophage infection. J Virol. **69:**3949-3954.

60. **Freed, E. O., and Martin, M. A.** 1996. Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions. J Virol. **70:**341-51.

61. **Freed, E. O., and Martin, M. A.** 1994. Evidence for a functional interaction between the V1/V2 and C4 domains of human immunodeficiency virus type 1 envelope glycoprotein gp120. J Virol. **68:**2503-12.

62. **Freed, E. O., and Martin, M. A.** 1995. Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails blocked by specific, single amino acid substituitions in the human immunodeficiency virus type 1 matrix. J Virol. **69:**19841989.

63. **Freed, E. O., Orenstein, J. M., Buckler-White, A. J., and Martin, M. A.** 1994. Single amino acid changes in the immunodeficiency virus type 1 matrix blocks virus particle production. J Virol. **68:**5311-5320.

64. **Fu, Y. X.** 1994. Estimating effective population size or mutation rate using the frequencies of mutations of various classes in a sample of DNA sequences. Genetics. **138:**1375-1386.

65. **Furtado, M. R., Kingsley, L. A., and Wolinsky, S. M.** 1995. Changes in the viral mRNA expression pattern correlate with a rapid rate of CD4+ T-cell number decline in human immunodeficiency virus type 1 infected individuals. J Virol. **69:**2092-2100.

66. **Gallay, P., Swingler, S., Aitken, C., and Trono, D.** 1995. HIV-1 infection of non-deviding cells: C-terminal phosphorylation of the viral matrix protein is a key regulator. Cell. **80:**379-388.

67. **Gallay, P., Swingler, S., Song, J., Bushman, F., and Trono, D.** 1995. HIV-1 nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. Cell. **83:**569-576.

68. **Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., Safai, B., and et al.**

1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science. **224:**500-3.

69. **Ganeshan, S., Dickover, R. E., Korber, B. T., Bryson, Y. J., and Wolinsky, S. M.** 1997. Human immunodeficiency virus type 1 genetic evolution in children with different rates of development of disease. J Virol. **71:**663-677.

70. **Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M., Shaw, G., Sharp, P. M., and Hahn, B. H.** 1999. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglogytes*. Nature. **387:**436-441.

71. **Gartner, S., Markovits, P., Markovitz, D. M., Kaplan, M. H., Gallo, R. C., and Popovic, M.** 1986. The role of mononuclear phagocytes in HTLV-III/LAV infection. Science. **233:**215-9.

72. **Geyer, M., Munte, C. E., Schorr, J., Kellner, R., and Kalbitzer, H. R.** 1999. Structure of the anchor-domain of myristoylated and non-myristoylated HIV-1 Nef protein. J Mol Biol. **289:**123-38.

73. **Ghys, P. D., Fransen, K., Diallo, M. O., Ettiegne-Traore, V., Coulibaly, I. M., Yeboue, K. M., Kalish, M. L., Maurice, C., Whitaker, J. P., Greenberg, A. E., and Laga, M.** 1997. The associations between cervicovaginal HIV shedding, sexually transmitted diseases and immunosuppression in female sex workers in Abidjan, Cote d'Ivoire. AIDS. **11:**F85-93.

74. **Gichangi, P. B., Nyongo, A. O., and Temmerman, M.** 1993. Pregnancy outcome and placental weights: their relationship to HIV-1 infection. East Afr Med J. **70:**85-9.

75. **Gottlinger, H. G., Sodroski, S. G., and Haseltine, W. A.** 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proceedings of the National Academy of Science, USA. **86.**

76. **Greenberg, M., DeTulleo, L., Rapoport, I., Skowronski, J., and Kirchhausen, T.** 1998. A dileucine motif in HIV-1 Nef is essential for sorting into clathrin- coated pits and for downregulation of CD4. Curr Biol. **8:**1239-42.

77. **Greenberg, M. E., Bronson, S., Lock, M., Neumann, M., Pavlakis, G. N., and Skowronski, J.** 1997. Co-localization of HIV-1 Nef with the AP-2 adaptor

protein complex correlates with Nef-induced CD4 down-regulation. Embo J. **16**:6964-76.

78. **Greenberg, M. E., Iafrate, A. J., and Skowronski, J.** 1998. The SH3 domain-binding surface and an acidic motif in HIV-1 Nef regulate trafficking of class I MHC complexes. Embo J. **17**:2777-89.

79. **Greenway, A., Azad, A., and McPhee, D.** 1995. Human immunodeficiency virus type 1 Nef protein inhibits activation pathways in peripheral blood mononuclear cells and T-cell lines. J Virol. **69**:1842-50.

80. **Griffiths, R. C., and Tavare, S.** 1994. Sampling theory for neutral alleles in a varying environment. Phil Trans Royal Soc London, Series B. **334**:403-410.

81. **Grzesiek, S., Stahl, S. J., Wingfield, P. T., and Bax, A.** 1996. The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. Biochemistry. **35**:10256-61.

82. **Guatelli, J. C.** 1997. The positive influence of Nef on viral infectivity. Res Virol. **148**:34-7.

83. **Guy, B., Kieny, M., Riviere, Y., Peuch, C., Dott, K., Girard, M., Montagnier, L., and Lecocq, J.** 1987. HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling an oncogene product. Nature. **330**:266-269.

84. **Hahn, T., and Ahmad, N.** 2001. Genetic characterization of HIV type 1 gag p17 matrix genes isolates from infected mothers lacking perinatal transmission. AIDS Res Hum Retroviruses. **17**:1673-1680.

85. **Hahn, T., Matala, E., Chappey, C., and Ahmad, N.** 1999. Characterization of mother-infant HIV type 1 gag p17 sequences associated with perinatal transmission. AIDS Res Hum Retroviruses. **15**:875-88.

86. **Hanna, Z., Kay, D. G., Rebai, N., Guimond, A., Jothy, S., and Jolicoeur, P.** 1998. Nef harbors a major determinant of pathogenicity for an AIDS-like disease induced by HIV-1 in transgenic mice. Cell. **95**:163-75.

87. **Harada, S., Koyanagi, Y., and Yamamoto, N.** 1985. Infection of HTLV-III/LAV in HTLV-I-carrying cells MT-2 and MT-4 and application in a plaque assay. Science. **229**:563-6.

88. **Hart, C. A., Palmore, M., Wight, T., and al, e.** 1997. Correlation of Cell-free and cell-associated HIV RNA levels in plasma and vaginal secretions.

89. **Heinzinger, N. K., Bukrinsky, M. I., Haggerty, S. A., Ragland, A. M., Kewalramani, V., Lee, M.-A., Gendelman, H. E., Ratner, L., Stevenson, M., and Emerman, M.** 1994. The Vpr protein of human immunodeficiency virus type 1influences nuclear localization of viral nucleic acids in nondividing cells. Proceedings of the National Academy of Science, USA. **91:**7311-7315.

90. **Henin, Y., Mandelbrot, L., Henrion, R., Pradinoud, R., Coulad, J., and Montagnier, L.** 1993. Virus excretion in the cervicovaginal secretions of pregnant and non-pregnant HIV-infected women. J AIDS. **6:**72-75.

91. **Henrard, D. R., Phillips, J. F., Muenz, L. R., Blattner, W. A., Wiesner, D., Eyster, M. E., and Goedert, J. J.** 1995. Natural history of HIV-1 cell-free virema. JAMA. **274:**554-558.

92. **Hill, C. P., Worthylake, D., Bancroft, D. P., Christensen, A. M., and Sundquist, W. I.** 1996. Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. Proc Natl Acad Sci U S A. **93:**3099-104.

93. **Hodge, D. R., Dunn, K. J., Pei, G. K., Chakrabarty, M. K., Heidecker, G., Lautenberger, J. A., and Samuel, K. P.** 1998. Binding of c-Raf1 kinase to a conserved acidic sequence within the carboxyl-terminal region of the HIV-1 Nef protein. J Biol Chem. **273:**15727-33.

94. **Holland, J. J., de la Torre, J. C., and Steinhauer, D. A.** 1992. RNA virus populations as quasispecies. Curr Top Microbiol Immunol. **176:**1-20.

95. **Huang, M., Orenstein, J. M., Martin, M. A., and Freed, E. O.** 1995. p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. J Virol. **69:**6810-8.

96. **Huang, Y., Paxton, W. A., and Wolinsky, S. M.** 1996. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. Nat Med. **2:**1240-1243.

97. **Huang, Y., Zhang, L., and Ho, D.** 1998. Characterization of gag and pol sequenes from long-term survivors of human immunodeficiecny virus type 1 infection. Virology. **240:**36-49.

98. **Huang, Y., Zhang, L., and Ho, D. D.** 1995. Characterization of nef sequences in long-term survivors of human immunodeficiency virus type 1 infection. J Virol. **69:**93-100.

99. **Huelsenbeck, J. P., and Crandall, K. A.** 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu Rev Ecol Sys. **28:**437-466.

100. **Hughes, E. S., Bell, J. E., and Simmonds, P.** 1997. Investigation of the dynamics of the spread of human immunodeficiency virus to brain and other tissues by evolutionary analysis from p17 *gag* and *env* genes. J Virol. **71:**1272-1280.

101. **Husain, M., Hahn, T., Yedavalli, V. R., and Ahmad, N.** 2001. Characterization of HIV type 1 tat sequences associated with perinatal transmission. AIDS Res Hum Retroviruses. **17:**765-73.

102. **Hutto, C., Parks, W. P., and Lai, S. H.** 1991. A Hospital-based prospective study of perinatal infection with human immunodeficiency virus type 1. J Pediatr. **118:**347-353.

103. **Hwang, S. S., Boyle, T. J., Lyerly, H. K., and Cullen, B. R.** 1991. Identification of the envelope V3 loop as the primary determinant of cell tropism. Science. **253:**71-74.

104. **Iafrate, A. J., Bronson, S., and Skowronski, J.** 1997. Separable functions of Nef disrupt two aspects of T cell receptor machinery: CD4 expression and CD3 signaling. Embo J. **16:**673-84.

105. **Italian Multiculture Study.** 1988. Epidemiology, clinical features, and prognostic factors of pediatric HIV infection. Lancet. **ii:**1043-46.

106. **Jacque, J. M., Mann, A., Enslen, H., Sharova, N., Brichacek, B., Davis, R. J., and Stevenson, M.** 1998. Modulation of HIV-1 infectivity by MAPK, a virion-associated kinase. Embo J. **17:**2607-18.

107. **Javaherian, K., Langois, A. J., McDanal, D., Ross, K. L., Eckler, L. I., Jellis, C. L., Profy, A. T., Rusche, J. R., Bolognesi, D. P., Putey, S. D., and Mathews, T. J.** 1986. Principal neutralizing domain of the human immunodeficiency virus type 1 envelope protein. Proceedings of the National Academy of Science, USA. **86:**6768-6772.

108. **Kalish, L. A., Pitt, J., Lew, J., Landesman, S., Diaz, C., Hershow, R., Hollinger, F. B., Pagano, M., Smeriglio, V., and Moye, J.** 1997. Defining the time of fetal or perinatal acquisition of human immunodeficiency virus type 1 infection on the basis of age at first positive culture. Women and Infants Transmission Study (WITS). J Infect Dis. **175:**712-5.

109. **Karlsson, A. C., Lindback, S., Gaines, H., and Sonnerborg, A.** 1998. Characterization of the viral population during primary HIV-1 infection. AIDS. **12:**839-47.

110. **Kaslow, R. A., Ostrow, D. G., Detels, R., Phair, J. P., Polk, B. F., and Rinaldo, C. R.** 1987. The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. Am J Epidemiol. **26:**310-318.

111. **Katzenstein, T. L., Pedersen, C., Nielsen, C., Lundgren, J. D., Lakobson, P. H., and Gerstoft, J.** 1996. Longitudinal serum HIV RNA quantification: Correlation to viral phenotype at seroconversion and clinal outcome. AIDS. **10:**167-173.

112. **Kestler, H. W., 3rd, Ringler, D. J., Mori, K., Panicali, D. L., Sehgal, P. K., Daniel, M. D., and Desrosiers, R. C.** 1991. Importance of the nef gene for maintenance of high virus loads and for development of AIDS. Cell. **65:**651-62.

113. **Kim, Y. H., Chang, S. H., Kwon, J. H., and Rhee, S. S.** 1999. HIV-1 Nef plays an essential role in two independent processes in CD4 down-regulation: dissociation of the CD4-p56(lck) complex and targeting of CD4 to lysosomes. Virology. **257:**208-19.

114. **Klotman, M. E., Kim, S., Buchbinder, A., DeRossi, A., Baltimore, D., and Wong-Staal, F.** 1991. Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. Proc Natl Acad Sci USA. **88:**5011-5015.

115. **Koito, A., Harrowe, G., Levy, J., and Cheng-Mayer, C.** 1994. Functional role of the V1/V2 region of HIV-1 envelope glycoprotein gp120 in infection of primary macrophages and soluble CD4 neutralization. J Virol. **68:**200-219.

116. **Korber, B. T., Learn, G., Mulins, J. I., Hahn, B. H., and Wolinsky, S. M.** 1995. Protecting HIV databases. Nature. **378:**242-243.

117. **Korber, B. T., MacInnes, K., Smith, R. F., and Myers, G.** 1994. Mutational trends in V3 loop protein sequences observed in different genetic lineages of human immunodeficiency virus type 1. J Virol. **68:**6730-44.

118. **Kuhn, L., Steketee, R. W., Weedon, J., Abrams, E. J., Lambert, G., Bamji, M., Schoenbaum, E., Farley, J., Nesheim, S. R., Palumbo, P., Simonds, R. J., and Thea, D. M.** 1999. Distinct risk factors for intrauterine and intrapartum human immunodeficiency virus transmission and consequences for disease progression in infected children. Perinatal AIDS Collaborative Transmission Study. J Infect Dis. **179:**52-8.

119. **Kuhner, M. K., Yamato, J., and Felsenstein, J.** 1995. Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. Genetics. **140:**1421-1430.

120. **Kuhner, M. K., Yamato, J., and Felsenstein, J.** 1998. Maximum likelihood estimation of population growth rates based on the coalescent. Genetics. **149:**429-434.

121. **Lallemant, M., Baillou, A., Lallement-Le Coeur, S., Nzingoula, S., Mampaka, M., M'Pele, P., Brain, F., and Essex, M.** 1994. Maternal antibody response at delivery and perinatal transmission of human immunodeficiency virus type 1 in african women. Lancet. **343:**1001-1005.

122. **Lama, J., Mangasarian, A., and Trono, D.** 1999. Cell-surface expression of CD4 reduces HIV-1 infectivity by blocking Env incorporation in a Nef- and Vpu-inhibitable manner. Curr Biol. **9:**622-31.

123. **Lamers, S. L., Sleasman, J. W., She, J. X., Barrie, K. A., Pomeroy, S. M., Barrett, D. J., and Goodenow, M.** 1994. Persistence of multiple maternal genotypes of human immunodeficiency virus type 1 in infants by vertical transmission. J Clin Invest. **93:**380-90.

124. **Lasky, L. A., Groopman, J. E., Fennie, C. W., Benz, P. M., Capon, D. J., Dowbenko, D. J., Nakamura, G. R., Nunes, W. M., Renz, M. E., and Berman, P. W.** 1986. Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein. Science. **233:**209-12.

125. **Le Gall, S., Erdtmann, L., Benichou, S., Berlioz-Torrent, C., Liu, L., Benarous, R., Heard, J. M., and Schwartz, O.** 1998. Nef interacts with the mu

subunit of clathrin adaptor complexes and reveals a cryptic sorting signal in MHC I molecules. Immunity. **8:**483-95.

126. **Learmont, J. C., Geczy, A. F., Mills, J., Ashton, L. J., Raynes-Greenow, C. H., Garsia, R. J., Dyer, W. B., McIntyre, L., Oelrichs, R. B., Rhodes, D. I., Deacon, N. J., and Sullivan, J. S.** 1999. Immunologic and virlogic status after 14 to 18 years of infection with an attenuated strain of HIV-1. A report from the Sydney Blood Bank Cohort. N Engl J Med. **340**.

127. **Learn, G. H., Korber, B. T., Foley, B., Hahn, B. H., Wolinsky, S. M., and Mullins, J. I.** 1996. Maintaining the integrity of human immunodeficiency virus sequence databases. J Virol. **70:**5720-5730.

128. **Lee, C. H., Leung, B., Lemmon, M. A., Zheng, J., Cowburn, D., Kuriyan, J., and Saksela, K.** 1995. A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. Embo J. **14:**5006-15.

129. **Lee, C. H., Saksela, K., Mirza, U. A., Chait, B. T., and Kuriyan, J.** 1996. Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. Cell. **85:**931-42.

130. **Li, W. H.** 1997. Molecular Evolution, p. pp.487. Sinauer Associates, Inc., Sunderland, MA.

131. **Liu, S. L., Schacker, T., Musey, L., Shriner, D., McElrath, M. J., Corey, L., and Mullins, J. I.** 1997. Divergent patterns of progression to AIDS after infection from the same source: human immunodeficiency virus type 1 evolution and antiviral responses. J Virol. **71:**4284-95.

132. **Ljunggren, K., Moschese, V., Broliden, P. A., Giaquinto, C., Quinti, I., Fenyo, E. M., Wahren, B., Rossi, P., and Jondal, M.** 1990. Antibodies mediating cellular cytotoxicity and neutralization correlate with a better clinical stage in children born to human imunodeficiency virus-infected mothers. J Infect Dis. **191:**198-202.

133. **Lu, X., Yu, H., Liu, S. H., Brodsky, F. M., and Peterlin, B. M.** 1998. Interactions between HIV1 Nef and vacuolar ATPase facilitate the internalization of CD4. Immunity. **8:**647-56.

134. **Luciw, P. A.** 1996. Human Immunodeficiency Viruses and Their Replication, p. 845-916. *In* B. N. Fields and D. M. Knipe and P. M. Howley (ed.), Fundamental Virology, Third Edition ed. Lippincott-Raven Publishers, Philadelphia.

135. **Lukashov, V. V., Kuiken, C. L., and Goudsmit, J.** 1995. Intrahost human immunodeiciency virus type 1 evolution is related to length of the immunocompetent period. J Virol. **69:**6911-6916.

136. **Mangasarian, A., Foti, M., Aiken, C., Chin, D., Carpentier, J. L., and Trono, D.** 1997. The HIV-1 Nef protein acts as a connector with sorting pathways in the Golgi and at the plasma membrane. Immunity. **6:**67-77.

137. **Mangasarian, A., Piguet, V., Wang, J. K., Chen, Y. L., and Trono, D.** 1999. Nef-induced CD4 and major histocompatibility complex class I (MHC-I) down-regulation are governed by distinct determinants: N-terminal alpha helix and proline repeat of Nef selectively regulate MHC-I trafficking. J Virol. **73:**1964-73.

138. **Mansky, L. M., and Termin, H. M.** 1995. Lower in-viivo mutation rate of human immunodeficiency virus type 1 than that predicted by from fidelity of purified reverse transcriptase. J Virol. **69:**5087-5094.

139. **Mariani, R., Kirchhoff, F., Greenough, T. C., Sullivan, J. L., Desrosiers, R. C., and Skowronski, J.** 1996. High frequency of defective nef alleles in a long-term survivor with nonprogressive human immunodeficiency virus type 1 infection. J Virol. **70:**7752-64.

140. **Massiah, M. A., Starich, M. R., Paschall, C., Summers, M. F., Christensen, A. M., and Sundquist, W. I.** 1994. Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein. J Mol Biol. **244:**198-223.

141. **Matala, E. J.** 1999. University of Arizona.

142. **Matala, E. J., Crandall, K. A., Baker, R. C., and Ahmad, N.** 2000. Limited heterogeneity of HIV type 1 in infected mothers correlates with lack of vertical transmission. AIDS Res Hum Retroviruses. **16:**1481-1489.

143. **Matala, E. J., Hahn, T., Yedavalli, V. R. K., and Ahmad, N.** 2001. Biological Characterization of the HIV-1 env V3 regions from Mothers and Infants associated with perinatal Transmission. AIDS Res Hum Retroviruses. **17:**1725-1735.

144. **Matala, E. J., Husain, M., Crandall, K. A., and Ahmad, N.** 2002, manuscript in preparation. Genetic Characterization of Human Immunodeficiency Virus Type 1 gp120 isolates associated with and lack of vertical transmission.

145. **Mattern, C. F., Murray, K., Jensen, A., Farzadegan, H., Pang, J., and Modlin, J. F.** 1992. Localization of human immunodeficiency virus core antigen in term human placentas. Pediatrics. **89:**207-9.

146. **MaWhinney, S., Pagano, M., and Thomas, P. A.** 1993. Age at AIDS diagnosis for children with perinatally acquired HIV. J AIDS. **6:**1139-1144.

147. **Mayaux, M.-J., Blanche, S., Rouzioux, C., Le Chenadec, J., Chambrin, V., Firtion, G., Allemon, M.-C., Vilmer, E., Vigneron, N. C., Tricoire, J., Guillot, F., Courpotin, C., and French Pediatric HIV Infection Study Group.** 1995. Maternal factors associated with perinatal HIV-1 transmission: The French Cohort Study: 7 years of follow up observation. J Acquir Immune Defic Syndr Hum Retrovirol. **8:**188-194.

148. **McNearney, T., Hornickova, Z., Markham, R., Birdwell, A., Arens, M., Saah, A., and Ratner, L.** 1992. Relationship of human immunodeficiency virus type 1 sequence heterogeneity to stage of disease. Proceedings of the National Academy of Science, USA. **89:**10247-10251.

149. **Mehta, S. U., Rupprecht, K. R., Hunt, J. C., Kramer, D. E., McRae, B. J., Allen, R. C., Dawson, G. J., and Devane, S. G.** 1990. Prevalence of antibodies to the core protein p17, and a serlogical marker during HIV-1 infection. AIDS Res Hum Retroviruses. **6:**440-445.

150. **Miller, M. D., Farnet, C. M., and Bushman, F. D.** 1997. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. J Virol. **71:**5382-90.

151. **Mock, P. A., Shaffer, N., Bhadrakom, C., Siriwasin, W., Chotpitayasunondh, T., Chearskul, S., Young, N. L., Roongpisuthipong, A., Chinayon, P., Kalish, M. L., Parekh, B., and Mastro, T. D.** 1999. Maternal viral load and timing of mother-to-child HIV transmission, Bangkok, Thailand. Bangkok Collaborative Perinatal HIV Transmission Study Group. AIDS. **13:**407-14.

152. **Mok, J. Q., Giaquinto, C., DeRossi, A., Gurch-Worner, I., Ades, A. E., and Pekham, C. S.** 1987. Infants born to mothers seropositive for human

immunodeficiency virus - preliminary findings from a multiculture European study. Lancet. **1:**1164-1168.

153. **Myers, G., Korber, B., Foley, B., Jeang, K.-T., Mellors, J. W., and Wain-Hobson, S.** 1996. Human retroviruses and AIDS database. Thoeretical Biology, Los Alamos National Laboratory.

154. **Myers, G., MacInnes, K., and Korber, B.** 1992. The emergence of simian/human immunodeficiency viruses. AIDS Res Hum Retroviruses. **8:**373-386.

155. **Nardiello, S., Pizzella, T., Russo, M., and Galanti, B.** 1982. Esterase staining of monocytes in suspensions of ficoll-paque isolated mononuclear cells. J Immunol Methods. **52:**111-114.

156. **Narwa, R., Roques, P., Courpotin, C., Parnet-Mathieu, F., Boussin, F., Roane, A., Marce, D., Lasfargues, G., and Dormont, D.** 1996. Characterization of human immunodeficiency virus type 1 p17 matrix protein motifs associated with mother-to-child transmission. J Virol. **70:**4474-4483.

157. **Nei, M., and Gojobori, T.** 1986. Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. Mol Biol and Evol. **3.**

158. **Niederman, T. M., Hastings, W. R., and Ratner, L.** 1993. Myristoylation-enhanced binding of the HIV-1 Nef protein to T cell skeletal matrix. Virology. **197:**420-5.

159. **O'Shea, S., Newell, M. L., Dunn, D., and al, e.** 1997. HIV viral load in plasma and cervico-vaginal secretions of pregnant and non-pregnant women., p. 63. Conference on Global Strategies for the Prvention of HIV Transmission from Mothers to Infants.

160. **Oram, J. D., Downing, R. G., Roff, M., Sernankambo, N., Clegg, J. C. S., Featherstone, A.-S. R., and Booth, J. C.** 1991. Sequence analysis of the V3 loop regions of the env genes of Ugandian human immunodeficiency proviruses. AIDS Res Hum Retroviruses. **7:**605-614.

161. **Overbaugh, J., Anderson, R. J., Ndinya-Achola, J. O., and Kreiss, J. K.** 1996. Distinct but related human immunodeficiency virus type 1 variant

populations in genital secretions and blood. AIDS Res Hum Retroviruses. **12:**107-15.

162. **Palumbo, P. E., Kwok, S., and Waters, S.** 1995. Viral measurement by polymerase chain reaction-based assays in human immunodeficiency virus-infected infants. J Pediatr. **126:**592-595.

163. **Pasquier, C., Cayrou, C., Blancher, A., Tourne-Petheil, C., Berrebi, A., Tricoire, J., Puel, J., and Izopet, J.** 1998. Molecular evidence for mother-to-child transmission of multiple variants by analysis of RNA and DNA sequences of human immunodeficiency virus type 1. J Virol. **72:**8493-501.

164. **Perelson, A. S., Neuman, A. U., Markowitz, M., leonard, J. M., and Ho, D.** 1996. HIV-1 dynamics in-vivo: Virion clearance rate, infected cell life-span, and viral generation time. Science. **271**.

165. **Piguet, V., Gu, F., Foti, M., Demaurex, N., Gruenberg, J., Carpentier, J. L., and Trono, D.** 1999. Nef-induced CD4 degradation: a diacidic-based motif in Nef functions as a lysosomal targeting signal through the binding of beta-COP in endosomes. Cell. **97:**63-73.

166. **Pitt, J., P., B., Landay, A., McIntosh, K., Burns, D., Hillyer, G. V., Mendez, H., and and Fowler, M. G.** 1997. Maternal immunologic and virologic risk factors for infant human immunodeficiency virus type 1 infection: Findings from Women and Infants Transmission Study. J Infect Dis. **175:**567-575.

167. **Posada, D., and Crandall, K. A.** 1998. Modeltest: Testing the model of DNA substitution. Bioinformatics. **14:**817-818.

168. **Poss, M., Martin, H. L., Kreiss, J. K., Granville, L., Chohan, B., Nyange, P., Mandaliya, K., and Overbaugh, J.** 1995. Diversity in virus populations from genital secretions and peripheral blood from women recently infected with human immunodeficiency virus type 1. J Virol. **69:**8118-22.

169. **Reicin, A. S., Ohagen, A., Yin, L., Hoglund, S., and Goff, S. P.** 1996. The role of gag in human immunodeficiency virus type 1 virion morphogenesis and early steps in the viral life cycle. J Virol. **70:**8645-8652.

170. **Reicin, A. S., Paik, S., Berkowitz, R. D., Luban, J., Lowy, I., and Goff, S. P.** 1995. Linker insertion mutations in the human immunodeficiency virus type 1

gag gene: effects on virion particle assembly, release, and infectivity. J Virol. **69:**642-50.

171. **Rhee, S. S., and Marsh, J. W.** 1994. Human immunodeficiency virus type 1 Nef-induced down-modulation of CD4 is due to rapid internalization and degradation of surface CD4. J Virol. **68:**5156-63.

172. **Romiti, M. L., Colognesi, C., Cancrini, C., Mas, A., Berrino, M., Salvatori, F., Orlandi, P., Jansson, M., Palomba, E., Plebani, A., Bertran, J. M., Hernandez, M., de Martino, M., Amoroso, A., Tovo, P. A., Rossi, P., Espanol, T., and Scarlatti, G.** 2000. Prognostic value of a CCR5 defective allele in pediatric HIV-1 infection. Mol Med. **6:**28-36.

173. **Ross, T. M., Oran, A. E., and Cullen, B. R.** 1999. Inhibition of HIV-1 progeny virion release by cell-surface CD4 is relieved by expression of the viral Nef protein. Curr Biol. **9:**613-21.

174. **Rusche, J. R., Javaherian, K., McDanal, C., Petro, J., Lynn, D. L., Grimailia, R., Langois, A., Gallo, R. C., Arthur, L. O., Fischinger, P. J., Bolognesi, D. P., Putney, S. D., and Mathews, T. J.** 1988. Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120. Proceedings of the National Academy of Science, USA. **85:**3198-202.

175. **Ryder, R. W., Nsa, W., Hassig, S. E., Behets, F., Rayfield, M., Kungola, E., Nelson, A., Mulenda, U., Francis, H., Mwandagalirwa, F., Davachi, F., Rogers, M., Nzilambi, N., Greenberg, A., Mann, J., Quinn, T. C., Piot, P., and Curran, J. W.** 1988. Perinatal transmission of human immunodeficiency virus type 1 to infants of seropositive women in Zaire. N Engl J Med. **320:**1637-42.

176. **Saitou, N., and Nei, M.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol and Evol. **4:**406-425.

177. **Saksela, K., Cheng, G., and Baltimore, D.** 1995. Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4. Embo J. **14:**484-91.

178. **Salvatori, F., and Scarlatti, G.** 2001. Hiv type 1 chemokine receptor usage in mother-to-child transmission. AIDS Res Hum Retroviruses. **17:**925-35.

179. **Sambrook, J., Fritsch, E. F., and Maniatis, T.** 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

180. **Samson, M., Libert, F., and Doranz, B. J.** 1996. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR5 chemonkine receptor gene. Nature. **382:**722-725.

181. **Sanger, F., Nicklen, S., and Coulson, A. R.** 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Science, USA. **74:**5463-7.

182. **Sawai, E. T., Khan, I. H., Montbriand, P. M., Peterlin, B. M., Cheng-Mayer, C., and Luciw, P. A.** 1996. Activation of PAK by HIV and SIV Nef: importance for AIDS in rhesus macaques. Curr Biol. **6:**1519-27.

183. **Scarlatti, G., Leitner, T., Hapi, E., Wahlberg, J., Marchisi, P., Clerici-Schoeller, M. A., Wigzell, H., Fenyo, E. M., Albert, J., Uhlen, M., and Rossi, P.** 1993. Comparison of variable region 3 sequences of human immunodeficiency virus type 1 from infected children with RNA DNA sequences of virus population of their mothers. Proc Natl Acad Sci USA. **90:**1721-1725.

184. **Schrager, L. K.** 1994. Long-term survivors of HIV-1 infection: Definitions and research challanges. AIDS. **8:**S95-S108.

185. **Schwartz, D. A., and Nahmias, A. J.** 1991. Human immunodeficiency virus and the placenta. Current concepts of vertical transmission in relation to other viral agents. Ann Clin Lab Sci. **21:**264-74.

186. **Schwartz, D. H., Sharma, U. K., Perlman, E. J., and Blakemore, K.** 1995. Adherence of human immunodeficiency virus-infected lymphocytes to fetal placental cells: a model of maternal --> fetal transmission. Proc Natl Acad Sci U S A. **92:**978-82.

187. **Schwartz, O., Marechal, V., Le Gall, S., Lemonnier, F., and Heard, J. M.** 1996. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. Nat Med. **2:**338-42.

188. **Scott, G. B., Fischl, M. A., and Khmas, N.** 1985. Mothers of infants with aquired immunodeficiency syndrome: Evidence for both symptomatic and asymptomatic carriers. JAMA. **253:**363-366.

189. **Scott, G. B., Hutto, C., Makuch, R. W., Mastrucci, M. T., O'Connor, T., D.Mitchell, C., Trapido, E. J., and Parks, W. P.** 1989. Survival in children with perinatally acquired immunodeficiency virus type 1 infection. N Engl J Med. **321:**1791-1796.

190. **Scott, R. J.** 1994. Danforth's Obstetrics and Gynelcology, 7 ed. Lippincott-Raven, Philadelphia.

191. **Shankarappa, R., Gupta, P., Learn, G. H., Jr., Rodrigo, A. G., Rinaldo, C. R., Jr., Gorry, M. C., Mullins, J. I., Nara, P. L., and Ehrlich, G. D.** 1998. Evolution of human immunodeficiency virus type 1 envelope sequences in infected individuals with differing disease progression profiles. Virology. **241:**251-9.

192. **Shankarappa, R., Margolick, J. B., Gange, S. J., Rodrigo, A. R., Upchurch, D., Farzadegan, H., Gupta, P., Rinaldo, C. R., Learn, G. H., He, X., Huang, X., and Mullins, J.** 1999. Consistent viral evolutionary changes associated with progression of human immunodeficiency virus yype 1 inefction. J Virol. **73:**10489-10502.

193. **Sharp, P. M.** 1997. In search of molecular darwinism. Nature. **385:**111-112.

194. **Shioda, T., Levy, J. A., and Cheng-Mayer, C.** 1991. Macrophage and T cell line tropism of HIV-1 are determined by specific regions of the envelope gp120 gene. Nature. **349:**167-169.

195. **Shugars, D. C., Smith, M. S., Glueck, D. H., Nantermet, P. V., Seillier-Moiseiwitsch, F., and Swanstrom, R.** 1993. Analysis of human immunodeficiency virus type 1 nef gene sequences present in vivo. J Virol. **67:**4639-50.

196. **Simon, F., Mauclere, P., Roques, P., Loussert-Ajaka, I., Muller-Trutwin, M. C., Saragosti, S., Georges-Courbot, M. C., Barre-Sinoussi, F., and Brun-Vezinet, F.** 1998. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. Nat Med. **4:**1032-7.

197. **Simonds, R. J., Steketee, R., Nesheim, S., Matheson, P., Palumbo, P., Alger, L., Abrams, E. J., Orloff, S., Lindsay, M., Bardeguez, A. D., Vink, P., Byers, R., and Rogers, M.** 1998. Impact of zidovudine use on risk and risk factors for perinatal transmission of HIV. Perinatal AIDS Collaborative Transmission Studies. AIDS. **12:**301-8.

198. **Slatkin, M., and Hudson, R. R.** 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics. **129:**555-562.

199. **Smith, B. L., Krushelnycky, B. W., Mochly-Rosen, D., and Berg, P.** 1996. The HIV nef protein associates with protein kinase C theta. J Biol Chem. **271:**16753-7.

200. **Sonnerborg, A., and Sherefa, K.** 1997. HIV-1 subtype C epidemic in Ethiopia. AIDS. **11:**1897.

201. **Spearman, P., Wang, J.-J., Heyden, N. V., and Ratner, L.** 1994. Identification of human immunodeficiency virus type 1 Gag protein domains essential to membrane binding and particle assembly. J Virol. **68:**3232-3242.

202. **Sperling, R. S., Shapiro, D. E., and Coombs, R. W.** 1996. Maternal viral load, zidovudine treatment, and the risk of transmission of human immunodeficiency virus type 1 from mother to infant. Pediatric AIDS Clinical Trials Group Protocol 076 Study Group. N Engl J Med. **335:**1621-29.

203. **St Louis, M. E., Kamenga, M., Brown, C., Nelson, A. M., Manzila, T., Batter, V., Behets, F., Kabagabo, U., Ryder, R. W., Oxtoby, M., and et al.** 1993. Risk for perinatal HIV-1 transmission according to maternal immunologic, virologic, and placental factors. JAMA. **269:**2853-9.

204. **Stamatatos, L., and Cheng-Mayer, C.** 1993. Evidence that the structural conformation of the envelope gp120 affects HIV-1 infectivity, host range, and syncytium-forming ability. J Virol. **67:**5635-5639.

205. **Starcich, B. R., Hahn, B. H., Shaw, G. M., McNeely, P. D., Modrow, S., Wolf, H., Parks, E. S., Parks, W. P., Josephs, S. F., Gallo, R. C., and Wong-Staal, F.** 1986. Identification and characterization of conserved and variable regions in the envelope gene of HTLV III/LAV, the retorvirus of AIDS. Cell. **45:**637-648.

206. **Swingler, S., Gallay, P., Camaur, D., Song, J., Abo, A., and Trono, D.** 1997. The Nef protein of human immunodeficiency virus type 1 enhances serine phosphorylation of the viral matrix. J Virol. **71:**4372-7.

207. **Swingler, S., Mann, A., Jacque, J., Brichacek, B., Sasseville, V. G., Williams, K., Lackner, A. A., Janoff, E. N., Wang, R., Fisher, D., and**

Stevenson, M. 1999. HIV-1 Nef mediates lymphocyte chemotaxis and activation by infected macrophages. Nat Med. **5:**997-103.

208. **Swofford, D. L.** 1999. PAUP* Phylogenetic analysis using parsimony and other methods, 4.0.0b2 ed. Sinauer Associates, Sunderland, MA.

209. **Takeuchi, Y., Akutsu, M., Murayama, K., Shimizu, N., and Hoshino, H.** 1991. Host range mutant of human immunodeficiency virus type 1: Modification of cell tropism by a single point mutation at the neutralization epitope in the env gene. J Virol. **65:**1710-1718.

210. **The European Collaborative Study Group, Newell, M. L., Dunn, D. T., Peckham, C. S., Semprini, A. E., and Pardi, G.** 1996. Vertical transmission of HIV-1: maternal immune status and obstetric factors. The European Collaborative Study. AIDS. **10:**1675-81.

211. **Thea, D. M., Steketee, R. W., Pliner, V., Bornschlegel, K., Brown, T., Orloff, S., Matheson, P. B., Abrams, E. J., Bamji, M., Lambert, G., Schoenbaum, E. A., Thomas, P. A., Heagarty, M., and Kalish, M. L.** 1997. The effect of maternal viral load on the risk of perinatal transmission of HIV-1. New York City Perinatal HIV Transmission Collaborative Study Group. AIDS. **11:**437-44.

212. **Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G.** 1997. The clustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nuc Acids Res. **24:**4876-4882.

213. **Toohey, K., Wehrly, K., Nishio, J., Perryman, S., and Chesebro, B.** 1995. Human immunodeficiency virus envelope V1 and V2 regions influence replication efficiency in macrophages by affecting virus spread. Virology. **213:**70-79.

214. **Tovo, P. A., de Martino, M. D., and Gabiano, C.** 1992. Prognostic factors and survival in children with perinatal HIV-1 infection. Lancet. **339:**1249-1253.

215. **UNAIDS/WHO, and Joint United Nations Program on HIV/AIDS.** 2001. World AIDS Epedimic Report - Epidemic Update December 2001. UNAIDS/WHO.

216. **van de Peere, P., Lepage, P., Simonon, A., Desgranges, C., Hitimanna, D. G., and Bazubagria, A.** 1992. Biological markers associated with prolonged survival in African children maternally infected with human immunodeficiency virus type. AIDS Res Hum Retroviruses. **8:**435-422.

217. **van't Wout, A. B., Kootstra, N. A., Mulder-Kampinga, G. A., Albrecht-Van Lent, N., Scherpbier, H. J., Veenstra, J., Boer, K., Coutinho, R. A., Miedema, F., and Schuitmaker, H.** 1994. Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission. J Clin Invest. **94:**2060-2067.

218. **Vasco, D. A., Crandall, K. A., and Fu, Y.-X.** 2001. Molecular Population Genetics: Coalescent methods based on Summary Statistics, p. 173-216. *In* A. G. Rodrigo and J. G. H. Learn (ed.), Computational and Evolutionary Analysis of HIV Molecular Sequences. Kluwer Academic Publishers, Dordrecht, The Netherlands.

219. **Vodicka, M. A., Goh, W. C., Wu, L. I., Rogel, M. E., Bartz, S. R., Schweickart, V. L., Raport, C. J., and Emerman, M.** 1997. Indicator cell lines for detection of primary strains of human and simian immunodeficiency viruses. Virology. **233:**193-8.

220. **von Schwedler, U., Kornbluth, R. S., and Trono, D.** 1994. The nuclear localization signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescent T lymphocytes. Proc Natl Acad Sci U S A. **91:**6992-6.

221. **Watterson, G. A.** 1975. On the number of segregating sites in genetical models without recombination. Theor Pop Biol. **7:**256-276.

222. **Welker, R., Harris, M., Cardel, B., and Krausslich, H. G.** 1998. Virion incorporation of human immunodeficiency virus type 1 Nef is mediated by a bipartite membrane-targeting signal: analysis of its role in enhancement of viral infectivity. J Virol. **72:**8833-40.

223. **Westervelt, P., Gendelman, H. E., and Ratner, L.** 1991. Identification of a determinant within the human immunodeficiency virus 1 surface envelope glycoprotein critical for productive infection of primary monocytes. Proc Natl Acad Sci U S A. **88:**3097-101.

224. **Westervelt, P., Trowbridge, D. B., Epstein, L. G., Blumberg, B. M., Li, Y., Hahn, B. H., Shaw, G. M., Price, R. W., and Ratner, L.** 1992. Macrophage tropism determinants of human immunodeficiency virus type 1 *in-vivo*. J Virol. **66:**2577-2582.

225. **Willey, R. L., Ross, E. K., Buckler-White, A. J., Theodore, T. S., and Martin, M. A.** 1989. Functional interaction of constant and variable domains of human immunodeficiency virus type 1 gp120. J Virol. **63:**3595-600.

226. **Willey, R. L., Smith, D. H., Lasky, L. A., Theodor, T. S., Earl, P. L., Mars, B., Capon, D. J., and Martin, T. J.** 1988. In vitro mutagenesis identifies a region witht the envelope gene of human immunodeficiency virus that is critical for infectivity. J Virol. **62:**139-47.

227. **Willey, R. L., Theodore, T., and Martin, M.** 1994. Amino acid substitutions in the HIV-1 gp120 V3 loop that change viral tropism also alter physical and functional properties of the virion envelope. J Virol. **68:**4409-4419.

228. **Wills, J. W., and Craven, R. C.** 1991. Form, function, and use of retroviral gag proteins. AIDS. **5:**639-54.

229. **Wolfs, T. F. W., Zwart, G., Bakker, M., and Goudsmit, J.** 1992. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. Virology. **189:**103-110.

230. **Wolinsky, S., Korber, B. T. M., Neumann, A. E., Daniels, M., Kunstman, K. J., Whetsell, A. J., Furtado, M. R., Cao, Y., Ho, D. D., Safrit, J. T., and Koup, R. A.** 1996. Adaptive evolution of human immunodeficiency virus type-1 during the natural course of infection. Science. **272:**537-542.

231. **Wolinsky, S., Wike, C. M., Korber, B. T. M., Hutto, C., Parks, W. P., Rosenblum, L. L., Kunstman, K. J., Furtado, M. R., and Munzo, I. L.** 1992. Selective transmission of human immunodeficiency virus type-1 variants from mother to infants. Science. **255:**1134-1137.

232. **Xu, X. N., Laffert, B., Screaton, G. R., Kraft, M., Wolf, D., Kolanus, W., Mongkolsapay, J., McMichael, A. J., and Baur, A. S.** 1999. Induction of Fas ligand expression by HIV involves the interaction of Nef with the T cell receptor zeta chain. J Exp Med. **189:**1489-96.

233. **Xu, X. N., Screaton, G. R., Gotch, F. M., Dong, T., Tan, R., Almond, N., Walker, B., Stebbings, R., Kent, K., Nagata, S., Stott, J. E., and McMichael, A. J.** 1997. Evasion of cytotoxic T lymphocyte (CTL) responses by nef-dependent induction of Fas ligand (CD95L) expression on simian immunodeficiency virus-infected cells. J Exp Med. **186:**7-16.

234. **Yang, X., and Gabuzda, D.** 1999. Regulation of human immunodeficiency virus type 1 infectivity by the ERK mitogen-activated protein kinase signaling pathway. J Virol. **73:**3460-6.

235. **Yang, Z.** 1996. Maximum-Likelihood Models for Combined Analyses of Multiple Sequence Data. J Mol Evol. **42:**587-96.

236. **Yedavalli, V. R., Husain, M., Horodner, A., and Ahmad, N.** 2001. Molecular characterization of HIV type 1 vpu genes from mothers and infants after perinatal transmission. AIDS Res Hum Retroviruses. **17:**1089-98.

237. **Yedavalli, V. R. K., and Ahmad, N.** 2001. Genetic Characterization of HIV-1 Type 1 vif and vpr Genes from infected Mothers Lacking Vertical Transmission. AIDS Res Hum Retroviruses. **17:**911-923.

238. **Yedavalli, V. R. K., Chappey, C., and Ahmad, N.** 1998. Maintenance of an intact human immunodeficiency virus type 1 *vpr* gene folowing mother-to-infant transmission. J Virol. **72:**6937-6943.

239. **Yedavalli, V. R. K., Chappey, C., Matala, E., and Ahmad, N.** 1998. Conservation of an intact *vif* gene of human immunodeficiency virus type 1 during maternal-fetal transmission. J Virol. **72:**1092-1102.

240. **Yokoyama, S., and Gojobori, T.** 1987. Molecular evolution and phylogeny of the human AIDS viruses LAV, HTLV-III and ARV. J Mol Evol. **24:**330-336.

241. **Yu, X., Yu, Q.-C., Lee, T.-H., and Esex, M.** 1992. The C-terminus of human immunodeficiency virus type 1 matrix protein is involved in early steps of the virus life cycle. J Virol. **66:**5657-5670.

242. **Yu, X., Yuan, X., Matsuda, Z., Lee, T.-H., and Essex, M.** 1992. The matrix protein of human immunideficiency virus type 1is required for incorporation of viral evelope protein into mature virions. J Virol. **66.**

243. **Yuan, X., Yu, X., Lee, T.-H., and Essex, M.** 1993. Mutations in the N-terminal region of the human immunodeficiency virus type 1 matrix protein block intracellular transport of the Gag precursor. J Virol. **67:**6387-6394.

244. **Zazopoulos, E., and Haseltine, W. A.** 1993. Disulfide bond formation in the human immunodeficiency virus type 1 Nef protein. J Virol. **67:**1676-80.

245. **Zhang, L. Q., MacKenzie, P., Cleland, A., Holmes, E. C., Brown, A. J., and Simmonds, P.** 1993. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J Virol. **67:**3345-56.

246. **Zhou, W., Parent, L. J., Wills, J. W., and Resh, M. D.** 1994. Identification of a membrane binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J Virol. **68:**2556-69.

247. **Zhu, T., Mo, H., Wang, N., Nam, D. S., Cao, Y., Koup, R. A., and Ho, D. D.** 1993. Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. Science. **261:**1179-1181.

248. **Zhu, T., Wang, N., Carr, A., Nam, S., Moor-Jankowski, R., Cooper, D., and Ho, D. D.** 1996. Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: Evidence for viral compartmentalization and selection during sexual transmission. J Virol. **70:**3098-3107.