SIV envelope evolution and virus virulence

Valli, P.J.S.

Publication date
2000
Document Version
Final published version

SIV Envelope Evolution and Virus Virulence

Peter John Spencer Valli

SIV Envelope Evolution and Virus Virulence

Peter John Spencer Valli

SIV Envelope Evolution and Virus Virulence

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam,
op gezag van de Rector Magnificus Prof. Dr. J.J.M. Franse
ten overstaan van een door het college voor promoties ingestelde
commissie in het openbaar te verdedigen in de Aula der Universiteit
Oude Lutherse Kerk, ingang Singel 411, hoek Spui,
op woensdag 7 juni 2000 te 10.00 uur

door

Peter John Spencer Valli

geboren te Brooks

Promotor	Prof. Dr. J. Goudsmit
Copromotor	Dr. J.L. Heeney
Beoordelingscommissie	Prof. Dr. R A. Coutinho Prof. Dr. S. Danner Prof. Dr. J. M. A. Lange Prof. Dr. F. Miedema Prof. Dr. A. D. M. E. Osterhaus

This research described in this thesis was conducted in the Department of Human Retrovirology (head Prof. Dr. Goudsmit) of the University of Amsterdam, Academic Medical Centre.

Without scholarship there is no history, only fashion.

Sir Geoffrey Elton

Contents

Chapter 1

SIV and HIV: A General Introduction

SIV and HIV: A General Introduction

The Start of the AIDS Epidemic and the Discovery of HIV:

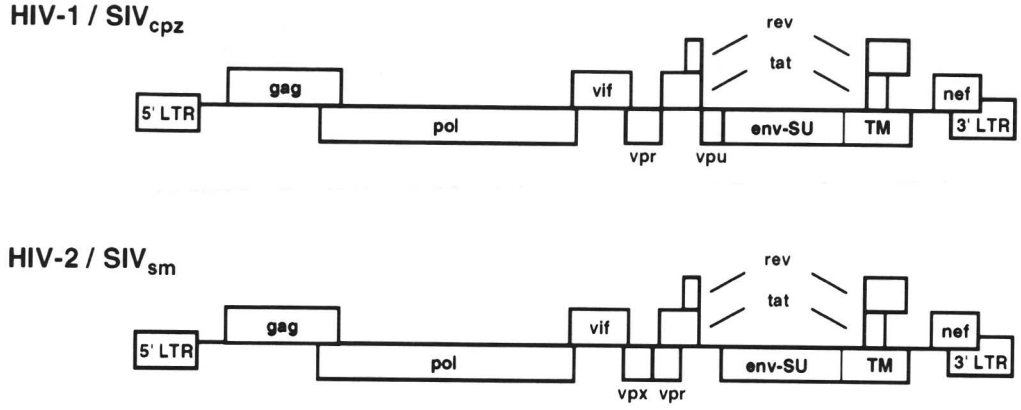
In the early 1980's an epidemic of Kaposi's Sarcoma and other opportunistic infections emerged among homosexual men in the United States (13). Originally considered a rare syndrome, AIDS subsequently spread among intravenous drug users, haemophiliacs, partners of infected individuals, blood product recipients, and also the infants of mothers with AIDS. The epidemiology of AIDS suggested an infectious etiology. In 1983 a French group led by Montagnier first isolated a lentivirus from patient material which then became known as LAV (now known as Human Immunodeficiency Virus 1 or HIV-1 (5)). First recognized in the United States, reports of HIV-1 seronegative AIDS later came from West-Africa, and in 1986 the same French team isolated another HIV-1 related virus designated HIV-2 (18). The gene order as well as phylogenetic relationship of HIV-1 and HIV-2 (Figure #1) made it clear that the two viruses were of the same family. Phenotypic differences were observed with HIV-2 infection appearing to be much less virulent, and less infectious according to lower transmission rates.

AIDS in Non-Human Primates:

In the 1980's an increase in the incidence of lymphoma in rhesus macaques in the California Regional Primate Research Center and subsequently in the New England and Tulane primate centers were observed. Unlike the first increase in lymphomas in the 1960's and 1970's these tumors were not consistently positive for SRV (Simian Retrovirus (typeD)(53)), certain serotypes of which cause retroperitoneal fibromas in rhesus monkeys. Years earlier leprosy in African sooty mangabeys was being studied at the Tulane Primate Research Center (6). As this species was endangered there was an effort to develop another animal model for leprosy using rhesus macaques. Tissue from lepromatous leprosy lesions from sooty mangabeys was used to inoculate Asian rhesus macaques. It is highly likely that these transmission experiments led to the transmission of Simian Immunodeficiency Virus (SIV) from these sooty mangabeys to rhesus macaques. Unaware of the asymptomatic infection of sooty mangabeys with SIVsm (the small letters denote the species specific strain of the virus) the housing of these animals with rhesus macaques likely led to the spread of the infection. Simultaneously, similar outbreaks in rhesus macaques were observed at the California Regional Primate Research Center (35). Evidence suggests that asymptomatic of SIVsm carriers were transferred from Tulane and then from California to other research centers, precipitating the outbreaks of lymphomas at Tulane, California, and the New England Primate Research Center. Antibodies of affected animals reacted with some antigens of HIV-1, and a subsequent study revealed a significant disease association with transmission of tissues from animals which developed lymphoid neoplasms (42,55). The first virus isolations (21,44,54,60), and their use in experimental infections, revealed that an immunodeficiencey syndrome could be reproduced in Asian nonhuman primates with an etiologic agent related to HIV-2 (subsequently called SIV). Certain variants of SIVsm were found that were extremely pathogenic (31), and were studied extensively although the disease characteristics were not very characteristic of those of AIDS in humans. Designated SIVsmPBj14 this isolate caused enormous immune activation (23) characterized by extensive lymphoid hyperplasia and haemorrhagic enteritis (32), displaying unique lentivirus-host interactions resembling superantigen-like activity (67). The linking of the epidemic of lymphomas, opportunistic infections, and AIDS in macaques with a lentivirus recovered from asymptomatic sooty mangabeys (30,60) allowed investigations to trace the origin and spread of the primate

epidemic in the major primate centres. The viruses isolated from sooty mangabeys (6,30,60) were later found to be natural asymptomatic infections in this species. These findings sparked the search for the origin of the lineage of primate lentiviruses and lead to the discovery of more and more primate specific lentiviruses causing asymptomatic infections in African nonhuman primates.

Figure #1: The genomic organization of the primate lentiviruses HIV-1/SIVcpz and HIV-2/SIVsm. Both have similar patterns of genomic organization and length (between 9 and 10 kb) of linear double-stranded proviral DNA. Structural genes are shaded; regulatory and accessory genes are are represented by black squares. *Vpu* is found exclusively in HIV-1 and SIVcpz.



Natural Asymptomatic Infection in Africa:

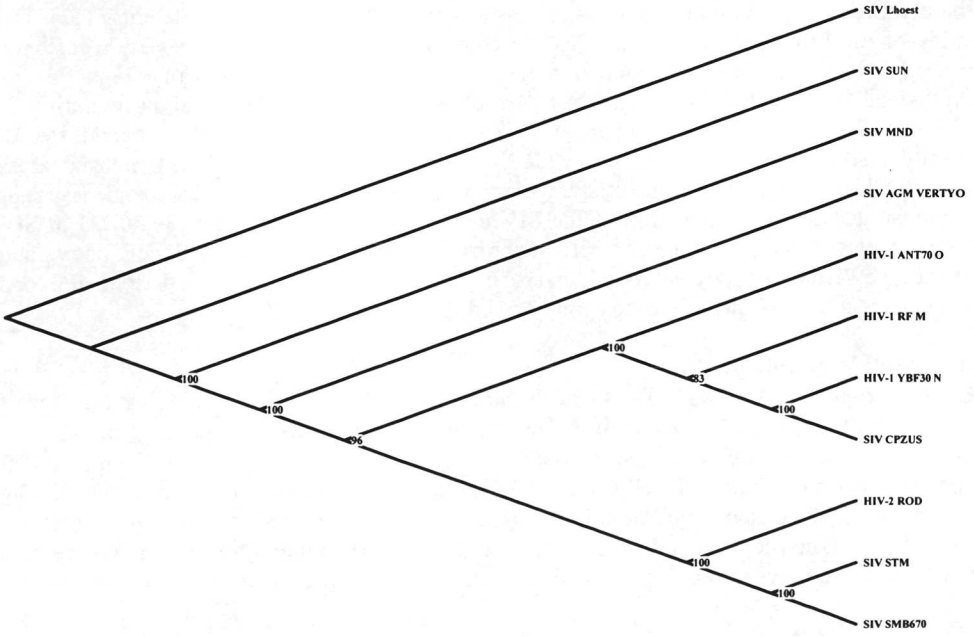
The five major lineages of currently known primate lentiviruses are designated SIVcpz from chimpanzees, SIVagm from African green monkeys, SIVsm from sooty mangabeys, SIVsun from sun-tailed monkeys, and SIVl'hoest from L'hoest monkeys (together with SIVmnd from mandrills). Based upon genetic analysis the origin of HIV-2 was traced to SIVsm transmitted to humans (34,40). Virus characterization studies revealed the geographic clustering of HIV-2 strains in humans with SIVsm in feral sooty mangabeys, supporting the hypothesis that each HIV-2 subtype originated from widely divergent SIVsm strains (16,37). This led to an extensive, and still ongoing, search of the African continent for new virus strains from all of the nonhuman primate species. Knowledge of the ability or inability to infect other species has added to the understanding of the transmission of lentiviruses between man and the nonhuman primates. The ability to efficiently infect other species is essential to the spread from one species to another. Of the 92 species of monkeys in Africa currently only seven (possibly eight, SIVrcm from a Red capped mangabey has an original *env* but appears to be a chimaera (36,69)) are known to be infected with distinct viruses of the HIV/SIV family. The greatest problem is the difference between viral phylogenies and the actual distribution history of the virus itself. The extinction of lineages during history and the fragmentary recompilation of the development of the pandemic from the information available only shows a fraction of what may have taken place. The recent discovery of a new SIVcpz isolate (SIVcpzUS) which clusters closer to HIV-1 than another isolate from chimpanzees (SIVcpzant) is important as it is most related to a new group of human isolates (type N (33)).

Phylogenetic analysis has provided data supporting the relationship of the human viruses HIV-1 with SIVcpz, and HIV-2 with SIVsm (Chapter 2, Figure 2 (38,61)). The nonhuman primate viruses being nonvirulent in their natural hosts and the human viruses being highly pathogenic.

Evidence for Cross-species Transmission between Primates:

Aside from the experimental or nosocomial transmission of SIVsm from African sooty mangabeys to Asian rhesus macaques, there are cases of transmission of SIVagm from sympatric African green monkeys to a yellow baboon, a chacma baboon, and a patas monkey (7). Studies in West Africa had shown a distinct clustering of the different HIV-2 strains in humans to be related to several sooty mangabey strains (34). Marx et al studied 197 captive and wild sooty mangabeys, and local inhabitants in West Africa. Each separate subtype of HIV-2 found in villagers clustered geographically with the SIV strain found in sooty mangabeys from that locale (16). There was greater variation between the viruses found within a feral troup of sooty mangabeys than there was between selected variants with individual sooty mangabeys and the HIV-2 found in the human inhabitants. Based on this observation they proposed that multiple transmissions had occurred between man and sooty mangabeys. More recently evidence for cross species transmission of SIVcpzUS a virus of Pan troglodytes troglodytes has been proposed as the closest lentivirus relative to HIV-1 in man, and a likely source for the current HIV-1 epidemic in humans ((33,37) Figure #2).

Figure #2: A phylogenetic tree of the complete outer membrane or envelope gene (*env* or gp160) of representative isolates of the HIV/SIV lentiviral family. *Env* is approximatley 3.5 kb in length and contains the regions of the virus which interact with the cellular membrane during binding, entry and infection, and is one of the main targets of the immune response. The relationship of the various genes is shown by the length of the branches of the tree, with a longer branch denoting a less related relative. Representative isolates of the virus, and their relationship to HIV-1 types M, N, and O, SIVcpz, HIV-2, and SIVsm are observed. The closely related viruses are believed to be the result of the cross species transmission from chimpanzee (SIVcpz) and sooty mangabey (SIVsm) to man (HIV-1 and HIV-2 respectively).



Assessment of the *in vitro* capabilities for transmission into cells of other species led to several set backs in models for HIV-1 infection. The known isolates from chimpanzees resemble that of HIV-1 but are of little experimental value as they do not infect lower species of primates (8). The only strains of the virus that do grow in human cells are SIVcpz, SIVsm, SIVl'hoest, and SIVsun. The cross species barriers of infection became even more complicated. This coupled with the species barriers between the African nonhuman primates (8) and their species specific strains (2) points to these four viruses as the existing relatives of the progenitors of the human lentiviruses HIV-1, and HIV-2.

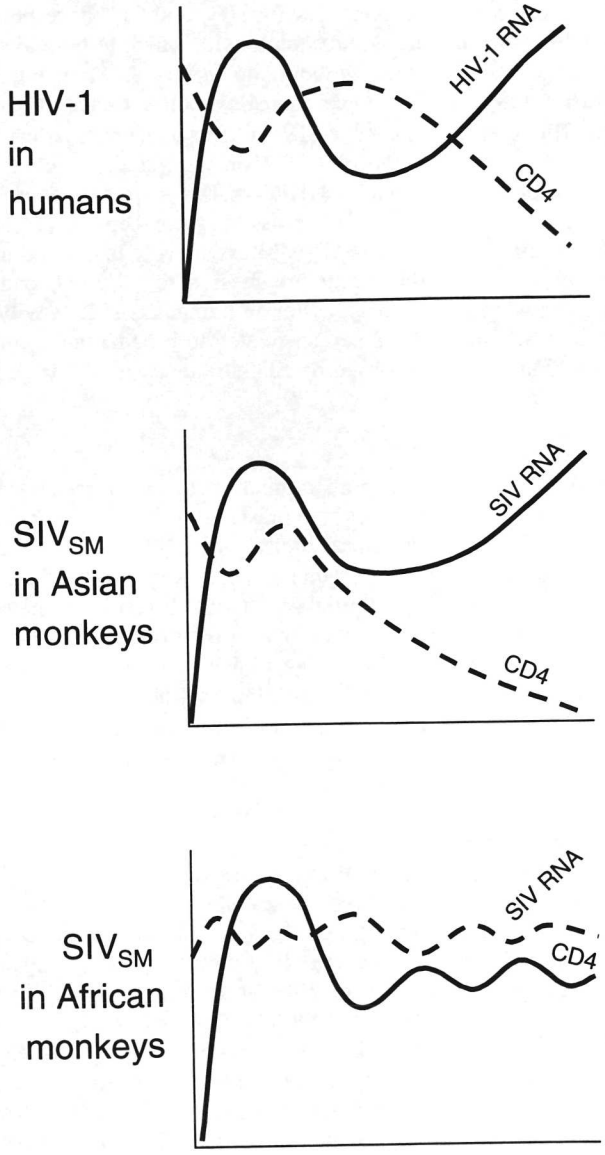
Molecular Evolution of SIV:

The finding that variation of only nine of the thousand amino acids in the envelope of the virus (Env, or gp160 (the translated protein), or *env* (the viral gene) could lead to a one hundred fold difference in replication between SIV variants (3) revealed the importance of small numbers of amino acid changes encoded in *env*. Large scale sequencing and analysis of *env* clones demonstrated that the hydrophillic and externalized regions were highly variable (V regions). The internalized and hydrophobic regions were found to be less variable (constant or C regions (24)) and have altered codon usage (70). The first variable regions of the SIV *env* were suggested to play a role in mucosal transmission (V1 and V2 (4,57)). The effect of glycosylation on the tertiary structure of *env* became apparent when the Ab binding profiles to successive Envs during infection were found to vary according to the addition or deletion of glycosylation sites in the V1 and V4 regions (64,65). The spontaneous changes brought about by the infidelity of the reverse transcriptase of the virus was thought to inadvertantly allow the selection of escape mutants (63) which Abs could no longer bind. The complexity of the folded structure of the HIV and SIV *env* and it's biological function in binding and infection showed that various nonadjacent areas of the nascent Env participate in the conformation and binding to CD4, accessory coreceptors and to facilitate entry (12). The study of the Env structure showed that the cooperation of numerous amino acids scattered throughout the internal and external areas of the folded protein cooperate to define its biological function (43,48,49,56,58,59). Antibody-antigen interactions followed during *in vivo* infection indicated that mutations in the *env* occur in a site specific manner, and these Ab sensitive sites were mapped to the V1,2,3,4 and V5 (virtually all the externalized areas (11,47)). Using mutagenesis and studying variants generated *in vivo,* two independent groups demonstrated that the V3 region of the SIV *env* did play a definite role (39,50,71) in SIV pathogenesis. Several mutations were mapped which markedly affected cell entry, and phenotype (from monocyte-derived macrophage tropic to infectious in peripheral blood mononuclear cells (PBMCs) or the complete abrogation of either (39,71)).

Biological Variation of SIV:

Cellular tropism of HIV and SIV is largely controlled by the *env* gene and is currently one of the most intensely studied areas of AIDS research. HIV and SIV infection is initiated by viruses with a macrophage tropic phenotype (Mϕ). Frequently during progression to AIDS HIV strains will switch to T cell tropism (TT). This TT phenotype characterized *in vitro* by growth in MT-2 cells and the formation of syncytia in these cells (SI for syncytium inducing, and NSI for nonsyncytium inducing viruses (20,22,51)). The cloning of the first coreceptors which HIV-1 utilizes during binding and entry (CCR5 and CXCR4), in conjunction with the

Figure #3: Different patterns of HIV and SIV infection *in vivo*. Shown (upper) is the typical picture observed of a progressor of HIV-1 infection in man. An initial peak in virus load is seen, followed by containment of viral replication, and then decreasing CD4 T-cell counts, eventual rise in viral load, and progression to AIDS. In long term nonprogressors of HIV-1 infection in man (middle) the initial virus load peak is contained for an indeterminate amount of time and no or minimal loss of CD4 T-cells is seen. The natural infection of nonhuman primates is characterized by to moderate high virus loads, and sustained normal amounts of CD4 T-cell counts with no progression to AIDS (lower).



CD4 T-cell receptor, was a major breakthrough in AIDS research (1,25,26,29). The family of G-protein-coupled seven-transmembrane-domain proteins (19,29) are the mediators of chemokine activity. The majority of Mϕ isolates appear to specifically use CCR5, a receptor for the CC chemokines RANTES, MIP-1α, and MIP-1β (19). The TT viruses predominantly use CXCR4, a receptor for the CXC chemokine SDF-1 (9). Subsequently the V3 mediated SI and NSI virus infections were shown to be mediated by a TT coreceptor (1,25), and a Mϕ coreceptor respectively (29). More recently coreceptors for HIV and SIV have been cloned (27). HIV and some SIV isolates use a myriad of coreceptors with which to bind to a cell and gain entry, and some isolates can do so even without the use of CD4, using only the coreceptor to gain cellular fusion (28). Human coreceptors allow cell entrance to SIV (14,17) without any loss of infection efficiency. In the case of SIVsm Mϕ and TT viruses can both use CCR5. Usage of the same coreceptor by the two types SIVsm virus is differential with two different domains of the coreceptor being utilized (26). HIV-2\SIVsm viruses generally prefer the CCR5 coreceptor, although they can utilize numerous other coreceptors (15,17,27,66). HIV-1 generally uses CCR5 in early infection, and CXCR4 exclusively or as well as CCR-5 (dual tropism) in late symptomatic infection during progression to AIDS. Broadening of coreceptor usage and its importance to disease progression in humans is still heavily debated (20,22). Progression to AIDS in SIV infection is neither related to broadening nor switching of coreceptor use (45). In fact CXCR4 is never used by SIV strains early or late in infection (14,45,46).

Determinants of Transmission of SIV:

The entry of SIV through non-abrasive mucosal oral, vaginal or rectal surfaces results in the selection of a small number of genomes (bottleneck transmission) and results in the decrease in genomic heterogeneity (10,57,62). Even with intravenous, mucosal or infected microgolial cell inoculation of virus the resulting effect is the same (52,62). Regardless of the route of infection, or the size of the inoculum, there is an initial selection that results in the decreased heterogeneity of virus. Later in the infection variants found in the primary inoculum but not predominant early in infection are frequently re-emerge (4,57). Transmission studies with SIVsmB670 have shown genotypic and phenotypic selection of Mϕ variants in the initial infection in neonates (4,57). The genomes present in the blood as measured by genotypic and phenotypic characteristics are influenced by humoral and cellular immune responses and the stage of infection.

The SIV Model Used in the Present Study:

Holterman and coworkers from the Biomedical Primate Research Centre (Rijswijk, The Netherlands) performed a passage experiment in rhesus macaques that was recently published with the department of Human Retrovirology (AMC, Amsterdam, The Netherlands (41)). Initially ten Asian rhesus macaques of an average age of 18 months were inoculated with a SIV B670 stock (6,60) originating from a virus infecting sooty mangabeys. It took seven months before the first of these rhesus macaques became ill, or showed signs of progression to AIDS. During a total of five passages of large inoculum the time to death decreased from seven to thirty three months during the first passage to one to two months during passages four and five respectively. A correlation was shown in this study between the plasma viral load and monkey survival. Plasma load as indicated by this study is clearly related to the virulence of the virus and/or the susceptibility of the host (68). The present thesis includes a genetic analysis of the *env* of the *in vivo* serial passage of SIVsmB670.

Scope of this Thesis:

Chapter two describes the use of Likelihood Mapping and Quartet Puzzling to unveil the family tree of SIVs isolated from different species of African monkeys. The evolution of viral genes having different roles in the asymptomatic infection of African feral nonhuman primates is analysed. The next chapters three to six describe the SIV *env* evolution during serial *in vivo* passage of SIVsmB670 in rhesus macaques. Chapter three describes the amino acid changes in the SIV *env* gene related to increased virulence and changes (or the the absence thereof) in coreceptor usage. Chapter Four describes the shortening of the asymptomatic period during serial population passage and how it is related to changes in the variable versus the constant *env* regions. The phylogenetic methods used to determine the course of natural SIV infections were subsequently applied to the SIV data set in Chapter Five. Chapter Six compares the viral load data and *env* evolution during serial passage and the relationship between viral load and genetic diversity. Finally in chapter seven the relationship between SIV virulence and *env* variation is discussed.

References:

1. **Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and E. A. Berger.** 1996. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. Science **272**:1955-1958.

2. **Allan, J. S., M. Short, M. E. Taylor, S. Su, V. M. Hirsch, P. R. Johnson, G. M. Shaw, and B. H. Hahn.** 1991. Species-specific diversity among simian immunodeficiency viruses from African green monkeys. Journal of Virology **65**:2816-2828.

3. **Almond, N., A. Jenkins, A. Slade, A. Heath, M. Cranage, and P. Kitchin.** 1992. Population sequence analysis of a simian immunodeficiency virus (32H reisolate of SIVmac251): a virus stock used for international vaccine studies. AIDS Research & Human Retroviruses **8**:77-88.

4. **Amedee, A. M., N. Lacour, J. L. Gierman, L. N. Martin, J. E. Clements, R. Bohm,Jr., R. M. Harrison, and M. Murphey-Corb.** 1995. Genotypic selection of simian immunodeficiency virus in macaque infants infected transplacentally. Journal of Virology **69**:7982-7990.

5. **Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science **220**:868-871.

6. **Baskin, G. B., L. N. Martin, S. R. Rangan, B. J. Gormus, M. Murphey-Corb, R. H. Wolf, and K. F. Soike.** 1986. Transmissible lymphoma and simian acquired immunodeficiency syndrome in rhesus monkeys. Journal of the National Cancer Institute **77**:127-139.

7. **Beer, B., E. Bailes, R. Goeken, G. Dapolito, C. Coulibaly, S. G. Norley, R. Kurth, J. P. Gautier, A. Gautier-Hion, D. Vallet, P. M. Sharp, and V. M. Hirsch.** 1999. Simian Immunodeficiency Virus (SIV) from Sun-Tailed Monkeys (Cercopithecus solatus): Evidence for Host-Dependent Evolution of SIV within the C. lhoesti Superspecies. Journal of Virology **73**:7734-7744.

8. **Beer, B., M. Peeters, S. G. Norley, G. van der Groen, L. K. Fransen, H. Willkommen, and R. Kurth.** 1995. Failure to infect lower primate species with SIVcpz. AIDS **9**:527-528.

9. **Bleul, C. C., H. Frazan, H. Choe, C. Parolin, I. Clark-Lewis, J. G. Sodroski, and T. A. Springer.** 1996. The Lymphocyte Chemoattractant SDF-1 is a Ligand for LESTR/fusin and Blocks HIV-1 Entry. Nature **382**:829-833.

10. **Bosch, M. L., A. Schmidt, M. B. Agy, L. E. Kimball, and W. R. Morton.** 1997. Infection of Macaca nemestrina neonates with HIV-1 via different routes of inoculation. AIDS **11**:1555-1563.

11. **Burns, D. P. and R. C. Desrosiers.** 1994. Envelope sequence variation, neutralizing antibodies, and primate lentivirus persistence. Current Topics in Microbiology & Immunology **188**:185-219.

12. **Burns, D. P. W., C. Collignon, and R. C. Desrosiers.** 1993. Simian immunodeficiency virus mutants resistant to serum neutralization arise during persistent infection of rhesus monkeys. Journal of Virology **67**:4104-4113.

13. **CDC.** 1981. Pneumocystis pneumonia - Los Angeles. M. M. W. R **30**:250-252.

14. **Chackerian, B., E. M. Long, P. A. Luciw, and J. Overbaugh.** 1997. Human immunodeficiency virus type 1 coreceptors participate in postentry stages in the virus replication cycle and function in simian immunodeficiency virus infection. Journal of Virology **71**:3932-3939.

15. **Chen, Z., A. Gettie, D. D. Ho, and P. A. Marx**. 1998. Primary SIVsm isolates use the CCR5 coreceptor from sooty mangabeys naturally infected in west Africa: a comparison of coreceptor usage of primary SIVsm, HIV-2, and SIVmac. Virology **246**:113-124.
16. **Chen, Z., P. Telfer, A. Gettie, P. Reed, L. Zhang, D. D. Ho, and P. A. Marx**. 1996. Genetic Characterization of New West African Simian Immunodeficiency Virus SIVsm: Geographic Clustering of Household-Derived SIV Strains with Human Immunodeficiency Virus Type 2 Subtypes and Genetically Diverse Viruses from a Single Feral Sooty Mangabey Troop. Journal of Virology **70**:3617-3627.
17. **Chen, Z., P. Zhou, D. D. Ho, N. R. Landau, and P. A. Marx**. 1997. Genetically Divergent Strains of Simian Immunodeficiency Virus Use CCR5 as a Coreceptor for Entry. Journal of Virology **71**:2705-2714.
18. **Clavel, F., D. Guetard, F. Brun-Vezinet, S. Chamaret, M. A. Rey, M. O. Santos-Ferreira, A. G. Laurent, C. Dauguet, C. Katlama, C. Rouzioux, D. Klatzmann, J. L. Champalimaud, and L. Montagnier**. 1986. Isolation of a new human retrovirus from West African patients with AIDS. Science **233**:343-346.
19. **Cocchi, F., A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, and P. Lusso**. 1995. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. Science **270**:1811-1815.
20. **Connor, R. I., K. E. Sheridan, D. Ceradini, S. Choe, and N. R. Landau**. 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1-Infected Individuals. Journal of Experimental Medicine **185**:621-628.
21. **Daniel, M. D., N. L. Letvin, N. W. King, M. Kannagi, P. K. Sehgal, R. D. Hunt, P. J. Kanki, M. Essex, and R. C. Desrosiers**. 1985. Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. Science **228**:1201-1204.
22. **de Roda Husman, A. M., R. P. van Rij, H. Blaak, S. Broersen, and H. Schuitemaker**. 1999. Adaptation to Promiscuous Usage of Chemokine Receptors Is Not a Prerequisite for Human Immunodeficiency Virus Type 1 Disease Progression. J. Infect. Dis. **180**:1106-1115.
23. **Dewhurst, S., J. E. Embretson, D. C. Anderson, J. I. Mullins, and P. N. Fultz**. 1990. Sequence analysis and acute pathogenicity of molecularly cloned SIVSMM- PBj14. Nature **345**:636-640.
24. **Douglas, N. W., G. H. Munro, and R. S. Daniels**. 1997. HIV/SIV glycoproteins: structure-function relationships. Journal of Molecular Biology **273**:122-149.
25. **Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton**. 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature **381**:667-673.
26. **Edinger, A. L., A. M. Amedee, K. Miller, B. J. Doranz, M. Endres, M. Sharron, M. Samson, Z. H. Lu, J. E. Clements, M. Murphey-Corb, M. Parmentier, S. C. Peiper, C. C. Broder, and R. W. Doms**. 1997. Differential utilization of CCR5 by macrophage and T cell tropic simian immunodeficiency virus strains. Proceedings of the National Academy of Sciences of the United States of America **94**:4005-4010.
27. **Edinger, A. L., J. E. Clements, and R. W. Doms**. 1999. Chemokine and orphan receptors in HIV-2 and SIV tropism and pathogenesis. Virology **260**:211-221.
28. **Edinger, A. L., T. L. Hoffman, M. Sharron, B. Lee, Y. Yi, W. Choe, D. L. Kolson, B. Mitrovic, Y. Zhou, D. Faulds, R. G. Collman, J. Hesselgesser, R. Horuk, and R. W. Doms**. 1998. An orphan seven-transmembrane domain receptor expressed widely in the brain functions as a coreceptor for human immunodeficiency virus type 1 and simian immunodeficiency virus. Journal of Virology **72**:7934-7940.
29. **Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger**. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven- transmembrane, G protein-coupled receptor. Science **272**:872-877.
30. **Fultz, P. N., H. M. McClure, D. C. Anderson, R. B. Swenson, R. Anand, and A. Srinivasan**. 1986. Isolation of a T-lymphotropic retrovirus from naturally infected sooty mangabey monkeys (Cercocebus atys). Proc. Natl. Acad. Sci. U. S. A. **83**:5286-5290.
31. **Fultz, P. N., H. M. McClure, D. C. Anderson, and W. M. Switzer**. 1989. Identification and biologic characterization of an acutely lethal variant of simian immunodeficiency virus from sooty mangabeys (SIV/SMM). AIDS Res. Hum. Retroviruses **5**:397-409.
32. **Fultz, P. N. and P. M. Zack**. 1994. Unique lentivirus--host interactions: SIVsmmPBj14 infection of macaques. Virus Research **32**:205-225.
33. **Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn**. 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature **397**:436-441.
34. **Gao, F., L. Yue, A. T. White, P. G. Pappas, J. Barchue, A. P. Hanson, B. M. Greene, P. M. Sharp, G. M. Shaw, and B. H. Hahn**. 1992. Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. Nature **358**:495-499.
35. **Gardner, M. B.** 1996. The history of simian AIDS. Journal of Medical Primatology **25**:148-157.

36. **Georges-Courbot, M. C., C. Y. Lu, M. Makuwa, P. Telfer, R. Onanga, G. Dubreuil, Z. Chen, S. M. Smith, A. Georges, F. Gao, B. H. Hahn, and P. A. Marx.** 1998. Natural infection of a household pet red-capped mangabey (Cercocebus torquatus torquatus) with a new simian immunodeficiency virus. Journal of Virology **72**:600-608.
37. **Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp.** 2000. AIDS as a zoonosis: scientific and public health implications. Science 2000. Jan. 28;287(5453):607-14. **287**:607-614.
38. **Hirsch, V. M., G. A. Dapolito, S. Goldstein, H. McClure, P. Emau, P. N. Fultz, M. Isahakia, R. Lenroot, G. Myers, and P. R. Johnson.** 1993. A distinct African lentivirus from Sykes' monkeys. Journal of Virology **67**:1517-1528.
39. **Hirsch, V. M., J. E. Martin, G. Dapolito, W. R. Elkins, W. T. London, S. Goldstein, and P. R. Johnson.** 1994. Spontaneous substitutions in the vicinity of the V3 analog affect cell tropism and pathogenicity of simian immunodeficiency virus. Journal of Virology **68**:2649-2661.
40. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson.** 1989. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature **339**:389-392.
41. **Holterman, L., H. Niphuis, P. ten Haaft, J. Goudsmit, G. B. Baskin, and J. L. Heeney.** 1999. Specific Passage of Simian Immunodeficiency Virus from End-Stage Disease Results in Accelerated Progression to AIDS in Rhesus Macaques. Journal of General Virology 3089-3097.
42. **Homma, T., P. J. Kanki, N. W. King,Jr., R. D. Hunt, M. J. O'Connell, N. L. Letvin, M. D. Daniel, R. C. Desrosiers, C. S. Yang, and M. Essex.** 1984. Lymphoma in macaques: association with virus of human T lymphotrophic family. Science **225**:716-718.
43. **Joag, S. V., M. G. Anderson, J. E. Clements, M. F. McEntee, D. P. Sharma, R. J. Adams, and O. Narayan.** 1993. Antigenic variation of molecularly cloned SIVmac239 during persistent infection in a rhesus macaque. Virology **195**:406-412.
44. **Kanki, P. J., M. F. McLane, N. W. King,Jr., N. L. Letvin, R. D. Hunt, P. Sehgal, M. D. Daniel, R. C. Desrosiers, and M. Essex.** 1985. Serologic identification and characterization of a macaque T-lymphotropic retrovirus closely related to HTLV-III. Science **228**:1199-1201.
45. **Kimata, J. T., J. J. Gosink, V. N. KewalRamani, L. M. Rudensey, D. R. Littman, and J. Overbaugh.** 1999. Coreceptor specificity of temporal variants of simian immunodeficiency virus Mne. Journal of Virology **73**:1655-1660.
46. **Kimata, J. T., L. Kuller, D. B. Anderson, P. Dailey, and J. Overbaugh.** 1999. Emerging cytopathic and antigenic simian immunodeficiency virus variants influence AIDS progression. Nature Medicine **5**:535-541.
47. **Kinsey, N. E., M. G. Anderson, T. J. Unangst, S. V. Joag, O. Narayan, M. C. Zink, and J. E. Clements.** 1996. Antigenic variation of SIV: mutations in V4 alter the neutralization profile. Virology **221**:14-21.
48. **Kirchhoff, F. and R. C. Desrosiers.** 1993. A PCR-derived library of random point mutations within the V3 region of simian immunodeficiency virus. PCR Methods & Applications **2**:301-304.
49. **Kirchhoff, F., H. G. Morrison, and R. C. Desrosiers.** 1995. Identification of V3 mutations that can compensate for inactivating mutations in C4 of simian immunodeficiency virus. Virology **213**:179-189.
50. **Kirchoff, F., K. Mori, and R. C. Desrosiers.** 1994. The "V3" Domain Is a Determinant of Simian Immunodeficiency Virus Cell Tropism. Journal of Virology **68**:3682-3692.
51. **Koot, M., I. P. Keet, A. H. V. Vos, R. E. Y. de Goede, M. Roos, R. A. Coutinho, F. Miedema, P. T. A. Schellekens, and M. Tersmette.** 1993. Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS. Ann. Intern. Med. **118**:681-688.
52. **Lane, T. E., M. J. Buchmeier, D. D. Watry, D. B. Jakubowski, and H. S. Fox.** 1995. Serial passage of microglial SIV results in selection of homogeneous env quasispecies in the brain. Virology **212**:458-465.
53. **Letvin, N. L., M. D. Daniel, P. K. Sehgal, L. V. Chalifoux, N. W. King, R. D. Hunt, W. R. Aldrich, K. Holley, D. K. Schmidt, and R. C. Desrosiers.** 1984. Experimental infection of rhesus monkeys with type D retrovirus. Journal of Virology **52**:683-686.
54. **Letvin, N. L., M. D. Daniel, P. K. Sehgal, R. C. Desrosiers, R. D. Hunt, L. M. Waldron, J. J. MacKey, D. K. Schmidt, L. V. Chalifoux, and N. W. King.** 1985. Induction of AIDS-like disease in macaque monkeys with T-cell tropic retrovirus STLV-III. Science **230**:71-73.
55. **Letvin, N. L. and N. W. King.** 1984. Clinical and pathologic features of an acquired immune deficiency syndrome (AIDS) in macaque monkeys. Advances in Veterinary Science & Comparative Medicine **28:237-65**:237-265.
56. **Marthas, M. L., R. A. Ramos, B. L. Lohman, K. K. Van Rompay, R. E. Unger, C. J. Miller, B. Banapour, N. C. Pedersen, and P. A. Luciw.** 1993. Viral determinants of simian immunodeficiency virus (SIV) virulence in rhesus macaques assessed by using attenuated and pathogenic molecular clones of SIVmac. Journal of Virology **67**:6047-6055.
57. **Martin Amedee, A., N. Lacour, L. N. Martin, J. E. Clements, R. B. Bohm,Jr., B. Davison, R. Harrison,**

and M. Murphey-Corb. 1996. Genotypic analysis of infant macaques infected transplacentally and orally. Journal of Medical Primatology **25**:225-235.
58. **Morrison, H. G., F. Kirchhoff, and R. C. Desrosiers**. 1993. Evidence for the cooperation of gp120 amino acids 322 and 448 in SIVmac entry. Virology **195**:167-174.
59. **Morrison, H. G., F. Kirchhoff, and R. C. Desrosiers**. 1995. Effects of mutations in constant regions 3 and 4 of envelope of simian immunodeficiency virus. Virology **210**:448-455.
60. **Murphey-Corb, M., L. N. Martin, S. R. Rangan, G. B. Baskin, B. J. Gormus, R. H. Wolf, W. A. Andes, M. West, and R. C. Montelaro**. 1986. Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys. Nature **321**:435-437.
61. **Myers, G., K. MacInnes, and B. Korber**. 1992. The emergence of simian/human immunodeficiency viruses. AIDS Research & Human Retroviruses **8**:373-386.
62. **Neildez, O., R. Le Grand, P. Caufour, B. Vaslin, A. Cheret, F. Matheux, F. Theodoro, P. Roques, and D. Dormont**. 1998. Selective quasispecies transmission after systemic or mucosal exposure of macaques to simian immunodeficiency virus. Virology **243**:12-20.
63. **Nowak, M. A., R. M. Anderson, A. R. McLean, T. F. Wolfs, J. Goudsmit, and R. M. May**. 1991. Antigenic diversity thresholds and the development of AIDS. Science **254**:963-969.
64. **Overbaugh, J. and L. M. Rudensey**. 1992. Alterations in potential sites for glycosylation predominate during evolution of the simian immunodeficiency virus envelope gene in macaques. Journal of Virology **66**:5937-5948.
65. **Overbaugh, J., L. M. Rudensey, M. D. Papenhausen, R. E. Benveniste, and W. R. Morton**. 1993. Variation in simian immunodeficiency virus env is confined to V1 and V4 during progression to simian AIDS. Journal of Virology **65**:7025-7031.
66. **Owen, S. M., D. Ellenberger, M. Rayfield, S. Wiktor, P. Michel, M. H. Grieco, F. Gao, B. H. Hahn, and R. B. Lal**. 1998. Genetically divergent strains of human immunodeficiency virus type 2 use multiple coreceptors for viral entry. Journal of Virology **72**:5425-5432.
67. **Schwiebert, R. and P. N. Fultz**. 1994. Immune activation and viral burden in acute disease induced by simian immunodeficiency virus SIVsmmPBj14: correlation between in vitro and in vivo events. Journal of Virology **68**:5538-5547.
68. **ten Haaft, P., B. Verstrepen, K. Uberla, B. Rosenwirth, and J. Heeney**. 1998. A pathogenic threshold of virus load defined in simian immunodeficiency virus- or simian-human immunodeficiency virus- infected macaques. Journal of Virology **72**:10281-10285.
69. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins**. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research **25**:4876-4882.
70. **Valli, P. J. S., V. V. Lukashov, and J. Goudsmit**. 1998. Shortening of the symptom-free period in rhesus macaques related to decreasing non-synymous variation in the env variable regions of SIVsm during passage. Journal of Virology **72**:7494-7500.
71. **Watanabe, N., S. N. McAdam, J. E. Boyson, M. S. Piekarczyk, Y. Yasutomi, D. I. Watkins, and N. L. Letvin**. 1994. A simian immunodeficiency virus envelope V3 cytotoxic T-lymphocyte epitope in rhesus monkeys and its restricting major histocompatibility complex class I molecule Mamu-A*02. Journal of Virology **68**:6690-6696.

Chapter 2

Phylogenetic Analysis using Likelihood Mapping and Quartet Puzzling of the Common Ancestry of the SIVsm/HIV-2 and SIVcpz/HIV-1 Clusters

Phylogenetic Analysis using Likelihood Mapping and Quartet Puzzling of the Common Ancestry of the SIVsm/HIV-2 and SIVcpz/HIV-1 Clusters

P.J. Spencer Valli, Antoinette van der Kuyl, and Jaap Goudsmit

Department of Human Retrovirology, Academic Medical Centre, Amsterdam,
The Netherlands.

Department of Human Retrovirology
Academic Medical Centre
University of Amsterdam
Meibergdreef 15, 1105 AZ Amsterdam
The Netherlands
telephone: (31-20) 566-4853
fax: (31-20) 6916531

Abstract: Cross species transmission of a lentivirus from sooty mangabeys to humans gave rise to HIV-2 and similarly suggestive evidence has been presented that transmission from chimpanzees to humans gave rise to HIV-1. In order to identify the putative common ancestor of SIVsm, SIVcpz, and HIV-1/HIV-2 virus family we applied likelihood mapping and quartet puzzling. The *gag* and *env* of SIV isolates from eight species of monkey as well as the HIV-1 group M, N, and O viruses were analysed. Likelihood mapping showed that phylogenetic reconstruction of the SIV/HIV cluster was possible because the sequence data had a tree-like structure. The SIV isolates of *Cercopithecus l'hoesti l'hoesti* and *Cercopithecus l'hoesti solatus* were consistently placed in a separate cluster and possibly are genetically at the root of the primate lentivirus phylogenetic tree. After the SIV variants of mandrills diverged there were three separate clusters observed (when judged on the basis of the *env*): SIVsm/HIV-2, SIVcpz/HIV-1 M, N, O, and SIVsyk/SIVagm. A somewhat different branching pattern was seen with *gag*: SIVcpz/HIV-1, and SIVsab at the root of the SIVsyk/SIVagm, and the SIVsm/HIV-2 clusters.

Introduction: Eight of the ninety one species of African primates have been found to naturally harbour lentiviruses that have been partially or totally characterized (1). SIVsm from sooty mangabeys (*Cercocebus torquatus atys*), SIVmnd from the mandrill (*Mandrillus sphinx*), SIVsykes from sykes monkeys (*Cercopithecus mitis*), SIVcpz from chimpanzees (*Pan troglodytes*), SIVl'hoest from L'hoest monkeys (*Cercopithecus l'hoesti l'hoesti*), SIVsun from sun-tailed monkeys (*Cercopithecus l'hoesti solatus*), SIVrcm from red-capped mangabeys (*Cercocebus torquatus torquatus*, which appears to be a complex recombinant of SIVcpz and SIVagm (6)), and SIVagm from the four subspecies of the African green monkey (*Cercopithecus aethiops*). The SIVrcm, as well as SIVdrill, SIVwcm, and SIVtal are partially characterised (from drill, white-crowned mangabey, and talapoin monkeys (6,8,26,33)). Cross-species transmission of virus in wild monkeys has also been documented in several, sympatric species with the SIVagm from African green monkeys infecting a yellow chacma baboon, and a patas monkey (4,17,33,35). Of these species infected in the wild or in captivity no pathology resulting from these infections has been observed, except for the death of the patas monkey infected with SIVwcm which suffered from an AIDS-like syndrome (4). The newly found SIVrcm *gag* clusters with the SIVagm clade, and the *pol* gene clusters with SIVcpz (12).

The numbers of infected feral African monkeys varies greatly by species with Talapoins being tested at 70% seropositive, Sooty mangabeys at 57%, African green monkeys at 36%, and Sykes monkeys at 34% being the most seroprevalent species (Levy, The Retroviridae 4, and (11,15,18)). Other SIV infected species are found to be far less seropositive with Mandrills at 15%, Colobus at 10%, with Chimpanzees, and Baboons (other than mandrills) both having less than 1% seropositivity. There have been no Orangutans or Gorillas who have tested positively (1,3).

Transmission of SIVsm to Asian primates results in AIDS resembling that seen in humans infected with HIV. When pathogenicity is taken as evidence for recent viral introduction these data suggest the emergence of the virus after the separation of the Asian, and African primates. In humans AIDS is caused by two strains of the virus with HIV-1 and HIV-2 closely related to SIVcpz (11), and SIVsm (6) respectively. So far, of the eight SIV strains known, only four (SIVsm, SIVl'hoest, SIVsun, and SIVcpz) can infect human cells (1,6,15,23). The

inability of some SIV species specific strains to infect cells of even other non-human primates, than the species which the virus was isolated, as in SIVsykes (16) makes them, aside from the phylogenetic data, unlikely to have been the source of the human epidemic.

We used likelihood-mapping, which determines the best arrangement of the smallest information-containing subset of a tree (the quartet, or four taxa (30), to examine the branching pattern of the family of SIV and HIV strains. Our application of this method to the HIV/SIV alignments showed them to be highly informative compared to some HIV alignments, and the reconstruction of the phylogeny was tree-like (34). The analysis of the variation in the nonsynonymous substitution between the *env* and *gag* alignments showed a greater adaptive pressure upon the *env*.

Materials and Methods:

Sequences: The sequence accession numbers for the sequences used are: SIVl'hoest AF075269, SIVmnd M27470, SIVsun AF131870, SIVrcm AF1028, SIVcpzUS *gag* AF103818, HIV1,2 and SIV other *gag* and *env* sequences are available at http://hiv-web.lanl.gov/MAP/ hivmap.html. No envelope sequences for SIVrcm are presently available.

Phylogenetic Analysis: *Env*, *gag*, and mitochondrial 12S rRNA (Old World primate species) gene sequences were formatted into the Clustal format using Omiga 1.1.3 (Oxford Molecular Limited). Clones were aligned using ClustalX (32) and further optimised manually. Phylogenetic analysis was conducted using MEGA (20), ClustalX (32), and PUZZLE 4.02 (30,31). Neighbour-joining, maximum-likelihood and puzzle-based trees were produced using the Tamura-Nei distance estimation with pair-wise comparison and then boot-strapped with 1000 replications before tree construction with and without preselected outgroups. The maximum-likelihood mapping and quartet puzzling (QP) was carried out using Tamura-Nei distance estimation. Likelihood mapping was used to visualize the phylogenetic content of the sequence alignments (31). Briefly, likelihood-mapping is the construction of all quartets of sequences possible (e.g. A,B,C, and D, and then A,B,C, and E, etc. until all of the possible quartets of sequences in the analysis have been compared) and calculation of the relative weights of the probabilities for all of the three resolved tree structures (#1: A is like B, but not C or D, #2: A is like C, but not like B or D, and #3: A is like D, but not like B or C (29-31)) and the four unresolved structures (partially resolved #1: A is like B or C, but not D, partially resolved #2: A is like B or D, but not C, partially resolved #3: A is like C or D, but not D, and unresolved: A is not like B, C, or D). Thus, for the first four sequences analysed using likelihood-mapping, the seven possible locations of the probability vectors include three with a completely resolved placement of the four taxa, three with one taxon split between attraction to two others and not alike the fourth taxon (partially resolved), and one of star-like form or no favoured relation between the four taxa (unresolved). These seven relations can be plotted in a triangle to show the phylogenetic signal. The corners represent areas of completely resolved (tree-like) quartets (A is most like one of B,C or D, resolved #1,2, and 3); the sides represent quartets with a split between equal attraction of one taxon for two other taxa (partially resolved: A is like B and C, B and D, or C and D); and the centre of the triangle represents an area of equal likelihood or distance between the four taxa (A is not like B,C, or D, unresolved, giving a star-like formation). If each possible quartet vector (resolved, partially resolved and unresolved quartet comparison) is plotted as a point in the triangle it gives

a rapid visualisation of the phylogenetic content of the alignment data (robustness of the tree, or lack thereof).

Tree construction was carried out using the maximum-likelihood based PUZZLE 4.02 (31). The maximum-likelihood analysis is used for determination of a tree and its corresponding branch lengths that have the greatest likelihood of reconstructing the correct phylogeny. The numbers of possible topologies increases exponentially as the number of taxa rises, making heuristic searches very slow. Reconstructing a tree using all of the possible sets of four taxa, or quartets, allows the heuristic search to proceed efficiently in a ML procedure (30). Quartet puzzling reconstructs the ML tree for all possible quartets. The total set of quartet trees are subsequently combined to form a complete tree in the quartet puzzling step (QP). This procedure was carried out 1000 times. The (QP) tree is a majority-rule consensus (22) of the set number of puzzling steps. There are three main steps to the QP method, ML determination of all possible quartet trees, combining all of the quartet trees one thousand times to form a complete tree (puzzling step), and the final majority rule consensus computation of all intermediate trees (31). For the analysis of the progenitor sequences only the relationship of these groups are of interest and not details of the structure of the clusters themselves. Then a likelihood mapping analysis is sufficient. The clusters selected were SIVl'hoest and SIVsun as A and the rest of the aligned sequences (both *env* and *gag*) selected as B. The corresponding likelihood mapping triangle diagrams show the possible relationships in detail (19,31), and show the robustness of the internal edge of the tree. The analysis of the synonymous (ds) and nonsynonymous (dn) nucleotide substitutions was conducted using the automated SNAP system (found at http://hiv-web.lanl.gov (10,25)) using the Nei and Gojobori method. Alignments are from the Los Alamos web site.

The resolution of complex phylogenies is helped by using longer sequences, such as we have done here, to avoid trees containing little information and having star topologies. The use of quartet analysis, with its focus on small informative subtrees (29-31), has been applied to an inference method that takes advantage of the determination of all possible quartets, or heuristic search, unavailable to the other methods, to time consuming or available only as a statistical test of best fit (21). PUZZLE 4.0.2 incorporates a first-step plot of all the possible four taxa quartets as a means of weighting the best placement of them all within a phylogeny. With seven possible organizations for each quartet, calculation of all possibilities is feasible. The maximum likelihood of the grouping of a quartet can also be plotted graphically to give an idea of the "tree likeness" of the data known as likelihood mapping as shown in Figure 1A, B, and C. The four taxa are compared by grouping three and comparing the fourth to them by placing it in one of the seven possible maximum-likelihood areas as a dot: with taxon D most like taxa A,B or C (the three corners: tree-like), split between A and B or A and C or B and C (the three sides: unresolved between the two), and lastly, unlike A,B or C and placed in the middle as a neutral attractor (star formation). The centre of the maximum-likelihood triangles are inhabited by star-like phylogenies lacking tree-like content, i.e., having no attraction to cluster with the other taxa (31,34). The percentages of the contents of the complete set of quartets are displayed in the numbered triangles. Each of the seven areas has a finite number of the maximum likelihood dots, with the total of the seven being one hundred percent.

Results:

Likelihood Mapping: The phylogenies were shown to all contain large amounts of tree

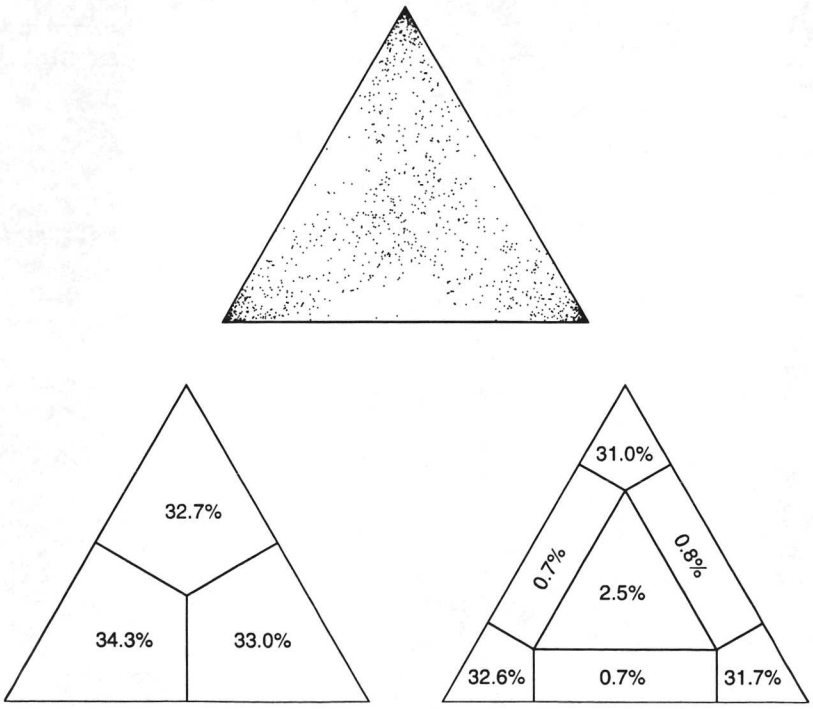
topology information. The *env*, *gag* and 12s alignments carried 95.3, 98.2, and 87.1 % respectively of resolved quartets (Figure 1A, B, 12s not shown). The top triangle shows the distribution pattern of the data in the seven basins of attraction or quartet possibilities. The three corners of all three likelihood-mapping analysis, where completely resolved trees are found, contain virtually all of the content. The analysis of the internal edge containing the SIVl'hoest and SIVsun (Figure 1C, quartet diagram) was used to ascertain the robustness of the separation of these two sequences as a cluster (of progenitor sequences), as opposed to a cluster containing all the remaining sequences. This algorithm analyses exhaustively all of the possible quartets of sequences. Two from one side of the branch (designated (a) and (a) for there is little separation between the two sequences) and two from the other side of the internal branch ((b) and (b)), chosen at random to encompass all of the possible subsets of the alignment taxa. This compares the branch the SIVl'hoest, and SIVsun are placed on against all of the combinations of all of the other sequences to analyse the probability that they are not the root. This allows the analysis of the possibility of the incorrect rooting of the trees, and in this case of the incorrect selection of possible progenitor sequences (Figure 1C). This was carried out for the *env* as well as the *gag* genes where the results were exactly the same (*env* analysis, and quartet designation are shown). The robustness of the branching separation of the two clusters is seen as the absence of any shared quartets which would appear in the centre of the triangles. The taxa were set in two distinct clusters, with 100% of the phylogenetic information confirming the separation of these subsets of the sequences as evolutionarily distinct.

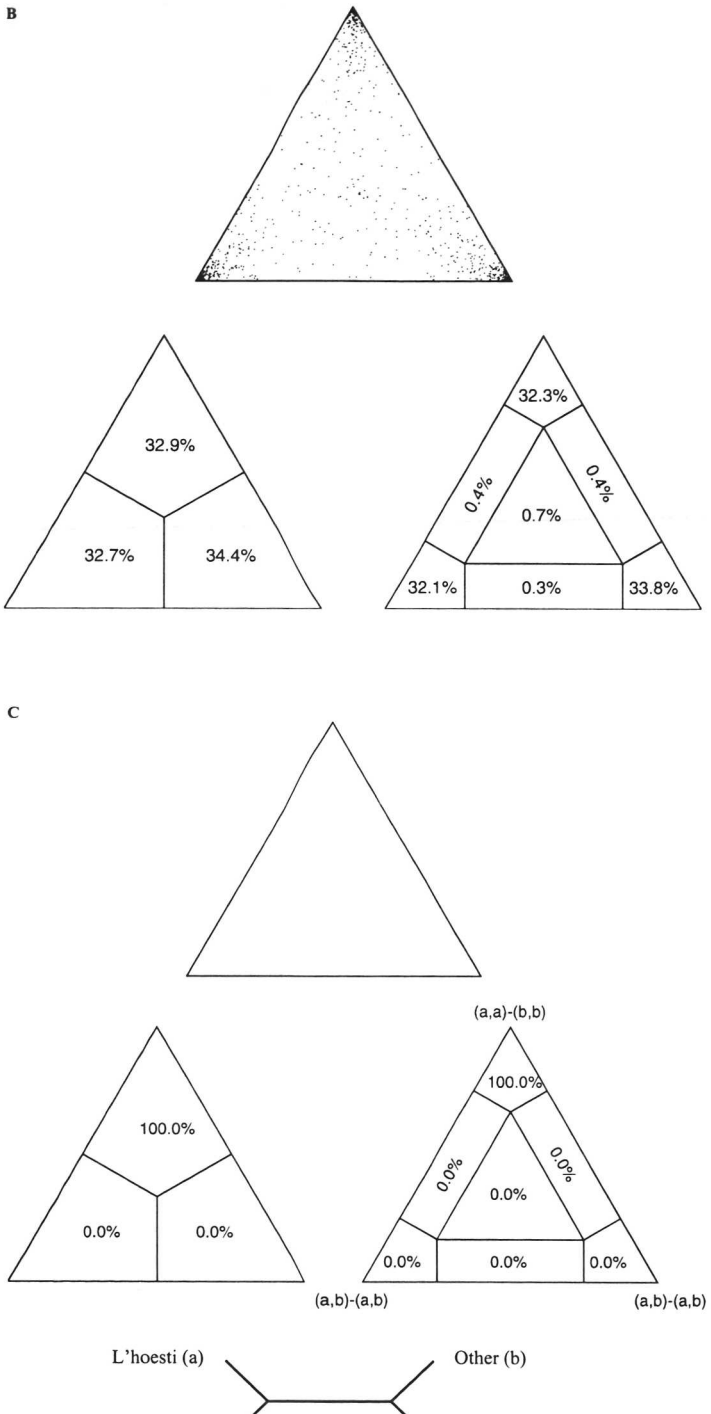
Phylogeny construction:

The gp160 *env* quartet puzzling tree (Figure 2A) shows the divergence of the SIVl'hoest and Sun from the Mnd as a branching pattern also found in the *gag* tree (Figure 2B), and supported by the likelihood-mapping. The *env* of the viruses has evolved further than the *gag* as is seen in the trichotomy of the Agm, Cpz, Sm families, plus the Sykes sequence alone, whereas in the case of the *gag* phylogeny the Agm and Sm sequences cluster together on a branch opposite the Cpz cluster. The tantalus *env* clusters with the other Agm substrains but has a very long branch, whereas the *gag* is short and could possibly be indicative of an earlier recombinant such as that seen in the genome of the YBF30 N strain (Figure 2A, and 2B (11)). The recombinant nature of YBF30 N is known, but has not been analysed previously using two complete different alignments of genes. The clustering of the *env* with the CpzUS, and the *gag* on a sublying branch of CpzUS and CpzGab is in accordance with the analysis done with half genome concatomers of protein sequences (11). The new CpzUS sequence does not vary markedly from the CpzGab in the *env* phylogeny and in the *gag* it clusters further away with the Ant70 sequence. The branching of the three (Agm, Sm, and Cpz) strains varies between the *env* with three distinct branches and the *gag* with two branches and the Agm and Sm splitting further on the branch together with the Sabaeus as outlier. The Sabaeus appears to cluster differently in the *env* and *gag* genes as well. The 12s mtRNA followed a different radiation of the species (not shown). Human 12s sequences cluster with the gorilla sequence and not with either species of chimpanzee as in the case of HIV-1. As the 12s evolution is not involved in the immune selection of the virus-host evolution and is inherited in a maternal fashion it is not surprising that the branching pattern varies.

Figure 1: Likelihood-mapping analysis for the *env* gp160 (A), and *gag* (B) alignments. The top triangle shows the distribution pattern of the quartet analysis, and the bottom triangles show the distribution of the data in the seven basins of attraction or quartet possibilities. The majority of the data clusters in the three corners, where completely resolved robust trees with large amounts of phylogenetic content are found. The likelihood-mapping analysis can also be applied to the testing of an internal edge of a tree in divided into two clusters. Two-cluster likelihood-mapping analysis (C) of the *env* gp160 alignment split into two disjoint groups: (a,a)=(SIVl'hoest and SIVsun) and (b,b)=(all possible combinations of a duo of the remaining sequences). Each quartet's likelihood is plotted with the three possible arrangements plotted in a triangle (C upper panel). The tips of the triangle display the support for one of the three arrangements as labelled in the lower right triangle, and the quartet diagram. The percentage of the quartets supporting the grouping of the SIVl'hoest with the SIVsun with the exclusion of all other combinations of two sequences is found in the upper corner of the three areas ((a,a)-(b,b)). The clusters chosen separate the possible progenitor sequences (SIVl'hoest and Sun), from the rest of the HIV/SIV sequences analyzed (C quartet diagram).

A




B
32.9%
32.7%
34.4%
32.3%
0.4%
0.4%
0.7%
32.1%
0.3%
33.8%
C
100.0%
0.0%
0.0%
(a,a)-(b,b)
100.0%
0.0%
0.0%
0.0%
0.0%
0.0%
0.0%
(a,b)-(a,b)
(a,b)-(a,b)
L'hoesti (a)
Other (b)
Sun (a)
Other (b)

The differences in the evolutionary patterns in the *env*, and *gag* was studied to illustrate the marked differences in evolutionary pressure between pathogenic infection in man and the ancient nonpathogenic infection in nonhuman primates. The ds/dn ratios (Table 1) clearly show the four fold rate of variation of HIV-1 *env* substitutions as compared to that of SIV *env*. The ratios of *gag* and *nef* are two fold higher in HIV-1 and the ratio for *pol* is less than one and a half.

Table 1: Mean synonymous nucleotide substitutions (ds), nonsynonymous nt substitutions (dn) shown as ds/dn. Estimated by the method of Gojobori and Nei (25) as implemented in SNAP (10) analysis software. The estimation was carried out on alignments of the *env* gp160 or complete *gag*, *nef*, or *pol*.

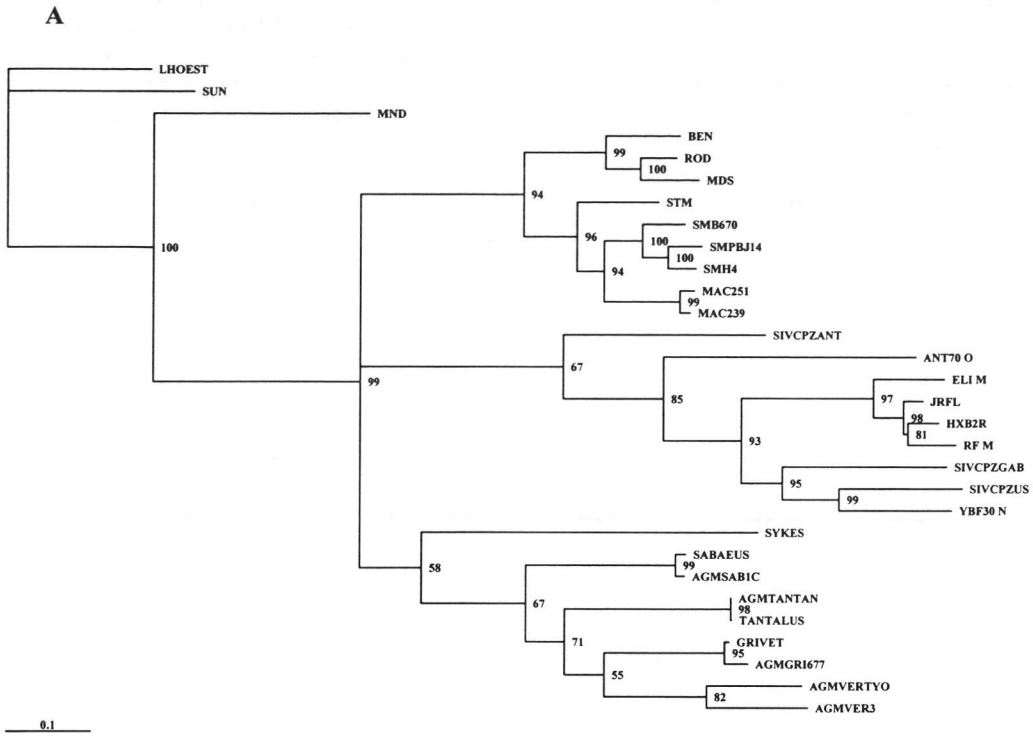
Ratios of Genetic Variation as Measured by Synonymous/Nonsynonymous Substitutions (ds/dn)

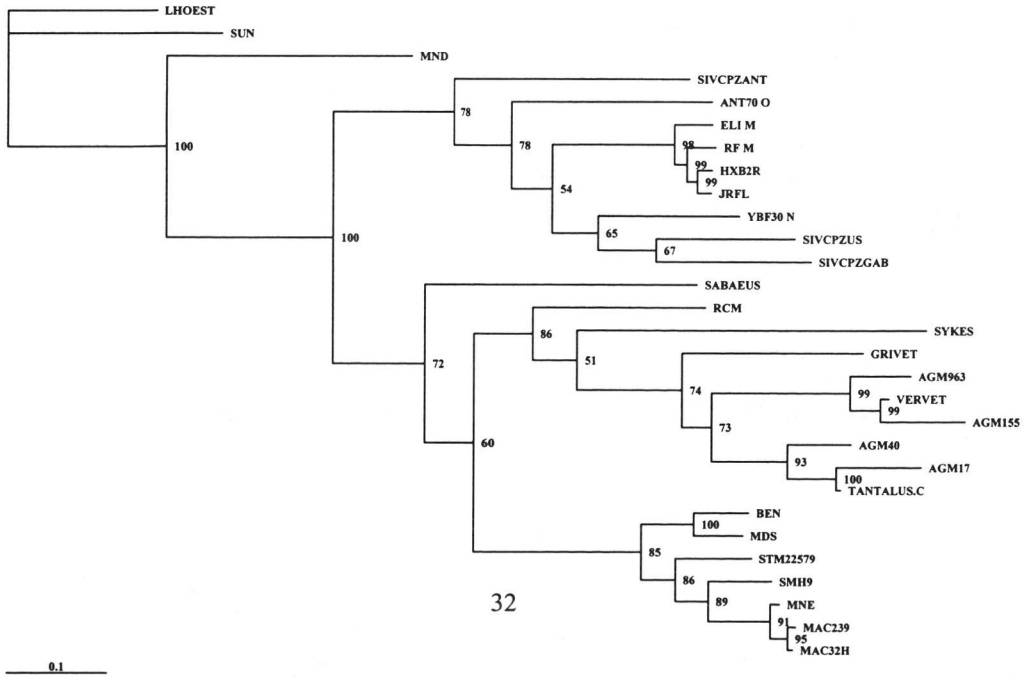
	env	*gag*	*nef*	*pol*
HIV-1	1.93	4.97	3.5	6.21
SIV	8.85	9.6	7.23	8.23

Discussion: A true primate lentivirus phylogeny is very difficult to reconstruct as ancient as well as recent cross-species transmissions have occurred, not to mention recombination events and missing strains. The branching of three (Agm, Sm, and Cpz) strains varies between the *env* (with three distinct branches) and the *gag* (with two branches and the Agm and Sm splitting further on).

Without causing any pathology SIVsm in its natural host replicates to high levels, 10^5 to 10^7 copies per millilitre plasma (27). This implies rapid replication kinetics and viremia. The branches of the SIVsm sequences (Figure 1A) are comparable to those of HIV-2 and implies a relatively equal rate of evolution (as compared to the longer branches of the SIVcpz and HIV-1 sequences). The branch lengths of the pathogenic SIVsm related viruses used in experimental infections (SIVsmB670, smH4, Mac239) do not produce longer branches than HIV-2 in the *env* or *gag*. The mutation rate of the *env* of SIVsm in Sooty mangabeys, SIVagm in African green monkeys, and SIVsm (B670, SMH4, and Mac239) strains in Rhesus macaques has been studied, and the rates are similar (10^{-2} to 10^{-3} substitutions per site per year (5,9,24,27)).

Figure 2: Phylogenetic inference for the *env* gp160 (A),and *gag* (B). Trees were constructed using Puzzle 4.0.2 with the Tamura and Nei distance approximation and 1000 tree puzzling steps. The quartet puzzling percentages of the important branches are in bold numbers. The scale bar is found in the lower left hand corner.


A
LHOEST
SUN
MND
BEN
ROD
MDS
STM
SMB670
SMPBJ14
SMH4
MAC251
MAC239
SIVCPZANT
ANT70 O
ELI M
JRFL
HXB2R
RF M
SIVCPZGAB
SIVCPZUS
YBF30 N
SYKES
SABAEUS
AGMSABIC
AGMTANTAN
TANTALUS
GRIVET
AGMGRI677
AGMVERTYO
AGMVER3
0.1


B
LHOEST
SUN
MND
SIVCPZANT
ANT70 O
ELI M
RF M
HXB2R
JRFL
YBF30 N
SIVCPZUS
SIVCPZGAB
SABAEUS
RCM
SYKES
GRIVET
AGM963
VERVET
AGM155
AGM40
AGM17
TANTALUS.C
BEN
MDS
STM22579
SMH9
MNE
MAC239
MAC32H
0.1

The SIVagm replicates to the lowest levels, and thus would presumably be less likely to rapidly adapt (2,14). The SIVagm cluster the furthest from the root *gag* branches and the branch lengths of the *gag* cluster are roughly equal to those of the SIVagm *env* cluster (Figure 2A and B). This dichotomy as compared to the SIVsm/HIV-2, and SIVcpz/HIV-1 clusters, where the *env* is more variable than the *gag* (presumably caused by the immune systems pressure) has been attributed to the lack of immune response (7). Studies comparing *env* and *gag* sequences derived sequentially from the same SIVagm infected African green monkey would be needed to resolve the branching pattern. Based on natural history data and our current analysis it appears that mutation rate and virus load are not the only predictors of virulence.

The large territory inhabited in Africa and the overlap with the other species would increase the chances of cross species transmission. In the case of SIVsykes the virus requires it's own host cells to infect and replicate (very host specific), which dramatically reduces the chances of finding these infections in other species. SIVsykes, a virus of a cercopithecus, was isolated from a hybrid monkey which was found to be a cross of mitis and ascaneus, and thus may explain it's odd host specific replication requirements. It is evident how the Cercopithecus L'hoest, Sun and Mnd would be subject to cross species transmission of SIV as they inhabit small regions which partially overlap. The geographic overlap may have been greater in the past but the habitat left is very small and may put the monkeys in close contact. This makes the relatedness of the three, and the stable changes between them in *env* and *gag* an example of a probably complete history. The long branches to the other strains suggesting missing isolates, or a case of very quick episodic evolution between the SIVl'hoest, SIVsun, SIVmnd and the rest of the known strains.

The HIV-1 YBF30 N isolate has been shown to be a recombinant of *env* and *gag-pol* or 3' and 5' to SIVcpzGAB (28) or to SIVcpzUS respectively (11,13,36). We analysed the *env* and *gag* genes separately. The branching of the *gag* outside the SIVcpzGAB and US branch had a far lower quartet puzzling probability than the clustering of the *env* (65 versus 95 respectively, Figures 2A and 2B). There may as yet be other subtypes that the Type N *gag* clusters together with having a more highly supported quartet puzzling probability, lying between the Type N and the Type M, and O.

The isolation of the SIVl'hoest and SIVsun led to the proposal due to their clustering that these SIV strains formed a new lineage evolved within the Cercopithecus l'hoesti superspecies (1). Seeming to predate the mandrill as a new lineage the distance between the SIVl'hoest, SIVsun and the SIVmnd may be due to separate origins in history and as the number of known SIV strains increases they will likely be separate related lineages, as is supported here by the branching analysis. The distances shown by the branching of the *env* and *gag* genes are almost identical and imply a common ancestral distance with no recombination between the two genes of the two separate families (SIVl'hoest and SIVsun are *Cercopithecus* and SIVmnd from a mandrill is a *Cercocebus* which is more related to sooty mangabeys (1)). The equal branch lengths of the two genes in the three viruses may imply equal evolutionary pressure upon them (1,15). This may be the fourth and fifth clusters of the SIV/HIV family but the distances between these three viruses easily fits into the species specific family of the four SIVagm strains, and may be another example of cross species transmission.

The data gained in recent years by the isolation of new strains of SIV have yielded more insight into the relation between the nonhuman primate viruses and those found in humans. The large gaps in the HIV\SIV tree result most likely from our snapshot of time sampling of the millennium old infection, and the surprisingly low rate of infection among some of the feral monkeys. The ability of only four of the known strains to infect human cells and their genetically relevant identities point to them as the putative progenitors of HIV infection in man.

Acknowledgements: I would like to thank Vanessa Hirsch and Brigitte Beer for sequences prior to publication and helpful discussions during the preparation of this manuscript.

References:

1. **Beer, B., E. Bailes, R. Goeken, G. Dapolito, C. Coulibaly, S. G. Norley, R. Kurth, J. P. Gautier, A. Gautier-Hion, D. Vallet, P. M. Sharp, and V. M. Hirsch**. 1999. Simian Immunodeficiency Virus (SIV) from Sun-Tailed Monkeys (*Cercopithecus solatus*): Evidence for Host-Dependent Evolution of SIV within the C. lhoesti Superspecies. Journal of Virology **73**:7734-7744.

2. **Beer, B., J. Scherer, J. zur Megede, S. Norley, M. Baier, and R. Kurth**. 1996. Lack of dichotomy between virus load of peripheral blood and lymph nodes during long-term simian immunodeficiency virus infection of African green monkeys. Virology **219**:367-375.

3. **Bibollet-Ruche, F., C. Brengues, A. Galat-Luong, G. Galat, X. Pourrut, N. Vidal, F. Veas, J. P. Durand, and G. Cuny**. 1997. Genetic diversity of simian immunodeficiency viruses from West African green monkeys: evidence of multiple genotypes within populations from the same geographical locale. J. Virol. **71**:307-313.

4. **Bibollet-Ruche, F., A. Galat-Luong, G. Cuny, P. Sarni-Manchado, G. Galat, J. P. Durand, X. Pourrut, and F. Veas**. 1996. Simian immunodeficiency virus infection in a patas monkey (Erythrocebus patas): evidence for cross-species transmission from African green monkeys (*Cercopithecus aethiops sabaeus*) in the wild. J. Gen. Virol. 77:773-781.

5. **Campbell, B. J. and V. M. Hirsch**. 1994. Extensive envelope heterogeneity of simian immunodeficiency virus in tissues from infected macaques. Journal of Virology **68**:3129-3137.

6. **Chen, Z., P. Telfer, A. Gettie, P. Reed, L. Zhang, D. D. Ho, and P. A. Marx**. 1996. Genetic Characterization of New West African Simian Immunodeficiency Virus SIVsm: Geographic Clustering of Household-Derived SIV Strains with Human Immunodeficiency Virus Type 2 Subtypes and Genetically Diverse Viruses from a Single Feral Sooty Mangabey Troop. Journal of Virology **70**:3617-3627.

7. **Cichutek, K. and S. Norley**. 1993. Lack of immune suppression in SIV-infected natural hosts. AIDS **7 Suppl 1:S25-35**:S25-35.

8. **Clewley, J. P., J. C. Lewis, D. W. Brown, and E. L. Gadsby**. 1998. A novel simian immunodeficiency virus (SIVdrl) pol sequence from the drill monkey, Mandrillus leucophaeus. Journal of Virology **72**:10305-10309.

9. **Fomsgaard, A., P. R. Johnson, W. T. London, and V. M. Hirsch**. 1993. Genetic variation of the SIVagm transmembrane glycoprotein in naturally and experimentally infected primates. AIDS 7:1041-1047.

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

11. **Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn**. 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature **397**:436-441.

12. **Georges-Courbot, M. C., C. Y. Lu, M. Makuwa, P. Telfer, R. Onanga, G. Dubreuil, Z. Chen, S. M. Smith, A. Georges, F. Gao, B. H. Hahn, and P. A. Marx**. 1998. Natural infection of a household pet red-capped mangabey (*Cercocebus torquatus torquatus*) with a new simian immunodeficiency virus. Journal of Virology **72**:600-608.

13. **Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp**. 2000. AIDS as a zoonosis: scientific and public health implications. Science 2000. Jan. 28;287(5453):607-14. **287**:607-614.

14. **Hartung, S., K. Boller, K. Cichutek, S. G. Norley, and R. Kurth**. 1992. Quantitation of a lentivirus in its natural host: simian immunodeficiency virus in African green monkeys. Journal of Virology **66**:2143-2149.

15. **Hirsch, V. M., B. J. Campbell, E. Bailes, R. Goeken, C. Brown, W. R. Elkins, M. Axthelm, M. Murphey-Corb, and P. M. Sharp**. 1999. Characterization of a novel simian immunodeficiency virus (SIV) from L'Hoest monkeys (*Cercopithecus l'hoesti*): implications for the origins of SIVmnd and other primate

lentiviruses. Journal of Virology 73:1036-1045.
16. **Hirsch, V. M., G. A. Dapolito, S. Goldstein, H. McClure, P. Emau, P. N. Fultz, M. Isahakia, R. Lenroot, G. Myers, and P. R. Johnson**. 1993. A distinct African lentivirus from Sykes' monkeys. Journal of Virology 67:1517-1528.
17. **Jin, M. J., J. Rogers, J. E. Phillips-Conroy, J. S. Allan, R. C. Desrosiers, G. M. Shaw, P. M. Sharp, and B. H. Hahn**. 1994. Infection of a yellow baboon with simian immunodeficiency virus from African green monkeys: evidence for cross-species transmission in the wild. J. Virol. 68:8454-8460.
18. **Kingdon, J.** 1997. The Kingdon Field Guide to African Mammals. Academic Press Limited, London.
19. **Krings, M., A. Stone, R. W. Schmitz, H. Krainitzki, M. Stoneking, and S. Paabo**. 1997. Neandertal DNA sequences and the origin of modern humans. Cell 90:19-30.
20. **Kumar, S., K. Tamura, and M. Nei**. 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Computer Applications in the Biosciences 10:189-191.
21. **Leitner, T., D. Escanilla, C. Franzev, U. Mathias, and J. Albert**. 1996. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. Proceedings of the National Academy of Sciences of the United States of America 93:10864-10869.
22. **Margush, T. and F. R. McMorris**. 1981. Consensus n-trees. Bull. Math. Biol. 239-244.
23. **Marx, P. A., Y. Li, N. W. Lerche, S. Sutjipto, A. Gettie, J. A. Yee, B. H. Brotman, A. M. Prince, A. Hanson, and R. G. Webster**. 1991. Isolation of a simian immunodeficiency virus related to human immunodeficiency virus type 2 from a west African pet sooty mangabey. Journal of Virology 65:4480-4485.
24. **Muller-Trutwin, M. C., S. Corbet, M. D. Tavares, V. M. A. Herve, E. Nerrienet, M. C. Georges-Courbot, W. Saurin, P. Sonigo, and F. Barre-Sinoussi**. 1996. The Evolutionary Rate of Nonpathogenic Simian Immunodeficiency Virus (SIVagm) Is in Agreement with a rapid and Continuous Replication in vivo. Virology 223:89-102.
25. **Nei, M. and T. Gojobori**. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418-426.
26. **Osterhaus, A. D., N. Pedersen, G. van Amerongen, M. T. Frankenhuis, M. Marthas, E. Reay, T. M. Rose, J. Pamungkas, and M. L. Bosch**. 1999. Isolation and partial characterization of a lentivirus from talapoin monkeys (*Myopithecus talapoin*). Virology 260:116-124.
27. **Rey-Cuille, M. A., J. L. Berthier, M. C. Bomsel-Demontoy, Y. Chaduc, L. Montagnier, A. G. Hovanessian, and L. A. Chakrabarti**. 1998. Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. Journal of Virology 72:3872-3886.
28. **Simon, F., P. Mauclere, P. Roques, I. Loussert-Ajaka, M. C. Muller-Trutwin, S. Saragosti, M. C. Georges-Courbot, F. Barre-Sinoussi, and F. Brun-Vezinet**. 1998. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. Nat. Med. 4:1032-1037.
29. **Strimmer, K., N. Goldman, and A. von Haeseler**. 1997. Bayesian Probabilities and Quartet Puzzling. Molecular Biology & Evolution 14:258-259.
30. **Strimmer, K. and A. von Haeseler**. 1996. Quartet Puzzling: A Quartet Maximum-Likelihood Method for Reconstructing Tree Topologies. Journal of Molecular Evolution 13:964-969.
31. **Strimmer, K. and A. von Haeseler**. 1997. Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. Proceedings of the National Academy of Sciences 94:6815-6819.
32. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins**. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25:4876-4882.
33. **Tomonaga, K., J. Katahira, M. Fukasawa, M. A. Hassan, M. Kawamura, H. Akari, T. Miura, T. Goto, M. Nakai, and M. Suleman**. 1993. Isolation and characterization of simian immunodeficiency virus from African white-crowned mangabey monkeys (*Cercocebus torquatus lunulatus*). Arch. Virol. 129:77-92.
34. **Valli, P. J. S. and J. Goudsmit**. 1998. Structured Tree Topology and Adaptive Evolution of the SIVsm Envelope During Serial Passage in Rhesus Macaques According to Likelihood Mapping and Quartet Puzzling. Journal of Virology 72:3673-3683.
35. **van Rensburg, E. J., S. Engelbrecht, J. Mwenda, J. D. Laten, B. A. Robson, T. Stander, and G. K. Chege**. 1998. Simian immunodeficiency viruses (SIVs) from eastern and southern Africa: detection of a SIVagm variant from a chacma baboon. Journal of General Virology 79:1809-1814.
36. **Weiss, R. A. and R. W. Wrangham**. 1999. From Pan to pandemic. Nature 397:385-386.

Chapter 3

Genotypic Variation in SIVsm DeltaB670 *Env* During the Shift to High Virulence

Genotypic Variation in SIVsm DeltaB670 *Env* During the Shift to High Virulence

P.J. Spencer Valli, Jonathan Heeney[1], and Jaap Goudsmit

Department of Human Retrovirology, Academisch Medisch Centrum, Amsterdam, [1]Department of Virology, Biomedical Primate Research Centre, P.O.Box 5815, 2280HV Rijswijk, The Netherlands

Department of Human Retrovirology
Academisch Medisch Centrum
Universiteit van Amsterdam
Meibergdreef 15, 1105 AZ Amsterdam
The Netherlands
telephone: (31-20) 566-4853
fax: (31-20) 691-6531

Abstract: The SIVsm DeltaB670 strain was serially passaged *in vivo* five times (Holterman *et al*, Journal of General Virology 80:3089-3097,1999). During passage the viral variants shifted from moderate (passages 1 and 2) to highly virulent (passages 3 to 5). This shift was not accompanied by a change in coreceptor use, as exclusive CCR5 usage was demonstrated for viruses with both modest and high virulence. The C1 to C4 region of 50 *env* clones derived from serum was sequenced. During the evolution towards increasing virulence there were no alterations in glycosylation sites, charge, hydrophobicity, secondary structure prediction, of the envelope region studied between the early and the late passages. Selection and fixation of nonsilent mutations found in the *env* of the heterogenous uncloned SIVsm DeltaB670 primary inoculum were observed during passage. Selection was most clearly marked by four amino acid alterations in the V1,V3,C3, and V4 region. The V1 contains a threonine (T) stretch possibly containing an O-linked glycosylation site(s) which a substitution (T134I) may alter. In the V3 region fixation of a valine (V) in place of a leucine (L), or isoleucine (I), at *env* position 327 is associated with increased virulence, as is K358R in the C3, and T423K in the V4.

Introduction: Primate lentiviral infection is characterized by extensive virus replication resulting in a quasispecies (32) consisting of a rated distribution of fitness variants. The numerous genetic variants produce a group of viral genomes that are similar but not alike. Substitutions in the viral proteins may cause an alteration in their folding, conformation and antigenic sites (49). Different variants may have different abilities to interact with the receptor, and coreceptors used to bind, enter, infect and replicate within different cell types (2). Evolution of the envelope of primate lentiviruses is correlated with changes in cell tropism (22). Although an expanded coreceptor repertoire is not a prerequisite for progression to AIDS (26), viral variants with virulent properties do emerge (42,53). During adaptation to a new species the viral quasispecies provides a variety of different and selectable candidates for the various genes which must vie for survival (77). The acquisition of virulence during infection with a large primary incoculum is known to be the effect of the competition between the many genomes present and the selection and or recombination of these during infection. As the average virulence increases so does the inherent advantage of being the first to infect the diminishing number of uninfected target cells

The outer envelope of the virus (*env*) is responsible for the binding and infection of cells, and contains determinants important in viral pathogenesis (62,80). The role of the outer protein in receptor binding, infection, syncytium induction and host immune response belies it's importance in the progression of SIV and HIV infection to AIDS. Analysis of the quasispecies variability during cross-species passage displays the molecular footprint of adaptation occurring due to the fixation of mutations specific to function in the new host environment. Study of the HIV and SIV *env* has led to much attention being placed upon the roles of glycosylation (1,7,14,15,51,55,73), charge (36,57), secondary structure (28,38,49,60), amino acid variation (12,13,23,25,54,55,61,69,71,74,88), tropism (4,5,16,57,62, 68,78,79,82), and coreceptor usage (18,30,31,46,66,75,92) upon escape from either B(11,20,48-50,60), T cell responses (17,38,90), or from increased chemokine concentrations (21).

The infection of a cell by HIV-1,2, and SIV begins with the binding of CD4 by the viral Env (*env* gene protein). For entry and infection, the fusion of the envelope and the cellular membrane must occur which is mediated by CD4 and a suitable coreceptor expressed externally on the host cell. The discovery that CD4 alone was not sufficient for binding but required a coreceptor (29,33) led to the identification of multiple coreceptors (31). Macrophage tropic strains (Mϕ) of HIV-1 use CCR5 and T-cell tropic strains (T-tropic) use CXCR4 in addition to CD4. Other coreceptors such as CCR2b and CCR3 also function in membrane fusion and infection and some viruses can use

multiple coreceptors (92). Repertoires of viral coreceptors play an important role in viral tropism and progression to AIDS. The CCR5 using strains of HIV-1 are either Mϕ or dual tropic (T cell or Mϕ) viruses usually found during the asymptomatic period (22), but can also predominate during disease progression (26). Evolution of the virus causes the acquisition of mutations in the *env* which can lead to the emergence of T-tropic viruses using CXCR4 (22). The broadened usage of coreceptors is believed to decrease the immune pressure on the virus by providing multiple escape routes into the target cells. Viruses found early in infection typically are slowly replicating Mϕ and non-syncytium inducing (NSI) whereas the viruses that emerge late in infection or during progression are usually T-tropic rapidly replicating, syncytium inducing (SI), and are associated with CD4+ T cell decline and progression to AIDS (22,81). Progression to AIDS without the SI phenotype and altered coreceptor use does occur, but less commonly (26).

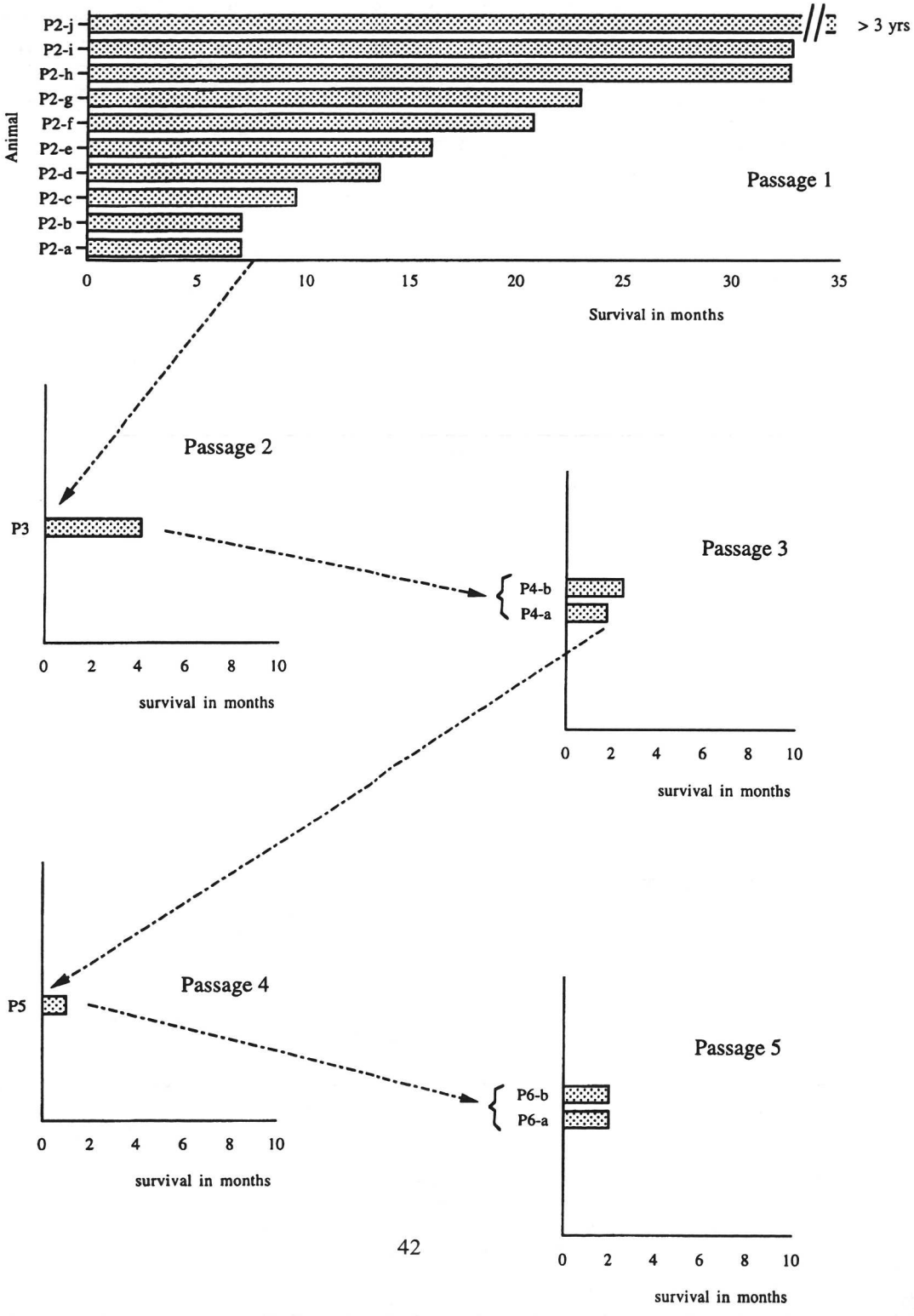
The coreceptors used by SIV are manifold as in HIV infection (31). The majority of SIV strains tested use CCR5 as the major receptor, but also use the growing myriad of other coreceptors in a strain dependent manner (31). The regions of the *env* used in binding the coreceptors by SIV is as yet undetermined, although it was shown that Mϕ and T tropic SIVs use CCR5 differentially, and bind the coreceptor in two different areas (30). It has been suggested that during disease progression there is evolution of particular pathogenic "end stage" variants, with particular sequences, which become fixed (42,53). In this study following *in vivo* passage we found that there was a narrowing of the quasispecies and the acquisition of particular sequences in the V1 to V4 domains. Four amino acid substitutions were associated with the increasing virulence which was apparent with each passage. The marked decrease in the asymptomatic phase of the infection after each passage resulted in increased virulence with no qualitative changes in the manifestations of disease (persistence of antigenemia and rapid loss of T helper/memory cells) during adaptation (47). No alterations were found in the exclusive CCR5 coreceptor usage during the acquisition of increased virulence.

Materials and Methods:

Virus. The SIVsm DeltaB670 strain used is one of the commonly used members SIVsm/HIV-2 family of lentiviruses (SIVmac, SIVsmmH4 and SIVPBj14) originally derived from sooty mangabeys when captive-bred Asian macaques were accidentally infected with SIVsm (23,24,44,70). The SIVsm DeltaB670 strain used is well studied and has been used in previous experimental infections of rhesus macaques because it accurately reproduces the clinical and pathological manifestations of AIDS seen in humans (3,5,6,8,41,52,65,70,71,85,91). The virus stock has been analysed as to genotypic and phenotype variation during experimental infection of non-human primates and during *in vitro* infection of rhesus macaque and human PBMCs (3,64).

Passage. As reported elsewhere (47), selective in vivo passage was carried out with sixteen age matched Asian rhesus macaques (*Macaca mulatta*) of one and a half years of age. Passage with SIVsm DeltaB670 was performed by intravenous inoculation, with the first rhesus macaque receiving $5x10^2$ infectious doses of cell free virus. The following rhesus macaques (P2 to P6A and P6B) were inoculated with uncultured cryopreserved peripheral blood mononuclear cells (PBMC) taken at the end stage of disease from the preceding rhesus macaque (Figure 1). Routine clinical biochemistry and haemotological analysis was performed at two week intervals, and following two months, at four week intervals. The CD4+ T cell changes in concentration were monitored by FACScan (Becton Dickinson) using double labelling with Leu3a and 4B4, as described earlier (47).

Figure 1: The survival in months of animals in the serial passage of SIVsm DeltaB670.


P2-j
P2-i
P2-h
P2-g
P2-f
P2-e
P2-d
P2-c
P2-b
P2-a
Animal
> 3 yrs
Passage 1
0 5 10 15 20 25 30 35
Survival in months
Passage 2
P3
0 2 4 6 8 10
survival in months
Passage 3
P4-b
P4-a
0 2 4 6 8 10
survival in months
Passage 4
P5
0 2 4 6 8 10
survival in months
Passage 5
P6-b
P6-a
0 2 4 6 8 10
survival in months

***Env* RT-PCR and Cloning Strategy.** Genomic viral RNA was extracted from macaque serum using silica beads and a chaotropic agent according to the Boom method (9), and used as template in a reverse transcriptase reaction followed by PCR amplification. The 3' PCR primer was also used to prime the reverse transcriptase reaction, followed by a single PCR reaction. Extraction of the viral RNA from 15 ul volumes of serum was completed by eluting in 10 ul H_2O containing 40 units of RNAsin (Promega, Madison, Wis.). The 10 ul volume of nucleic acids was then divided into two, and 3.0 ul of 5 x OF (One-Tube (83)) buffer (67mM Tris (pH 8.9), 50mM KCl, 6 uM EDTA, 0.1% Triton X-100, and 0.5 mM DTT) was added to each aliquot as well as 4 units of SuperScript I (Boehringer Mannheim, Germany), 1.5mM Mg^{2+}, 10 mM of mixed deoxynucleoside triphosphates (dNTPs) and 1 ng of 3' primer (5'TATAGCGGCCGCCTACTTTGTGCCACGTGTTG(8173)) with a final volume of 15 ul. Reagents were assembled at room temperature and incubated at 37°C for 90 minutes. Primers for RT-PCR (Applied Biosystem Inc., The Netherlands) correspond to nucleotides in the *env* region of SIVsmH4 (44), sequence coordinates are given following the 3' base, underlined nucleotides represent restriction enzyme sites included to facilitate cloning (5'GC GAAGCTTGGGATAATACAGTCACAGAAC (7089)). The 3' primer for PCR being the same as used in the RT reaction. The template for the PCR amplification was composed of 5 ul each of the duplicate RT reactions, 10 ng each of 3' and 5' PCR primer, 10 ul of OT buffer, 25 mM dNTPs, 1.5 units of Taq polymerase (Perkin Elmer Cetus) and H_2O to a final volume of 100 ul. PCR cycling consisted of 1 minute at 95°C, 1 minute at 55°C and 2 minutes at 72°C (repeated 35 times), final extension was carried out for 10 minutes at 72°C. PCR-amplified products were once again purified with the Boom method (9), eluted and then digested with Not1 and HindIII. Purification on an 0.8% agarose gel followed, and extracted with the Boom method (9), as above. The eluted fragments were ligated overnight into plasmid pSP73 (with an added Not1 site between the Kpn1 and Sac1 sites (Promega, Madison,WIS.) and transformed into *E.coli* DH5a by electroporation (Cellporator, Bethesda Research Laboratories, Gaithersburg, Md.).

Sequence Analysis. Double-stranded DNA was harvested by Qiagen column, and used for cycle-sequencing using the ABI PRISM Sequence Reagents (Applied Biosystem Inc., The Netherlands). Sequence primers were the SP6 and T7 standards and four SIVsm *env* specific designed from SIVsm DeltaB670 sequences provided by Chris Contaq (Stanford) and Michel Murphey-Corb (Tulane Regional Primate Center). The primers are 3'svseqV2V1 CAGGAGCAAAGCATAACCTGGCGGTGCAC, 5'svseq V2V3 CAGGAGAATCCTGTGACAAGCAT TATTGG, 3'svseq V3V2 CAGGACCTGATTGGG TCTCTCATTTATGG, and 5'svseqV3V4 CAGGAGTCATACCAGTCACCATTATGTCAGG. The timepoints for sampling by sequencing of serum viral RNA were seroconversion and death. The sequenced clones were named by passage number (P1 to P6A and B) and seroconversion (S) or death (D). The sequences were aligned using CLUSTAL W(41) and adjusted by eye. The proposed amino acid alignments were translated using MEGA (59). Sequence analyses were performed with the MacVector (Kodak) program using windows of hydrophilicity and amphiphilicity of 8 and 11 respectively. Secondary structure predictions were produced by the neural network protein secondary structure program PredictProt of the Protein Design Group EMBL using the MSF format of analysis.

Chemokine Receptor Usage. The coreceptor usage of the P5 (isolate 8980) and P5 clone (smF539) and other strains (SIV mac 251, and macBK28) will be published in full (46,66). Briefly, CD4 transformed HOS cell lines expressing CCR1, CCR2b, CCR3, CCR4, CCR5, CCR8, CX3CR1, CXCR4, BOB, and Bonzo were used. Reported here are only the CCR5, BOB, and Bonzo determinations as the passaged virus used only CCR5 and the comparable strains used CCR5, BOB and Bonzo. The CD4 HOS cells containing HIV-2 LTR driving green fluorescence protein (GFP) via the CMV promotor driving a hygromycin-resistant construct were grown in DMEM, 10% FCS, 0.5 mg/ml

neomycin, and 100ug/ml hygromycin, with the pBABE-puro introduced coreceptors held under 1ug/ml puromycin selection. HOS-CD4 coreceptor expressing cells were seeded in wells at $2x10^4$ per well and infected with 0.5 to $800x10^3$ TCID of the stocks with 20ug/ml Polybrene for 2 hours at 37°C. Cells were washed, cultured for 48 hours and analysed for GFP fluorescence with FACS (46,66). Shown are only the coreceptors relevant to the SIVsm DeltaB670 strain and only as positive or negative. The coreceptor usage for the P1 to P5 viruses did not vary and was exclusively CCR5 (Table 1 (46,66)).

Table 1: Summary of CD4-coreceptor mediated infection assays. HOS-CD4 cells containing the HIV-2 LTR driving GFP and expressing the indicated pBABE-puro introduced coreceptor were infected with virus stocks and 48 hours later were analysed for GFP fluorescence with FACS (46,66), data shown are only positive or negative and limited to the coreceptors used by the B670 derivatives. The SIV mac239 and smLib-1(18) are also in HOS-CD4 cells and SIV B670CL12 and 3 are in QT6 cells (31).

Summary of Coreceptor Usage of Passaged SIV and Related Strains				
Virus Isolate	**CCR5**	**CXCR4**	**BOB**	**BONZO**
B670 (P1)	+	-	-	-
B670 CL12	+	-	-	-
B670 CL3	+	-	-	-
P5	+	-	-	-
smF359 (P5)	+	-	-	-
smPBj6	+	-	nt	nt
MAC BK28	+	-	+	+
MAC251	+	-	+	+
MAC239	+	-	+	+
smLib-1	+	-	+	+

*Is infectious for MΦ but not productive (46).

Disease Characteristics. The passage of SIVsm DeltaB670 in juvenile rhesus monkeys resulted in the decrease of the asymptomatic period from eighteen months to one and two months, by the fifth and sixth passage respectively (Figure 1 (47)). Disease development was characterised by rapid loss of CD4+ T cells persistent and persistently high level plasma viremia. A marked reduction of the asymptomatic period occurs within three consecutive passages, dropping to weeks after P2. The P3 inoculum resulted in the decrease of the asymptomatic period to weeks, and the increase of the p27 concentration to more than 25 nanograms per millilitre within two weeks. The increasing virulence peaks at P4, showing a radically aggressive progression with almost one hundred nanogram per millilitre concentrations of p27 and progression to AIDS within one month. Further passage of the virus did not result in any increased gains in virulence. The P6A and P6B animals progressed to AIDS within two months with rapid CD4+ T cell loss and high persistent plasma antigenemia levels (47).

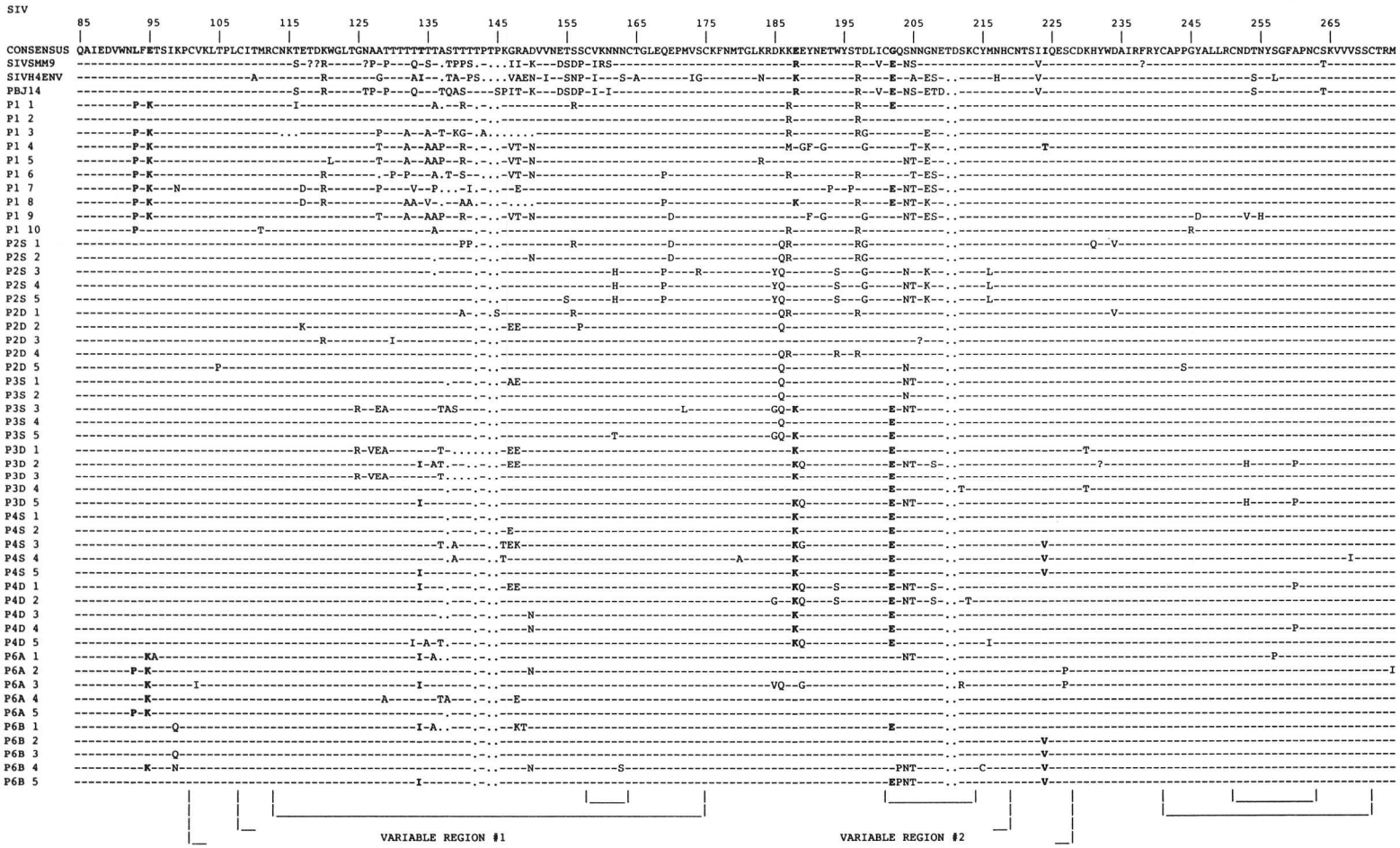
Results:

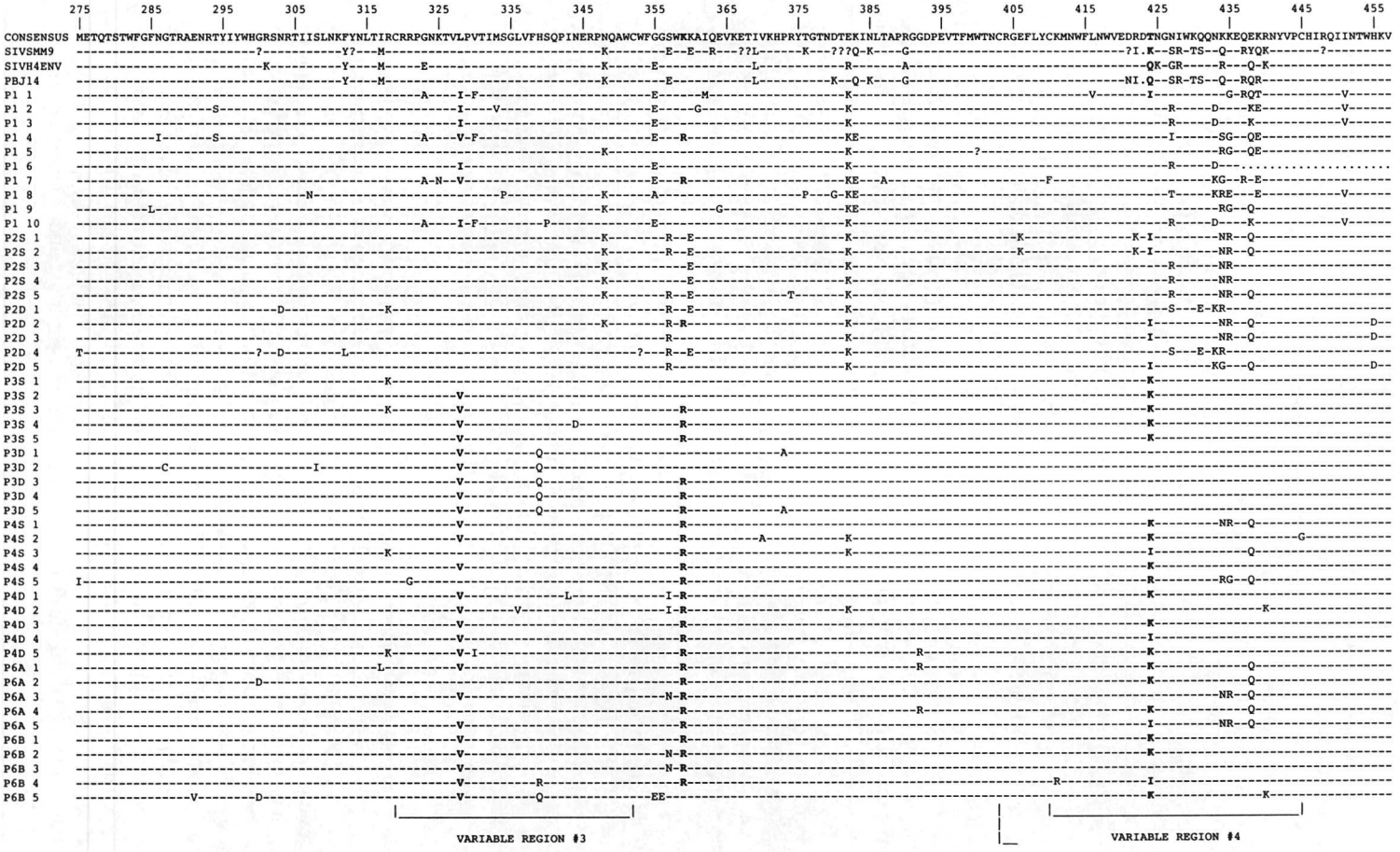
The *env* was sequenced from multiple clones at seroconversion and death in each rhesus macaque (*Macaca mulatta*), after the initial inoculation and passage of SIVsm DeltaB670. The pattern of *env* mutation and fixation was found to occur as a stepwise selection during adaptation of a number of variations found in the heterogenous primary inoculum, but also included changes not found earlier (Figure 2). The variable (V) domains contained the majority of the substitutions, although the constant (C) domains contained fixed and spontaneous mutations as well. The V1 and V4 were found to be the most variable of the *env* domains with accompanying mutations in the V2 and V3 as well as the constant regions. Glycosylation is reported to play a role in infection (1,7,15,51,55,71,73) because of it's effect on conformation, and on pathogenesis by means of epitope masking (10,15,50). However, no alterations in this regard were observed in the sites contained in the *env* region sequenced. Neither the charge of the predicted translation products or the secondary structure changed. The sequence variations were compared to sequences from the published literature on experimental infections and similarities were found in the specific sites with fixations and sporadic mutations.

The major sequence variation occurred in the first (V1) and fourth variable region (V4), as is common to SIV infection in non-human primates (4,10,12-14,20,23,35,43,55,56,68,71,73,74,76,77,87,88). Alterations were also seen to occur in the other *env* regions sequenced, including the constant domains. All fifty clones sequenced were found to be unique, but there were predominantly founder derived (early passage or P1) substitutions (Figure 2) selected and fixed during passage. The *env* substitutions seen were traceable to the multiple alignments produced by sequential isolate sequencing from the published literature of experimental infection of Asian monkeys with the most commonly used strains of SIV (see Table 2 for references as to relations). There was one stop codon introduced presumably by a Taq or a reverse transcriptase error and a deletion C terminal to the V4 region in one clone (P1-6).

The selection of amino acid substitutions found within the heterogenous P1 sequences is scattered throughout the *env* regions sequenced. There were fourteen sites of recurrent substitutions where the mutation selected later in the passage could also be found as one of the many variations seen in the heterogenous P1 sequence population (Figure 2). Eleven of these were sites of charged amino acids and two were substitutions between aliphatic amino acids. The PFK motif (amino acids 93-95) of the C1 is seen in the majority sequence of P1 sequences and is found again in the P6 sequences, but not elsewhere in between. The D150N motif is found in half the P1 clones and then reappears in a minority of the P2 and further sequences. The E188K is present in one P1 sequence and then dominant after P3S. This charge changing is also seen in SMM9,H4 and PBj14, although for E188R. Selected from the P1 sequences were G202E, I223V, L327V G355E, K358R, K381E, T423K,I and R, K432,433N or R, and lastly E437Q, R439K.

Figure 2: Alignment of the predicted translation products of the SIVsm DeltaB670 envelope sequences V1 to V4 of clones from the viral RNA in the serum of the serially infected rhesus monkeys. The amino acid sequence of the passage clones are compared in order with a consensus sequence, SIVsmm9, SIVsmH4 and SIVsmPBj14. Amino acid substitutions are indicated, similarity by a dash, a dot indicates a gap introduced to maximize alignment, and an asterisk indicates an in-frame stop codon. The cysteine loops contained within this envelope fragment are indicated below the alignment by dashed lines and labelled by the variable regions contained within them. Numbering is from the beginning of the gp120 envelope protein.


SIV
85
95
105
115
125
135
145
155
165
175
185
195
205
215
225
235
245
255
265
CONSENSUS QAIEDVWNLFETSIKPCVKLTPLCITMRCNKTETDKWGLTGNAATTTTTTTASTTTTPTPKGRADVVNETSSCVKNNNCTGLEQEPMVSCKFNMTGLKRDKKEEYNETWYSTDLICGQSNNGNETDSKCYMNHCNTSIIQESCDKHYWDAIRFRYCAPPGYALLRCNDTNYSGFAPNCSKVVVSSCTRM
SIVSMM9
SIVH4ENV
PBJ14
P1 1
P1 2
P1 3
P1 4
P1 5
P1 6
P1 7
P1 8
P1 9
P1 10
P2S 1
P2S 2
P2S 3
P2S 4
P2S 5
P2D 1
P2D 2
P2D 3
P2D 4
P2D 5
P3S 1
P3S 2
P3S 3
P3S 4
P3S 5
P3D 1
P3D 2
P3D 3
P3D 4
P3D 5
P4S 1
P4S 2
P4S 3
P4S 4
P4S 5
P4D 1
P4D 2
P4D 3
P4D 4
P4D 5
P6A 1
P6A 2
P6A 3
P6A 4
P6A 5
P6B 1
P6B 2
P6B 3
P6B 4
P6B 5
VARIABLE REGION #1
VARIABLE REGION #2


275 285 295 305 315 325 335 345 355 365 375 385 395 405 415 425 435 445 455
CONSENSUS METQTSTWFGFNGTRAENRTYIYWHGRSNRTIISLNKFYNLTIRCRRPGNKTVLPVTIMSGLVFHSQPINERPNQAWCWFGGSWKKAIQEVKETIVKHPRYTGTNDTEKINLTAPRGGDPEVTFMWTNCRGEFLYCKMNWFLNWVEDRDTNGNIWKQQNKEQEKRNYVPCHIRQIINTWHKV
SIVSMM9
SIVH4ENV
PBJ14
P1 1
P1 2
P1 3
P1 4
P1 5
P1 6
P1 7
P1 8
P1 9
P1 10
P2S 1
P2S 2
P2S 3
P2S 4
P2S 5
P2D 1
P2D 2
P2D 3
P2D 4
P2D 5
P3S 1
P3S 2
P3S 3
P3S 4
P3S 5
P3D 1
P3D 2
P3D 3
P3D 4
P3D 5
P4S 1
P4S 2
P4S 3
P4S 4
P4S 5
P4D 1
P4D 2
P4D 3
P4D 4
P4D 5
P6A 1
P6A 2
P6A 3
P6A 4
P6A 5
P6B 1
P6B 2
P6B 3
P6B 4
P6B 5
VARIABLE REGION #3
VARIABLE REGION #4

The variation in the V1 loop produced no extra N-linked glycosylation sites although there were three P1 clones with sites additional to those in the prevailing P1 sequences (P1-5,6 and 10) that were not seen in later clones. No additional glycosylation sites were seen in the adapted sequences as found in other experimental infections (16,73,74). Most variation in the V1 was Ala to Thr (eg. A135T) or the reverse with selection for the stable Pro (P135) in the tip of the V1 loop. The variation in the V1 region is seemingly strain or host dependent as there was little in common between the sorts of variants (Table 2). The passage V1 substitutions being mostly A,T,P,Q or S single substitutions within the serine-threonine rich region, with no obvious pattern except for the absence of glycosylation site alterations. The only alteration seen in the passage being the N164S mutation in clone P6B-4 which shifts a NXT site to a NXS site one amino acid further. No pattern was seen in the variation of charged sites with substitutions seen between positive, negative and neutral amino acids. The T134I substitution seen in the later passages is found commonly in SIV SMH4 clones (43) and although not fixed in all clones alterations in this site may involve O-linked glycosylation variations (15).

The V2 region variations were more limited with charged amino acids being substituted for like-charged amino acids (such as R187K), with the same exchange seen in SIVMne infection (74). The R197T mutation removes a charge that exists in SIVsmH4 but not in the other strains. The eleven amino acid V2 accessory loop is highly variable with most variations being size conservative. The D198G is found in some SIVmac239 clones (6/16 (61)). Variation in the V2 region consisted of four amino acid changes, two of which were variations away from charged amino acids (R-T and E-G). Half of the variation was found in the small "accessory" loop of the V2 consisting of only ten amino acids (amino acids 201-214).

The C2 region was the most stable domain having only four amino acid changes in the P6B sequences (I223V), as seen in the other SIVsm strains. The I-V or V-I variation was seen three times in the passage (C-terminal to the V2 region, mid in the V3 loop and in the CD4 binding region) and is a conservation of aliphatic amino acids. The I223V substitution is seen in four of the five P6B clones, producing the motif NTSIVQ. All other SIVs have the motif NTSVIQ with no variation of the glycosylation site or of the V or I to give the NTSIIQ seen in the other fifteen clones or the P6B clones. Situated between two disulphide bonds eight amino acids apart containing an N-linked glycosylation site there is possibly a constraint for aliphatic amino acids as no change in flexibility or antigenicity was seen. The two T294S substitutions shift an N-linked glycosylation site one amino acid .

Variation in the V3 loop was minimal compared with that in HIV-1 infections and consisted of one site with four G323A changes, both being tiny hydrophobic amino acids. The L327V mutation is a mix of all three aliphatic amino acids, found to be important at the tip of the SIV V3 as well as in HIV-2 (57).This selection and fixation occurs concomitantly with the persistent antigenemia seen in these later infections, variation in the V3 of quickly progressing monkeys was also found in SIVsmH4 infection (43), but not at the same site. The variation at this site was not found elsewhere with other strains having a conserved L at this site except for one SIVmac239 clone having a V at this site (1/16 (12)). The single change in the V3 domain is one of the three I-V selections in the P6 clones but two clones were also found to encode a Lys (L) at the same position but also being aliphatic. The post-V3 K358R selected in almost all clones mimics that variation seen in *env* clones from the brain of a SIVsmF236-infected macaque (13). The unchanged predicted structure for the V3 loop (between early and late passage virus), although more hydrophobic, was similar to the sheet-turn-sheet form of the HIV-1 V3 (40,60), and supported by the similarities and conserved

elements in the V3 sequences of HIV and SIV (49).

The start of the C3 region is a site of variation common between SIV and HIV-1 strains (67) and also during the passage. The G355E being in the P1 clones is also selected in SIVsmH4 and SIVsm/F236 infections (13,43), but arguably not important in adaptation. The S357N mutation is the opposite of that seen in SIVsmH4 infection (13) where the N was selected with 1/25 reverting to the S. The K358R variation was not found elsewhere but the opposite charge effect was seen (K-E) in the SIVsmH4 and the SIVmac239 infections (12,43).

The V4 region is variable in the same manner as the V1 region, in that no discernible pattern can be found in the various strains. The SIVMne infections predominate with NXT or NXS shifts (14,73,74) but there are no variations as such within the passage or other strains. The V4 region of the P6 sequences is more hydrophilic due to the T423K and the R421K substitutions although no other changes were predicted and the changes seen in the smH4 sequences is a general tendency toward positively charged amino acids while the mac variations were neutral. The T423K substitution is found in the SIVSMM9 *env* as an invariant position (37). The passage P6 sequences all increased in charge by two in the P6A clones whereas the P6B clones increased by one due to the Q432E mutation. The C4 which contains the CD4 binding region had one variation in the passage sequences with the V450I substitution selected in all ten P6 clones and is the norm in this region of all SIV strains and is invariant.

Tissue dependent changes found in the V1 region glycosylation sites in smH4 infections were not encountered (13,43). Adding or deleting trios of T and S rich motifs and with the loss or gain of NXT or NXS glycosylation sites as seen in most SIVMne infections (14,73) differ in character from those in mac142,239,251 and Bk28 infections (4,11,12,20), and also to those in our passage study. The only V1 glycosylation alteration was at A128T in three P1 sequences. This change is only found in sequences with an added proline (L93P), charge reversing E95K, and further only with in clones with the "AAP" motif of SIVB670-CL6 (Figure 2 (3)). Another site of additional glycosylation is the N163S of P6B-4, which is also found at this site in SIVsmH4. In the V2 accessory loop a site is created, N208S, also seen in SIVsmH4 and SIVPbj14, and seen in minority in P1, P3D and P4D. In the tip of the V4 (N426T/S) a glycosylation site is found in minority in the P1, P3D, and P4D sequences (common to SIVSMM9 and SIVPbj14 as well). Although not selected for during passage and adaptation, it is of note that these seemingly extraneous substitutions sites are only found in the death samples, and not during seroconversion samples. The T134I found in clones after P3S is similar to the O-linked glycosylation sites found in SIVMne infections and proven to play a role in antibody neutralization (15,53).

Table 2: The variations seen in the passage sequences were compared to those seen in other experimental infections. Listed are the common variations or variations in similar motifs in the various SIV strains the numbering corresponds to that of B670 or smH4 although variation between strains is one to four amino acids in length at most.

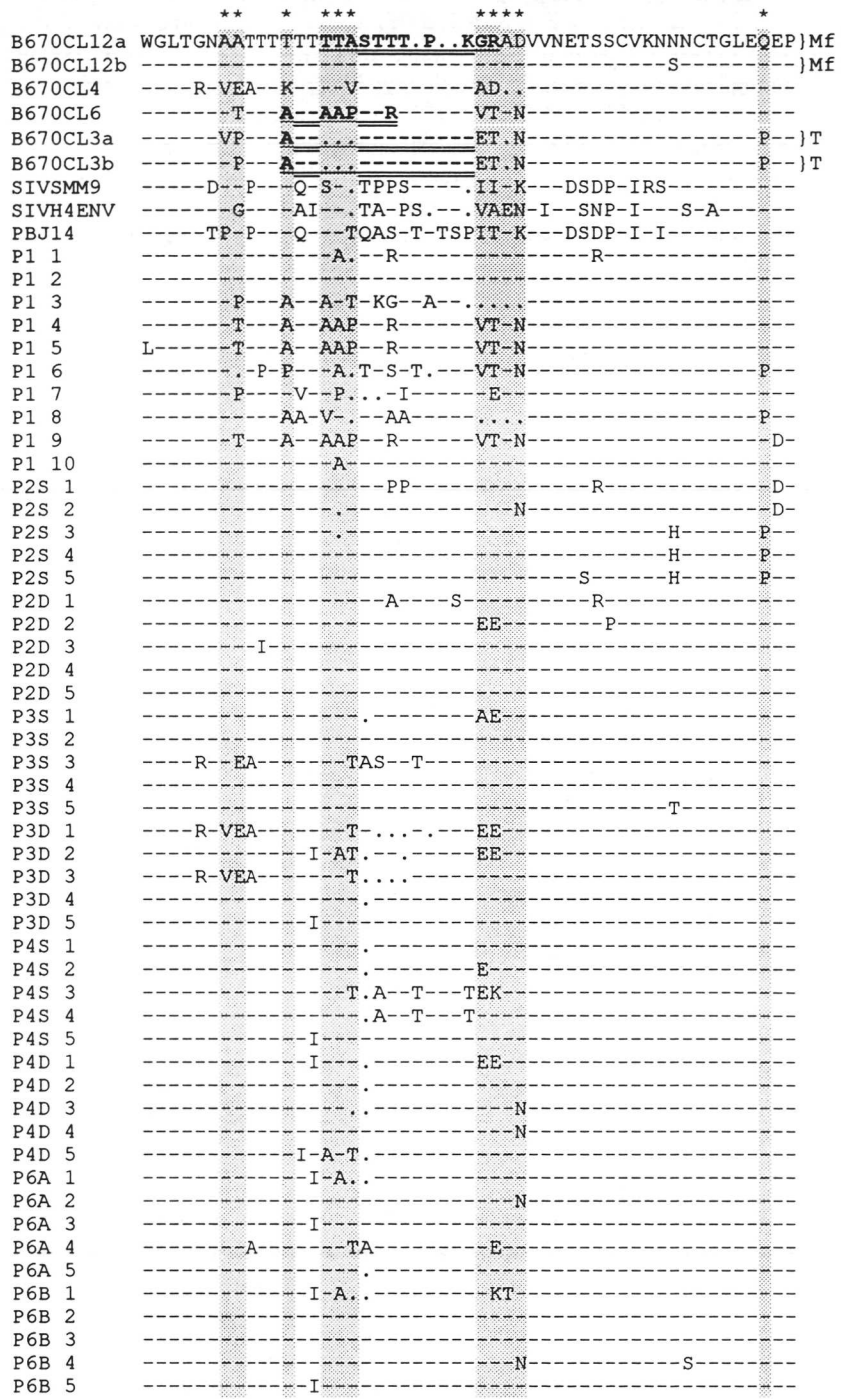
Mutation	Region of Mutation	SIV Passage Strain	Reference Number	Comments
L93P	C1	SMH4	(43)	seen in all SIVsm
E95K	C1	SMH4	(43)	seen in all SIVsm
K120R	V1	SMH4, MAC142,239&251 MNE	(43) (20) (73)	3/9 clones always K always R PBMCs only
T134I	V1	SM/F236	(13,43)	rhesus and nemestrina
T133,135&136A	V1	SMH4, F236& MNE	(13,43,73) (14,74)	not seen in SIVmac
T140R	V1	MAC239	(4)	neurotropic
K187R	V2	MNE & MAC239	(12,54,74)	14/18 clones in MNE, 1/16 MAC239
T197R	V2	MNE & MAC239A-T	(74) (12)	5/18 MNE 4/16 MAC239
S204N	V2	MAC142,239,251	(20) (11)	always S always N
N205T	V2	SMH4	(43)	N in others
G207E	V2	SMH4 E-K	(43)	G in all others
I223V	C2	4/5 P6B	-------	all others V
G322A	V3		-------	all others G
L327I,V	V3	MAC2391/16 L-V	(12)	all others V
S333P	V3	MNE	(74)	
N347K	V3	39/50	((34)K-R)	all others K
G354E	C3	SMH4 MAC(all) MACBK28, MNE	(43) (4,11,12) (11,20) (74)	1/11 G-E all G all G all E
S356N	C3	SMH4 MAC(all) MNE	(11-13,20) (11,20) (74)	24/25 S-N K-N N
K358R	C3	SMF236	(13)	16/54
K381E	C3	all	-------	K-R
G391R	C3	MAC239	(34,68)	linked to macrophage tropism
T423K	V4	SMM9	(37)	invariant in SMM9
K433R,N	V4	SMH4	(43)	P-R 6/11 P-K 1/11
E437K,Q	V4	SMH4	(43)	K-R 5/11
V450I	V4	all	-------	all I

The P6 sequences contain a relatively low level of intrasample variation as compared to the earlier sequences, although the P6 isolates are more virulent (Figure 2 and Table 2). The loss of immune control may explain the more homogenous, less variable P6 sequences as these animals do not seroconvert and develop no measurable immune response as is seen in SIVsmPBj14 infection of pig-tailed macaques (71,82). The changes in variable regions as reported in various experimental infections (4,11-14,20,43,68,73,74) related to the antigenicity or natural history are clearly ordered; V1>V4>V3>V2 (Figure 1). The variation seen in the C3 region just C-terminal to the V3 loop is the exception with the C3 being more variable than the V3 region, although the C3 constant domain is sometimes more variable than the V3 variable domain in HIV-1 infection (67). The C2 region is stable in all of the sequences. The four amino acid substitutions possibly linked to increased virulence are T134I, L327V, K358R, and T423K.

Alignment of the V1 variable domain sequences with SIVsm DeltaB670 sequences found to be crucial to transplacental transmission, macrophage tropism (Mϕ) and T cell tropism (3,30) and CCR5 coreceptor usage (30,31) showed that four main variants of virus (Figure 2, B670CL3a,b, and CL12a,b) found in the primary inoculum were also present up to P3D, and only one following that passage. This is likely due to a selective bottleneck event during transmission (Figure 1, and 2). The study of infection of pregnant monkeys and genotypic analysis of the viruses transmitted to offspring allowed Amadee et. al. to show genotypic selection and transmission of macrophage tropic viruses (Mϕ (3)). The motifs found in the V1 of the most transmissible Mϕ virus (B670-CL12 (clones L45 and L30 used as examples (64))) were the same motifs selected during this passage. The Mϕ virus marker sequences were found to increase from the P1 passage to a majority of the clones sequenced after the second passage (P2D to P6A and P6B, Figure 2). The other viruses seen in the P1 sequences were the T cell tropic B670-CL3 (clones B4 and B7), and three B670-CL6 sequences, although not seen again in the later passages where the heterogeneity of the quasispecies decreased as the virulence increased. The T cell tropic motifs of the SIVsm DeltaB670 variants found in infected rhesus and human PBMCs respectively (3,64), were not seen after P1 (Figure 2). The positive to negatively charged change of "GR" for "ET", "VT", and "EE" is in a hotspot of variation in the V1 during passage (Figure 2, fourth shaded column of amino acids). The B670-CL4 sequence not sampled earlier in the passage, but present in the primary inoculum, are found again in the P3S and P3D sequences (P3S-3, P3D-1&3), a sampling result also seen (3) to occur in mixed virus inoculations and multiple timepoint sequencing during infection. The predominating V1 sequence is the macrophage tropic one (AA-T-TTA-GRAD-Q, as seen in the shadowed columns of the alignment in Figure 2 (64)). A majority of the fifty clones sequenced contain the GRAD motif, with seven of the clones after P2S being exactly the same in all positions as B670CL12.

Figure 3: Alignment of the predicted translation products of the SIVsm DeltaB670 envelope V1 domain of clones from the viral RNA in serum of the serially infected rhesus monkeys, and three SIVsm DeltaB670 clones known to have specific tropisms and transmission characteristics (3). The predicted amino acid sequences of the passage clones are aligned in comparison to clone B670C-L12A and B (macrophage tropic, CCR5 dependent (31) and most frequently transmitted transplacentally), B670CL-4 (source of RNVEA motif seen in P3S and P3D), B670CL-6 (source of AAP motif seen in P1), and B670CL-3A and B (T-cell tropic and CCR5 dependent (31)). Amino acid substitutions are indicated using the one letter code and similarity to B670CL12A by a dash.

SIV B670 and Passage Sequence Env V1 Amino Acid Alignment



Discussion: Alignment of the *env* V1 variation during passage shows the presence early in the passage of clones similar to the B670-CL3A, 3B (T cell tropic, CCR5 dependent (3,31)), B670-CL4, and B670-CL6 sequences (Figure 2) present in the SIVsm DeltaB670 quasispecies. Later passages have more homology to the Mϕ B670-CL12 (CCR5 dependent (31)) sequences. The homologies of the selected sequences and coreceptor usage of the virus provide evidence for the tropism of the viruses, in relation to the variation observed. The passage in rhesus macaques showed the quasispecies heterogeneity decreased with the acquisition of increasing virulence and the predominant virus was the Mϕ or R5 (CCR5 coreceptor) tropic virus previously documented to be important in transplacental transmission (3,89).

The characterization of some of the coreceptors used by HIV-2 and SIV (27) has shown that coreceptor usage of the Mϕ and T cell tropic strains of these viruses (19,26,30,31,84) can differ during the stages of infection, but that it is not a requirement for progression to AIDS (26). The change in coreceptor usage from Mϕ to T cell tropic in HIV-1 infections is known to coincide with a switch in coreceptor usage (22,75). The coreceptors involved in HIV-1 infection being CCR5 and CCR3 for the macrophage dependent (R5) variants and CXCR4 for T cell infecting viruses (X4) and R5X4 which have dual tropism. Concomitant to the switch from the Mϕ tropic R5, and repopulation of the quasispecies with T cell X4, R5X4 variants, is the loss of CD4+ cells and phenotypic transition from non-syncytium inducing to syncytium inducing (NSI to SI (22,75)). The close relation of HIV-2 and SIVsm is shown by their high *env* sequence homology (up to 60 %), although differences have been found in the utilization of CXCR4 (27). The broadening or multi-usage of coreceptors may be a correlate of disease progression in HIV-2 (75). The tropism switch, phenotype alteration, and loss of CD4+ cells occurs in HIV-1,2 and some SIV infections (not including asymptomatic infection of Chimpanzees or African green monkeys). However of the SIVs evaluated so far all except SIVcpz use CCR5 (22,27,78,84). T cell tropic HIV-2 uses CXCR4 and not CCR5 (which are both utilized by T cell tropic HIV-1 (84)). The pattern of SIV coreceptor use also differs in using CCR5, Bonzo and BOB (27) for macrophage and T cell tropic viruses alike (30). Oddly, in this study the T cell tropic B670-CL3 CCR5 dependent genotypes are found only in the early samples. This does not mimic the phenotype switching seen in HIV-1 infections sometimes found during progression to AIDS in humans (22,78), but agrees with the finding of disease progression in the absence of SI viruses and without the associated broadening of the coreceptor repertoire (26). The lack of T cell tropic sequences in P6 may be due to the very short progression times after P4, and thus the of time and/or selection pressure to cause phenotype switching to exclusive T cell tropic variants.

The total *env* protein molecular weight is made up of 64% amino acid chain and the remaining 36% composed of N and O-linked glycosylation. Glycosylation sites play an important role in modulating humoral immune responses against envelope antigens. They may act to totally dissipate Ab binding after carbohydrate removal (1) and the alteration of these sites should be an effective method of immune evasion (11,20,39,49). In SIVMne infection the predominant variation was seen to be the addition/deletion/shifting of the NXT/S N-linked glycosylation sites in the V1 and V4 regions (14,15,53,73,91). The HIV-1 *env* is O-linked glycosylated and predicted to have eight O-linked glycosylated sites per gp160 by molecular weight (7). O-linked glycosylation sites are known to be rich in serine and threonine although the actual sequence needed is elusive. The changes seen in the V1 in the chain of threonines (notably the T134I substitution) could be alterations in the location and amount of O-linked glycosylation, as the role in antibody recognition or evasion played by O-linked glycosylation has been shown (15). There was little or no alteration of the N-linked glycosylation sites in the

V1 or the V4 of the P3 or later sequences as compared to those of the P1, P2, or P3 isolates. There are undoubtedly varying roles played by glycosylation in the various strains during infection, but it is unexpected that so little or no difference be found between the P1 and P6 clones in relation to the increased virulence. Due to the rapid progression of the infections after the first passages, the time needed for seroconversion (three to four weeks) was not reached. The mutations seen in this passage experiment were found to exist to some degree in all of the commonly used strains (4,11-14,35,43,68,73,74). Amino acid substitutions common to the present passage experiment and other experimental infections suggests that limited variation is permitted after a state of optimal fitness is reached (53). The similarities in the variable motifs within the common variable regions are to an extent similar within these infections. Most of the selected mutations were charge altering (positive to negative, negative to positive or charged to uncharged) making up the majority of the changes while composing only 15% of the amino acids (K, R and E). That charged amino acids commonly found exposed at the surface of the envelope could be the cause of the over-representation of these three amino acids in the variation during passage.

The antigenic area found to produce the most efficiently binding antibodies was referred to as the principal neutralizing domain (PND). The PND of SIV differs from that of HIV-1 and 2 is conformational and localised between the C2 and the C-terminus of the Env protein (48,50). The mutations seen here could be descriptive of a viral Env protein with an even less immunogenic tertiary structure so as to evade the humoral immune response lagging during the initial post-infection viremic stage. The decreased heterogeneity in the P3, P4 and P6 sequences gives support to the notion of driving viral variation by selection for variants present in the host (63,72). The need for a stable V3 loop in the conformational folding of the SIVsm Env may be why there are no linear-V3 peptide binding Abs in SIV infection as opposed to those seen in HIV-1 infection (50). If the V3 must remain in conformation, it could be that minute variations provided by the C3 variations are sufficient to allow correct orientation of the V3 loop in its function and structure.

Three studies have shown the importance of the SIVsm V3 in tropism and pathogenicity (43,57), and of CCR5 usage (30). The first change shown in the V3 loop (H342P) allowed the infection of PBMCs and macrophages in culture (43). The assumption in this latter study was that variation at this V3 residue, in addition to downstream substitutions, affected cell tropism of SIVsm and thus indirectly had an effect on virulence. In our passage experiment the P6 clones were found not to possess this mutation although the disease characteristics were the same. This proline is present in the B670 isolates and was found to be stable in the passage sequences throughout this study. Although a similar G to R substitution, as seen in the passage, was found to be necessary for the macrophage tropism of SIVmac239 (68), the V3 substitution here is just 3' of the loop (K358R), thirty residues away. The exact same G391R substitution is found in one P4D and two P6A sequences. The V3 and C3 changes seen (L327V and K358R) selected during passage may have effects upon altered coreceptor usage, or possibly even CTL epitopes. Reduced replication in macrophages, was shown to be the result of two substitutions in the V3 and these produced a virus that used CCR5 in a more T-tropic manner (30,57,58). Variations found in the V3 were shown to alter the region of CCR5 used with dependence on the second extracellular loop as do T tropic strains (30,57). The V4 mutation T423K varies between SIV SMM9 (K), H4 (Q), and PBJ14 (I) with the T423I found in some passage clones as well as in PBJ14, what role this mutation plays in the infection has not been studied.

During passage the virulence of this uncloned strain increased markedly. The "differential utilization" of Mϕ and T tropic SIV strains of CCR5 may explain the ability of two separate B670 clones (CL3

and CL12, or P1 versus P4) to have different *env* sequences and yet both utilize the same coreceptor exclusively (31). The selection during passage of a CCR5 dependent Mϕ tropic strain has been shown to be important in pathogenesis (45,86). Although the *env* variation found during passage may not lead to alterations in coreceptor usage, it may be involved in increased affinity for CCR5 or for the adaptation to receptors/coreceptors not yet known. The data presented support the concept that alterations in coreceptor repertoire are not important in AIDS pathogenesis caused by virulent strains of SIV. There were neither selected changes in glycosylation, nor charge. Detailed comparison of these changes within the passage revealed the commonality of these selected amino acid substitutions with the published literature, suggestive of a shared rule of *env* variation, probably under the biochemical constraints of tertiary structure and function. Selection and fixation of amino acid substitutions found in the heterogenous uncloned primary inoculum SIVsm DeltaB670 happened simultaneously with the development of increased virulence. The four substitutions selected occurred in the V1, V3, C3, and V4 (T134I, L327V, K358R, and T423K), and were all found to be minority substitutions acquired from SIVsm DeltaB670. These substitutions selected may together or in part be involved or associated with increased virulence.

Acknowledgements: Many thanks to Ineke Braakman for illuminating the protein folding, Bill Paxton for many helpful discussions, and Chris Contaq for unpublished sequences, pages and pages of faxed alignments and virtuoso sequence primer design help. We are indebted to Henk Niphuis expertise in performing the clinical aspects of this study, and Wim Koornstra for the coreceptor analysis.

References:

1. **Alexander, S. and J. H. Elder**. 1984. Carbohydrate Dramatically Influences Immune Reactivity of Antisera to Viral Glycoproteins. Science **226**:1328-1330.
2. **Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and E. A. Berger**. 1996. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. Science **272**:1955-1958.
3. **Amedee, A. M., N. Lacour, J. L. Gierman, L. N. Martin, J. E. Clements, R. Bohm,Jr., R. M. Harrison, and M. Murphey-Corb**. 1995. Genotypic selection of simian immunodeficiency virus in macaque infants infected transplacentally. Journal of Virology **69**:7982-7990.
4. **Anderson, M. G., D. Hauer, D. P. Sharma, S. V. Joag, O. Narayan, M. C. Zink, and J. E. Clements**. 1993. Analysis of envelope changes acquired by SIVmac239 during neuroadaption in rhesus macaques. Virology **195**:616-626.
5. **Baskin, G. B., L. N. Martin, M. Murphey-Corb, F. S. Hu, D. Kuebler, and B. Davison**. 1995. Distribution of SIV in lymph nodes of serially sacrificed rhesus monkeys. AIDS Research & Human Retroviruses **11**:273-285.
6. **Baskin, G. B., L. N. Martin, S. R. Rangan, B. J. Gormus, M. Murphey-Corb, R. H. Wolf, and K. F. Soike**. 1986. Transmissible lymphoma and simian acquired immunodeficiency syndrome in rhesus monkeys. Journal of the National Cancer Institute **77**:127-139.
7. **Bernstein, B. H., S. P. Tucker, E. Hunter, J. S. Schutzbach, and R. W. Compans**. 1994. Human Immunodeficiency Virus Type 1 Envelope Glycoprotein is Modified by O-linked Oligosaccharides. Journal of Virology **68**:463-468.
8. **Bohm, R. P., Jr., L. N. Martin, B. Davison-Fairburn, G. B. Baskin, and M. Murphey-Corb**. 1993. Neonatal disease induced by SIV infection of the rhesus monkey (*Macaca mulatta*). AIDS Research & Human Retroviruses **9**:1131-1137.
9. **Boom, R., C. J. Sol, M. M. Salimans, C. L. Jansen, P. M. Wertheim-van Dillen, and J. van der Noordaa**. 1990. Rapid and simple method for purification of nucleic acids. Journal of Clinical Microbiology **28**:495-503.
10. **Burns, D. P. and R. C. Desrosiers**. 1994. Envelope sequence variation, neutralizing antibodies, and primate lentivirus persistence. Current Topics in Microbiology & Immunology **188**:185-219.
11. **Burns, D. P. W., C. Collignon, and R. C. Desrosiers**. 1993. Simian immunodeficiency virus mutants resistant to serum neutralization arise during persistent infection of rhesus monkeys. Journal of Virology **67**:4104-4113.
12. **Burns, D. P. W. and R. C. Desrosiers**. 1991. Selection of genetic variants of simian imuunodeficiency virus in persistently infected rhesus monkeys. Journal of Virology **65**:1843-1854.
13. **Campbell, B. J. and V. M. Hirsch**. 1994. Extensive envelope heterogeneity of simian immunodeficiency virus in tissues from infected macaques. Journal of Virology **68**:3129-3137.
14. **Chackerian, B., W. R. Morton, and J. Overbaugh**. 1994. Persistence of simian immunodeficiency virus Mne

variants upon transmission. Journal of Virology 68:4080-4085.
15. **Chackerian, B., L. M. Rudensey, and J. Overbaugh**. 1997. Specific N-linked and O-linked Glycosylation Modifications in the Envelope V1 Domain of SIV variants that Evolve in the Host alter Recognition by Neutralizing Antibodies. Journal of Virology 71:7719-7727.
16. **Chakrabarti, L., P. Isola, M. C. Cumont, M. A. Claessens-Maire, M. Hurtrel, L. Montagnier, and B. Hurtrel**. 1994. Early stages of simian immunodeficiency virus infection in lymph nodes. Evidence for high viral load and successive populations of target cells. American Journal of Pathology 144:1226-1237.
17. **Chakrabarti, L. A., S. R. Lewin, L. Zhang, A. Gettie, A. Luckay, L. N. Martin, E. Skulsky, D. D. Ho, C. Cheng-Mayer, and P. A. Marx**. 2000. Normal T-Cell Turnover in Sooty Mangabeys Harboring Active Simian Immunodeficiency Virus Infection. Journal of Virology 74:1209-1223.
18. **Chen, Z., A. Gettie, D. D. Ho, and P. A. Marx**. 1998. Primary SIVsm isolates use the CCR5 coreceptor from sooty mangabeys naturally infected in west Africa: a comparison of coreceptor usage of primary SIVsm, HIV-2, and SIVmac. Virology 246:113-124.
19. **Chen, Z., P. Zhou, D. D. Ho, N. R. Landau, and P. A. Marx**. 1997. Genetically Divergent Strains of Simian Immunodeficiency Virus Use CCR5 as a Coreceptor for Entry. Journal of Virology 71:2705-2714.
20. **Choi, W. S., C. Collignon, C. Thiriart, D. P. W. Burns, E. J. Stott, and K. A. Kent**. 1994. Effects of Natural Sequence Variation on Recognition by Monoclonal Antibodies That Neutralize Simian Immunodeficiency Virus Infectivity. Journal of Virology 68:5395-5402.
21. **Cocchi, F., A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, and P. Lusso**. 1995. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. Science 270:1811-1815.
22. **Connor, R. I., K. E. Sheridan, D. Ceradini, S. Choe, and N. R. Landau**. 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1-Infected Individuals. Journal of Experimental Medicine 185:621-628.
23. **Courgnaud, V., F. Laure, P. N. Fultz, L. Montagnier, C. Brechot, and P. Sonigo**. 1992. Genetic differences accounting for evolution and pathogenicity of simian immunodeficiency virus from a sooty mangabey monkey after cross-species transmission to a pig-tailed macaque. Journal of Virology 66:414-419.
24. **Daniel, M. D., N. L. Letvin, N. W. King, M. Kannagi, P. K. Sehgal, R. D. Hunt, P. J. Kanki, M. Essex, and R. C. Desrosiers**. 1985. Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. Science 228:1201-1204.
25. **De Jong, J.-J., A. De Ronde, W. Keulen, M. Tersmette, and J. Goudsmit**. 1992. Minimal Requirements for the human immunodeficiency virus type 1 V3 domain to support the syncitium-inducing phenotype:analysis by single amino acid substitution. Journal of Virology 66:6777-6780.
26. **de Roda Husman, A. M., R. P. van Rij, H. Blaak, S. Broersen, and H. Schuitemaker**. 1999. Adaptation to Promiscuous Usage of Chemokine Receptors Is Not a Prerequisite for Human Immunodeficiency Virus Type 1 Disease Progression. J. Infect. Dis. 180:1106-1115.
27. **Deng, H. K., D. Unutmaz, V. N. KewalRamani, and D. R. Littman**. 1997. Expression cloning of new receptors used by simian and human immunodeficiency viruses. Nature 388:296-300.
28. **Douglas, N. W., G. H. Munro, and R. S. Daniels**. 1997. HIV/SIV glycoproteins: structure-function relationships. Journal of Molecular Biology 273:122-149.
29. **Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton**. 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature 381:667-673.
30. **Edinger, A. L., A. M. Amedee, K. Miller, B. J. Doranz, M. Endres, M. Sharron, M. Samson, Z. H. Lu, J. E. Clements, M. Murphey-Corb, M. Parmentier, S. C. Peiper, C. C. Broder, and R. W. Doms**. 1997. Differential utilization of CCR5 by macrophage and T cell tropic simian immunodeficiency virus strains. Proceedings of the National Academy of Sciences of the United States of America 94:4005-4010.
31. **Edinger, A. L., T. L. Hoffman, M. Sharron, B. Lee, Y. Yi, W. Choe, D. L. Kolson, B. Mitrovic, Y. Zhou, D. Faulds, R. G. Collman, J. Hesselgesser, R. Horuk, and R. W. Doms**. 1998. An orphan seven-transmembrane domain receptor expressed widely in the brain functions as a coreceptor for human immunodeficiency virus type 1 and simian immunodeficiency virus. Journal of Virology 72:7934-7940.
32. **Eigen, M.** 1993. Viral quasispecies. Scientific American 269:42-49.
33. **Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger**. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven- transmembrane, G protein-coupled receptor. Science 272:872-877.
34. **Flaherty, M. T., D. A. Hauer, J. L. Mankowski, M. C. Zink, and J. E. Clements**. 1997. Molecular and biological characterization of a neurovirulent molecular clone of simian immunodeficiency virus. Journal of Virology 71:5790-5798.
35. **Fomsgaard, A., P. R. Johnson, W. T. London, and V. M. Hirsch**. 1993. Genetic variation of the SIVagm transmembrane glycoprotein in naturally and experimentally infected primates. AIDS 7:1041-1047.
36. **Fouchier, R. A. M., M. Groenink, N. A. Kootstra, M. Tersmette, H. G. Huisman, F. Miedema, and H. Schuitemaker**. 1996. Phenotype-associated sequence in the third variable domainof the human immunodeficiency virus type 1 gp120

molecule. Journal of Virology **66**:3183-3187.
37. **Fultz, P. N., H. M. McClure, D. C. Anderson, R. B. Swenson, R. Anand, and A. Srinivasan**. 1986. Isolation of a T-lymphotropic retrovirus from naturally infected sooty mangabey monkeys (*Cercocebus atys*). Proc. Natl. Acad. Sci. U. S. A. **83**:5286-5290.
38. **Gairin, J. E. and M. B. Oldstone**. 1993. Virus and cytotoxic T lymphocytes: crucial role of viral peptide secondary structure in major histocompatibility complex class I interactions. Journal of Virology **67**:2903-2907.
39. **Garrity, R. R., G. Rimmelzwaan, A. Minassian, W. P. Tsai, G. Lin, J. Goudsmit, J. J. de Jong, and P. L. Nara**. 1997. Refocusing neutralizing antibody response by targeted dampening of an immunodominant epitope. Journal of Immunology **159**:279-289.
40. **Ghiara, J. B., E. A. Stura, R. L. Stanfield, A. T. Profy, and I. A. Wilson**. 1994. Crystal Structure of the Principal Neutralization Site of HIV-1. Science **264**:82-85.
41. **Higgins, D. G. and P. M. Sharp**. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73**:237-244.
42. **Hirsch, V. M.** 1999. Evolution of the fittest ends in tragedy. Nature Medicine **5**:488-489.
43. **Hirsch, V. M., J. E. Martin, G. Dapolito, W. R. Elkins, W. T. London, S. Goldstein, and P. R. Johnson**. 1994. Spontaneous substitutions in the vicinity of the V3 analog affect cell tropism and pathogenicity of simian immunodeficiency virus. Journal of Virology **68**:2649-2661.
44. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson**. 1989. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature **339**:389-392.
45. **Hirsch, V. M., M. E. Sharkey, C. R. Brown, B. Brichacek, S. Goldstein, J. Wakefield, R. Byrum, W. R. Elkins, B. H. Hahn, J. D. Lifson, and M. Stevenson**. 1998. Vpx is required for dissemination and pathogenesis of SIV(SM) PBj: evidence of macrophage-dependent viral amplification. Nature Medicine **4**:1401-1408.
46. **Holterman, L., R. Dubbes, J. I. Mullins, G. H. Learn, H. Niphuis, W. Koornstra, G. Koopman, E. Kuhn, A. Wade-Evans, B. Rosenwirth, J. Haaijman, and J. L. Heeney**. 1999. Distinct Characteristics of a Pathogenic Molecular Clone of Simian Immunodeficiency Virus Derived from Serum. submitted
47. **Holterman, L., H. Niphuis, P. ten Haaft, J. Goudsmit, G. B. Baskin, and J. L. Heeney**. 1999. Specific Passage of Simian Immunodeficiency Virus from End-Stage Disease Results in Accelerated Progression to AIDS in Rhesus Macaques. Journal of General Virology 3089-3097.
48. **Javaherian, K., A. J. Langlois, and G. J. LaRosa**. 1993. SIV neutralization epitopes. Chemical Immunology 78-90.
49. **Javaherian, K., A. J. Langlois, D. C. Montefiori, K. A. Kent, K. A. Ryan, P. D. Wyman, J. Stott, D. P. Bolognesi, M. Murphey-Corb, and G. J. LaRosa**. 1994. Studies of the conformation-dependent neutralizing epitopes of simian immunodeficiency virus envelope protein. Journal of Virology **68**:2624-2631.
50. **Javaherian, K., A. J. Langlois, S. Schmidt, M. Kaufmann, N. Cates, J. P. Langedijk, R. H. Meloen, R. C. Desrosiers, D. P. Burns, and D. P. Bolognesi**. 1992. The principal neutralization determinant of simian immunodeficiency virus differs from that of human immunodeficiency virus type 1. Proceedings of the National Academy of Sciences of the United States of America **89**:1418-1422.
51. **Joag, S. V., E. B. Stephens, R. J. Adams, L. Foresman, and O. Narayan**. 1994. Pathogenesis of SIVmac infection in Chinese and Indian rhesus macaques: effects of splenectomy on virus burden. Virology **200**:436-446.
52. **Khabbaz, R. F., T. Rowe, M. Murphey-Corb, W. M. Heneine, C. A. Schable, J. R. George, C. P. Pau, B. S. Parekh, M. D. Lairmore, and J. W. Curran**. 1992. Simian immunodeficiency virus needlestick accident in a laboratory worker. Lancet **340**:271-273.
53. **Kimata, J. T., L. Kuller, D. B. Anderson, P. Dailey, and J. Overbaugh**. 1999. Emerging cytopathic and antigenic simian immunodeficiency virus variants influence AIDS progression. Nature Medicine **5**:535-541.
54. **Kimata, J. T. and J. Overbaugh**. 1998. The Cytopathicity of a Simian Immunodeficiency Virus Mne Mutant is determined by Mutations in Gag and Env. Journal of Virology **(in press)**:
55. **Kinsey, N. E., M. G. Anderson, T. J. Unangst, S. V. Joag, O. Narayan, M. C. Zink, and J. E. Clements**. 1996. Antigenic variation of SIV: mutations in V4 alter the neutralization profile. Virology **221**:14-21.
56. **Kirchhoff, F. and R. C. Desrosiers**. 1993. A PCR-derived library of random point mutations within the V3 region of simian immunodeficiency virus. PCR Methods & Applications **2**:301-304.
57. **Kirchoff, F., K. Mori, and R. C. Desrosiers**. 1994. The "V3" Domain Is a Determinant of Simian Immunodeficiency Virus Cell Tropism. Journal of Virology **68**:3682-3692.
58. **Kodama, T., K. Mori, T. Kawahara, D. J. Ringler, and R. C. Desrosiers**. 1993. Analysis of simian immunodeficiency virus sequence variation in tissues of rhesus macaques with simian AIDS. Journal of Virology **67**:6522-6534.
59. **Kumar, S., K. Tamura, and M. Nei**. 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Computer Applications in the Biosciences **10**:189-191.
60. **La Rosa, G. J., J. P. Davide, K. Weinhold, J. A. Waterbury, A. T. Profy, J. A. Lewis, A. J. Langlois, G. R. Dreesman, R. N. Boswell, P. Shadduck, L. H. Holley, M. Karplus, D. P. Bolognesi, T. J. Mattews, E. A. Emini,**

and S. D. Putney. 1990. Conserved Sequence and Structural Elements in the HIV-1 Principal Neutralizing Determinant. Science **253**:932-935.
61. **LaBranche, C. C., M. M. Sauter, B. S. Haggarty, P. J. Vance, J. Romano, T. K. Hart, P. J. Bugelski, and J. A. Hoxie**. 1994. Biological, molecular, and structural analysis of a cytopathic variant from a molecularly cloned simian immunodeficiency virus. Journal of Virology **68**:5509-5522.
62. **Lane, T. E., M. J. Buchmeier, D. D. Watry, D. B. Jakubowski, and H. S. Fox**. 1995. Serial passage of microglial SIV results in selection of homogeneous env quasispecies in the brain. Virology **212**:458-465.
63. **Lukashov, V. V., C. L. Kuiken, and J. Goudsmit**. 1995. Intrahost human immunodeficiency virus type 1 evolution is related to length of the immunocompetent period. Journal of Virology **69**:6911-6916.
64. **Martin Amedee, A., N. Lacour, L. N. Martin, J. E. Clements, R. B. Bohm,Jr., B. Davison, R. Harrison, and M. Murphey-Corb**. 1996. Genotypic analysis of infant macaques infected transplacentally and orally. Journal of Medical Primatology **25**:225-235.
65. **Miller, M. A., M. Murphey-Corb, and R. C. Montelaro**. 1992. Identification of broadly reactive continuous antigenic determinants of simian immunodeficiency virus glycoproteins. AIDS Research & Human Retroviruses **8**:1153-1164.
66. **Mooij, P., W. M. J. M. Bogers, H. Niphuis, W. Koornstra, R. Dubbes, and J. L. Heeney**. 1999. Selective Croeceptor Use of SIV is Sufficient for, but not a Correlate of Accelerated Progression to AIDS in Rhesus Macaques. submitted
67. **Moore, J. P., F. E. McCutchan, S. W. Poon, J. Mascola, J. Liu, Y. Cao, and D. D. Ho**. 1994. Exploration of antigenic variation in gp120 from clades A through F of human immunodeficiency virus type 1 by using monoclonal antibodies. Journal of Virology **68**:8350-8364.
68. **Mori, K., D. J. Ringler, T. Kodama, and R. C. Desrosiers**. 1992. Complex determinants of macrophage tropism in env of simian immunodeficiency virus. Journal of Virology **66**:2067-2075.
69. **Morrison, H. G., F. Kirchhoff, and R. C. Desrosiers**. 1995. Effects of mutations in constant regions 3 and 4 of envelope of simian immunodeficiency virus. Virology **210**:448-455.
70. **Murphey-Corb, M., L. N. Martin, S. R. Rangan, G. B. Baskin, B. J. Gormus, R. H. Wolf, W. A. Andes, M. West, and R. C. Montelaro**. 1986. Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys. Nature **321**:435-437.
71. **Novembre, F. J., P. R. Johnson, M. G. Lewis, D. C. Anderson, S. Klumpp, H. M. McClure, and V. M. Hirsch**. 1993. Multiple viral determinants contribute to pathogenicity of the acutely lethal simian immunodeficiency virus SIVsmmPBj variant. Journal of Virology **67**:2466-2474.
72. **Nowak, M. A., R. M. Anderson, A. R. McLean, T. F. Wolfs, J. Goudsmit, and R. M. May**. 1991. Antigenic diversity thresholds and the development of AIDS. Science **254**:963-969.
73. **Overbaugh, J. and L. M. Rudensey**. 1992. Alterations in potential sites for glycosylation predominate during evolution of the simian immunodeficiency virus envelope gene in macaques. Journal of Virology **66**:5937-5948.
74. **Overbaugh, J., L. M. Rudensey, M. D. Papenhausen, R. E. Benveniste, and W. R. Morton**. 1993. Variation in simian immunodeficiency virus env is confined to V1 and V4 during progression to simian AIDS. Journal of Virology **65**:7025-7031.
75. **Owen, S. M., D. Ellenberger, M. Rayfield, S. Wiktor, P. Michel, M. H. Grieco, F. Gao, B. H. Hahn, and R. B. Lal**. 1998. Genetically divergent strains of human immunodeficiency virus type 2 use multiple coreceptors for viral entry. Journal of Virology **72**:5425-5432.
76. **Peeters, M., W. Janssens, K. Fransen, J. Brandful, L. Heyndrickx, K. Koffi, E. Delaporte, P. Piot, G. M. Gershy-Damet, and G. van der Groen**. 1994. Isolation of simian immunodeficiency viruses from two sooty mangabeys in Cote d'Ivoire: virological and genetic characterization and relationship to other HIV type 2 and SIVsm/mac strains. AIDS Research & Human Retroviruses **10**:1289-1294.
77. **Pelletier, E., W. Saurin, R. Cheynier, N. L. Letvin, and S. Wain-Hobson**. 1995. The tempo and mode of SIV quasispecies development in vivo calls for massive viral replication and clearance. Virology **208**:644-652.
78. **Rudensey, L. M., J. T. Kimata, R. E. Benveniste, and J. Overbaugh**. 1995. Progression to AIDS in macaques is associated with changes in the replication, tropism, and cytopathic properties of the simian immunodeficiency virus variant population. Virology **207**:528-542.
79. **Rudensey, L. M., J. T. Kimata, E. M. Long, B. Chackerian, and J. Overbaugh**. 1998. Changes in the extracellular envelope glycoprotein of variants that evolve during the course of SIVMne infection that affect neutralizing antibody recognition, syncytia formation and macrophage tropism, but not replication, cytopathicity or CCR-5 coreceptor recognition. Journal of Virology **(in press)**:
80. **Sakuragi, S., J. Sakuragi, and A. Adachi**. 1995. Both SU and TM env proteins are responsible for monkey cell tropism of simian immunodeficiency virus SIV mac. Archives of Virology **140**:2255-2260.
81. **Schuitemaker, H., M. Koot, N. A. Kootstra, M. W. Dercksen, R. E. Y. de Goede, R. P. van Steenwijk, J. M. Lange, J. K. Schattenkerk, F. Miedema, and M. Tersmette**. 1992. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. Journal of Virology **66**:1354-1360.

82. **Schwiebert, R. and P. N. Fultz**. 1994. Immune activation and viral burden in acute disease induced by simian immunodeficiency virus SIVsmmPBj14: correlation between in vitro and in vivo events. Journal of Virology **68**:5538-5547.
83. **Sellner, L. N., R. J. Coelen, and J. S. Mackenzie**. 1992. A one-tube, one manipulation RT-PCR reaction for detection of Ross River virus. Journal of Virological Methods **40**:255-263.
84. **Sol, N., F. Ferchal, J. Braun, O. Pleskoff, C. Treboute, A. Ansart, and M. Alizon**. 1997. Usage of the Coreceptors CCR-5, CCR-3, and CXCR-4 by Primary and Cell Line-Adapted Human Immunodeficiency Virus Type 2. Journal of Virology **71**:8237-8244.
85. **Spira, A. I., P. A. Marx, B. K. Patterson, J. Mahoney, R. A. Koup, S. M. Wolinsky, and D. D. Ho**. 1996. Cellular targets of infection and route of viral dissemination after an intravaginal inoculation of simian immunodeficiency virus into rhesus macaques. Journal of Experimental Medicine **183**:215-225.
86. **Stephens, E. B., D. Galbreath, Z. Q. Liu, M. Sahni, Z. Li, R. Lamb-Wharton, L. Foresman, S. V. Joag, and O. Narayan**. 1997. Significance of macrophage tropism of SIV in the macaque model of HIV disease. Journal of Leukocyte Biology **62**:12-19.
87. **Stephens, E. B., Z. Q. Liu, G. W. Zhu, I. Adany, S. V. Joag, L. Foresman, N. E. Berman, and O. Narayan**. 1995. Lymphocyte-tropic simian immunodeficiency virus causes persistent infection in the brains of rhesus monkeys. Virology **213**:600-614.
88. **Tao, B. and P. N. Fultz**. 1995. Molecular and biological analyses of quasispecies during evolution of a virulent simian immunodeficiency virus, SIVsmmPBj14. Journal of Virology **69**:2031-2037.
89. **Valli, P. J. S., V. V. Lukashov, and J. Goudsmit**. 1998. Shortening of the symptom-free period in rhesus macaques related to decreasing non-synymous variation in the env variable regions of SIVsm during passage. Journal of Virology **72**:7494-7500.
90. **Watanabe, N., S. N. McAdam, J. E. Boyson, M. S. Piekarczyk, Y. Yasutomi, D. I. Watkins, and N. L. Letvin**. 1994. A simian immunodeficiency virus envelope V3 cytotoxic T-lymphocyte epitope in rhesus monkeys and its restricting major histocompatibility complex class I molecule Mamu-A*02. Journal of Virology **68**:6690-6696.
91. **Zhang, J. Y., L. N. Martin, E. A. Watson, R. C. Montelaro, M. West, L. Epstein, and M. Murphey-Corb**. 1988. Simian immunodeficiency virus/delta-induced immunodeficiency disease in rhesus monkeys: relation of antibody response and antigenemia. Journal of Infectious Diseases **158**:1277-1286.
92. **Zhang, L., T. He, Y. Huang, Z. Chen, Y. Guo, S. Wu, K. J. Kunstman, R. C. Brown, J. P. Phair, A. U. Neumann, D. D. Ho, and S. M. Wolinsky**. 1998. Chemokine coreceptor usage by diverse primary isolates of human immunodeficiency virus type 1. Journal of Virology **72**:9307-9312.

Chapter 4

Shortening of the Symptom-Free Period in Rhesus Macaques Is Associated with Decreasing Nonsynonymous Variation in the *env* Variable Regions of Simian Immunodeficiency Virus SIVsm during Passage

JOURNAL OF VIROLOGY, Sept. 1998, p. 7494–7500 Vol. 72, No. 9
0022-538X/98/$04.00+0

Shortening of the Symptom-Free Period in Rhesus Macaques Is Associated with Decreasing Nonsynonymous Variation in the *env* Variable Regions of Simian Immunodeficiency Virus SIVsm during Passage

P. J. SPENCER VALLI,[1]* VLADIMIR V. LUKASHOV,[1] JONATHAN L. HEENEY,[2] AND JAAP GOUDSMIT[1]

Department of Human Retrovirology, Academic Medical Centre, Amsterdam,[1] and Biomedical Primate Research Centre, Rijswijk,[2] The Netherlands

Received 2 February 1998/Accepted 4 June 1998

During six blood passages of simian immunodeficiency virus SIVsm in rhesus macaques, the asymptomatic period shortened from 18 months to 1 month. To study SIVsm envelope gene (*env*) evolution during passage in rhesus macaques, the C1 to CD4 binding regions of multiple clones were sequenced at seroconversion and again at death. The *env* variation found during adaptation was almost completely confined to the variable regions. Intrasample sequence variation among clones at seroconversion was lower than the variation among clones at death. Intrasample variation among clones from a single time point as well as intersample variation decreased during the passage. In the variable regions, the mean number of intrasample nonsynonymous nucleotide substitutions decreased from the first passage ($5.26 \times 10^{-2} \pm 0.6 \times 10^{-2}$ per site) to the fifth passage ($2.24 \times 10^{-2} \pm 0.4 \times 10^{-2}$ per site), whereas in the constant regions, the mean number of intrasample nonsynonymous nucleotide substitutions differed less between the first and fifth passages ($1.14 \times 10^{-2} \pm 0.27 \times 10^{-2}$ and $0.80 \times 10^{-2} \pm 0.24 \times 10^{-2}$ per site). Shortening of the asymptomatic period coincided with a rise in the *Ks*/*Ka* ratio (ratio between the number of synonymous [*Ks*] and the number of nonsynonymous [*Ka*] substitutions) from 1.080 in passage one to 1.428 in passage five and mimicked the difference seen in the intrahost evolution between asymptomatic and fast-progressing individuals infected with human immunodeficiency virus type 1. The distribution of nonsynonymous substitutions was biphasic, with most of the adaptation of *env* variable regions occurring in the first three passages. This phase, in which the symptom-free period fell to 4 months, was followed by a plateau phase of apparently reduced adaptation. Analysis of codon usage revealed decreased codon redundancy in the variable regions. Overall, the results suggested a biphasic pattern of adaptation and evolution, with extremely rapid selection in the first three passages followed by an equilibrium or stabilization of the variation between *env* clones at different time points in passages four to six.

The discovery of lymphomas in macaques previously housed with sooty mangabeys or African green monkeys (7, 16, 18, 22) led to the isolation of the first simian immunodeficiency virus (SIV) isolates and characterization of the wasting diseases that they caused in macaques. Later, it was found that African feral monkeys were commonly SIV infected (22%) (14, 24, 30) and that SIV infection was endemic in sooty mangabeys housed at some centers. At the Tulane Regional Primate Research Center, tissue-derived inoculum from a sooty mangabey was used to inoculate several macaques intravenously; one died of a wasting syndrome 18 months later (rhesus macaque B670) (7). The transmission of this mangabey virus (SIVsm) to Asian macaques resulted in an infection characterized by a loss of $CD4^+$ T cells, persistent serum antigenemia, and trapping of virions in the follicular dendritic cell foot processes (6), all hallmarks of human immunodeficiency virus (HIV) infections.

Studies with molecular clones have shown that single nucleotide substitutions in *env* of HIV and SIV can cause changes in the biological phenotype, neutralizing antibody escape, and growth kinetics (1, 15, 25). The *env* nucleotide substitutions seen during HIV infection are concentrated in the variable regions. The predominance of nonsynonymous over synonymous substitutions is believed to be due to immune pressure (5, 11) on various viral proteins as well as on *env*. During cross-species transmission of a lentivirus, adaptation occurs for accommodation to the newly encountered nonhost environment. This process allows the viral proteins needed in the infective processes to adapt to the cellular composition of the new host.

The sequencing of multiple *env* clones at a particular time point in infection gives an approximation of the quasispecies present in the blood at that time. A comparison of the variation within these sequences gives an idea of the relative intrasample variation taking place at that time in the gene sequenced (*env*). An inverse correlation between virus variation and length of the immunocompetent period has been shown for asymptomatic carriers of HIV and for individuals progressing to AIDS (31). Progression to AIDS following HIV infection is known to be load dependent, with the more rapidly replicating and syncytium-inducing phenotypes being the most efficient in the pathway of events leading to immunodeficiency (26, 39).

To study the relationship between viral variation and length of the asymptomatic period, we passaged SIVsm in Asian macaques. The cross-species transmission was carried out with the Delta B670 SIV strain (7) as the primary inoculum followed by four serial intravenous inoculations with peripheral blood mononuclear cells (PBMC) taken from animals at the symptomatic stage of infection. The viral quasispecies present in the monkeys at seroconversion and death were sampled,

* Corresponding author. Mailing address: Department of Human Retrovirology, Academic Medical Centre, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands. Phone: (31-20) 5664853. Fax: (31-20) 6916531. E-mail: p.j.valli@amc.uva.nl.

and multiple *env* clones were sequenced. We report here on the shortening of the asymptomatic period from 18 months to a few weeks and the concomitant reduction of intrasample and intersample variations. Finally, we observed that the rate of nonsynonymous nucleotide substitutions during *env* adaptation of SIVsm in rhesus macaques was variable and that the changes seen occurred almost entirely in the *env* variable regions, where limited codon usage was found.

MATERIALS AND METHODS

Virus. The Delta B670 SIV strain (3, 6, 7, 9, 31a, 33, 41) is an SIV originally found in a sooty mangabey presenting with a cutaneous lepromatous lesion. When tissue from this animal was used to infect rhesus macaques, it closely reproduced AIDS in humans (33). The virus inoculum was proven to be free of type D retrovirus, which can cause such symptoms as well (28).

Passage. A full history of the passage is currently being prepared (unpublished data). Briefly, six Asian rhesus macaques, all 2 years of age, were used for the purpose of experimental infection with an SIVsm strain. The first virus sampled (P1) was the Delta B670 SIV stock from the macaque inoculated with the sooty mangabey virus (33). The second monkey (P2) was infected intravenously with 5×10^2 infectious doses of the Delta B670 SIV strain. The monkeys then were intravenously inoculated in a serial fashion with 2×10^6 PBMC taken at the symptomatic stage. Passage five symptomatic-stage PBMC were used to infect two monkeys (passages six A and six B). The time to death postinfection (tdpi) and moment of sampling (ms) for the animals were as follows: P1—tdpi, 18 months, and ms, 18 months; P2—tdpi, 12 months, and ms, 3 and 7 months; P3—tdpi, 9 months, and ms, 2 and 4.1 months; P4—tdpi, 4 months, and ms, 1 and 1.8 months; P5—tdpi, 2 months, and ms, not done, as no sample was available; P6A—tdpi, 2 months, and ms, 2 months; and P6B—tdpi, 2 months, and ms, 2 months. Animals were euthanized upon evidence of undue suffering. PBMC used for serial passages were not cultured or cocultivated.

RNA isolation and RT-PCR. Viral RNA was harvested with silica in the presence of a chaotropic agent (10) from the sera of experimentally infected Asian macaques and was used as a template in a reverse transcriptase (RT) PCR (RT-PCR). Viral RNA was isolated from 20-μl volumes of sera, resuspended in 20 μl of RNasin-containing H_2O (1 U/μl), and used in an RT reaction consisting of a mixture of 5 μl of viral RNA, 250 μM each deoxynucleoside triphosphate (dNTP), 2 ng of 3′ RT-PCR primer (SIV4Not1: TTATATGCGGCCGCCTAC TTTGTGCCACGTGTTG) per μl, 2.5 mM Mg^{2+}, 1 U of RNasin (Promega) per μl, 10 U of Super Script I (Gibco-BRL), and 1× reaction buffer (37) in a 20-μl volume. The components were assembled at 37°C and incubated at that temperature for 90 min. The PCR mixture consisted of 250 μM each dNTP, 2 ng of 3′ and 5′ (SIV1HIII: GTAGACAAGCTTGGGATAATACAGTCACAGAAC) PCR primers per μl, 1× reaction buffer, 2.25 mM Mg^{2+}, and 1.5 U of *Taq* polymerase (Perkin-Elmer Cetus) in a final volume of 100 μl including 5 μl of RT reaction mixture. The PCR mixture was overlaid with paraffin and heated to 95°C for 5 min followed by 35 cycles of 1 min at 95°C, 1 min at 55°C followed by 1 min at 72°C, and finally 10 min at 72°C in a Perkin-Elmer Cetus DNA Thermocycler. RT reactions and PCRs were carried out in duplicate for each sample to prevent mispriming and to preserve the fidelity of the virus genotypes sampled.

Samples were combined after PCR and size selected on 0.8% agarose gels followed by excision of the 1,151-bp band. The excised band was isolated from the gel slice and digested with *Not*I and *Hin*dIII, followed by agarose gel and silica gel fragment isolation (10). The size-selected, digested, purified RT-PCR product was ligated overnight into plasmid pSP64 (Promega) containing a *Not*I site. The ligated product was electroporated into electrocompetent *Escherichia coli* C600, and plasmid DNA from sequencing was isolated with Qiagen columns.

Sequencing and analysis. Double-stranded plasmid DNA was sequenced with custom labelled dye primers (ABI, Foster City, Calif.) by use of an ABI model 373A automated sequencer and version 1.2.0 software. Clones were assembled and aligned with the Sequence Navigator program (ABI). Nucleotide sequences were aligned with Sequence Navigator and Clustal V (21), with final adjustments being carried out visually. All positions with an alignment gap in at least one sequence were excluded from any pairwise sequence comparisons. *p* distances, defined as the number of synonymous or nonsynonymous substitutions divided by the total number of synonymous or nonsynonymous sites (34), were used to measure the relative genetic variation between clones. Synonymous and nonsynonymous nucleotide *p* distances (*Ks* and *Ka*, respectively) were calculated with the MEGA program (27). Intrasample calculations are the result of comparisons of clones from the same time point or quasispecies (within a sample); intersample calculations are the result of comparisons of clones one by one from two different time points or quasispecies (between two samples). Samples were named for their passage (P) position (from 1 to 6) and for the time of sampling, either at seroconversion (S) or death (D); e.g., P2S is the passage two seroconversion sample. P6A and P6B were considered one sample (P6) in the data calculations since there was no significant difference between them.

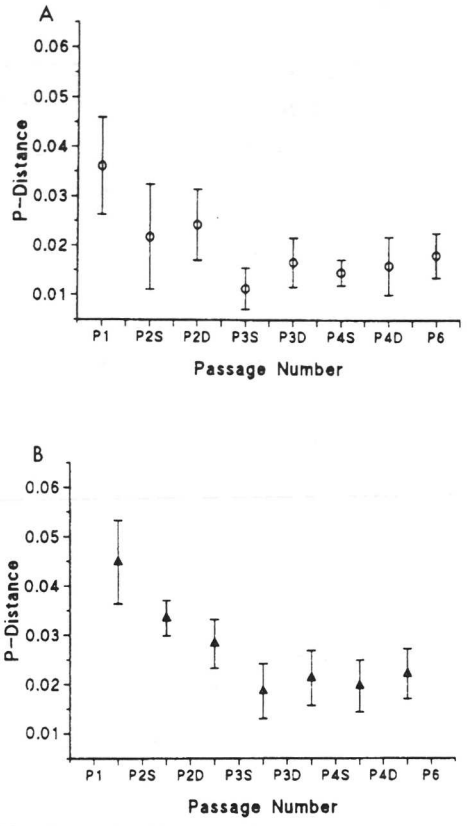


FIG. 1. Intrasample and intersample variations. Intrasample (A) and intersample (B) *p* distances are shown. Groups of five clones per sample were compared to each other (intrasample) or to the five clones in another sample (intersample). The resulting numbers are plotted against the position of the sample along the passage, from one to six, at seroconversion (S) and death (D). Samples without letters were at death. The symbols (○ and △) denote the means of all values, and the variance is shown by the vertical bars.

RESULTS

Intrasample genetic variation at seroconversion was lower than that at death. Declining genetic variation, as measured by *p* distances, was observed during the first three passages, coincident with a decrease in the sequence variation within the quasispecies (Fig. 1B). The values for the last three passages were not significantly different (0.0155 ± 0.0090, 0.0140 ± 0.0065, and 0.0154 ± 0.0092) and indicated that further adaptation to the host did not occur. The intrasample variation was higher at death than at seroconversion in all passages (Fig. 1A). The intersample variation decreased gradually from 0.0423 ± 0.015 to 0.0185 ± 0.0110 during the first three passages and then stabilized. The number of intrasample variations was one half the number of intersample variations, suggesting continuing replacement of genotypes (high interpassage *p* distances). Although the intersample variations decreased gradually during the first three passages, they remained higher than the intrasample variations.

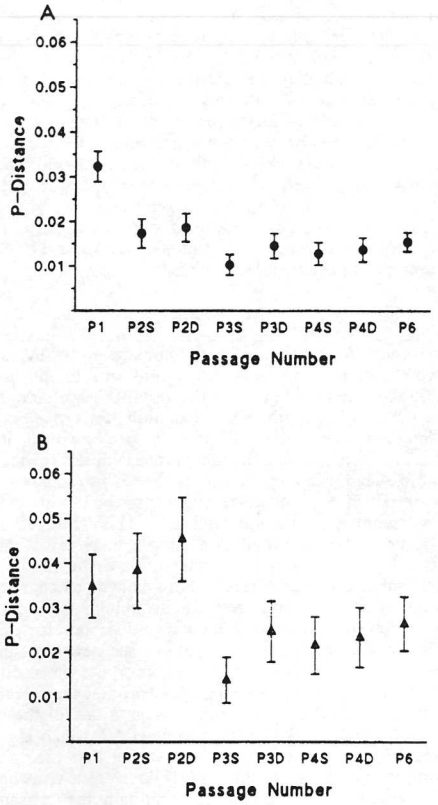


FIG. 2. Intrasample synonymous and nonsynonymous variations. Intrasample nonsynonymous (A) and synonymous (B) *p* differences are shown. The five clones of a sample were compared to one another individually, and the synonymous and nonsynonymous *p* distances were calculated. The symbols (○ and △) denote the means of all values, and the variance is shown by the vertical bars.

Nonsynonymous variations decreased during the first three passages and subsequently stabilized. Synonymous and nonsynonymous substitutions in passage one were almost equal in number, suggestive of a heterogeneous founder population adapting during the introduction of SIVsm to rhesus macaques (Fig. 2A and B). The intrasample nonsynonymous variation decreased threefold in the first three passages (from 0.0325 ± 0.0065 to 0.0110 ± 0.0004). In passages three to six, the synonymous variation was twice the nonsynonymous variation. The intrasample variation was lower at seroconversion than at death during the first two passages and to a lesser extent during the last two passages. The lowest values for synonymous and nonsynonymous variations were at seroconversion in the third passage, with the rate of adaptation leveling off in the following passages. Intersample nonsynonymous and synonymous variations were higher than intrasample variations (Fig. 3A and B), reflecting the adaptation that occurred over time during the infection (intrahost) and between the successive passages (interhost).

The *Ks* and *Ka* values for the *env* variable and constant regions followed dissimilar trends. Figure 4 shows the *Ks* and *Ka* values for the variable and constant regions. The *Ka* values for the variable (Fig. 4A) and the constant (Fig. 4B) regions decreased rapidly from P1 to P3S, when the *Ks* values increased. The *Ks* values for the variable and constant regions were similar from P1 to P3S, when the *Ka* values were more than fourfold higher for the variable regions than for the constant regions. At seroconversion in passage three, there was almost no nonsynonymous variation in the constant regions, when the *Ks* values were similar (variable, 0.0144; constant, 0.0133) and the *Ka* values for the variable regions were 13 times the *Ka* values for the constant regions (0.0187 and 0.0014). Variable-region *Ka* values from P3S onward remained fourfold those for the constant regions. The *Ks* values for the variable regions rose while those for the constant regions remained level in the last two passages. The *Ks*/*Ka* ratios for the constant regions were always greater than one, while those for the variable regions were less than one in five of the eight samples (Fig. 4A and B).

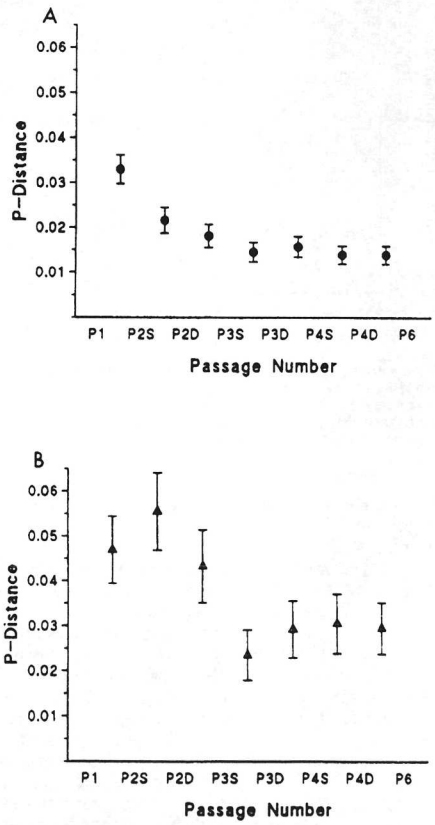


FIG. 3. Intersample synonymous and nonsynonymous variations. Intersample nonsynonymous (A) and synonymous (B) *p* differences are shown. The five clones of a sample were compared to those of the next sample individually, and the synonymous and nonsynonymous *p* distances were calculated. The symbols (○ and △) denote the means of all values, and the variance is shown by the vertical bars.

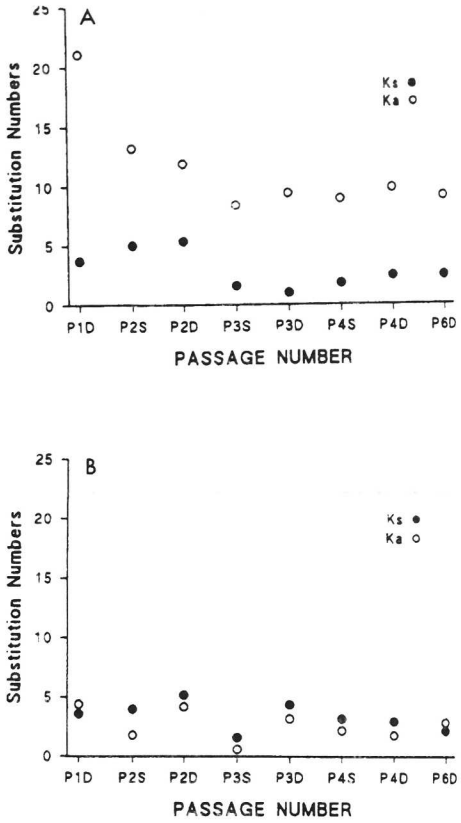


FIG. 4. Intrasample variations of constant and variable regions. *Ks* (synonymous) and *Ka* (nonsynonymous) substitutions per synonymous or nonsynonymous site were calculated by dividing the C1 to CD4 binding region clones into variable (A) and constant (B) regions and comparing these selected regions one at a time among the clones of each sample.

The variable regions of the envelope gene had a decreased codon redundancy. The alterations in *Ks* and *Ka* values led us to examine the nature of the substitutions and changes in nucleotide concentrations of the variable and constant regions of *env* (Fig. 5A to C). Differences found in the envelope sequences of SIVsm were compared to SIVmac251 (29) and to 55 HIV type 1 (HIV-1) subtype B gp120 sequences (Los Alamos Sequence Database). Figure 5 shows that *env* was rich in adenine (A), with increased concentrations in the variable regions compared to the constant regions. The ratios of the four nucleotides remained constant throughout the passages and within the HIV sequences. It has been reported that nucleotide composition affects the evolution of RNA viruses (12) and that individual genes of retroviruses have characteristic base compositions (8). We found that individual regions within retroviral genes, at least *env*, had characteristic base compositions. The A concentration in the variable regions was higher than that in the constant regions, except at the first nucleotide of SIVmac251 and the third nucleotide of HIV-1 subtype B (Fig. 5A and C). The greatest difference seen in nucleotide concentration was in HIV *env*, with its 20% discordance between the variable and constant regions. Nucleotides in position 1 of the SIV strains were almost equal between the two regions. In the SIV strains, the mostly synonymous third nucleotide showed higher A levels in the variable regions; HIV-1 subtype B displayed the opposite pattern.

The effects of the nucleotide concentrations found were examined with regard to codon usage within the two *env* domains, constant and variable. Although the nucleotide concentrations varied between viruses, the codon usage frequencies were similar (Fig. 5D). Of the 61 possible codons (there are 64 possible codons with three encoding stop messages not found within the envelope gene), the number used was 11% lower (on average) in the variable regions.

DISCUSSION

The endemic infection of feral monkeys with SIV has no known associated pathology, apparently due to the historic genetic accommodation (13) of the virus. The host adaptation of the virus, or host-pathogen coevolution, leads to asymptomatic infection of most African primate species with their own host-adapted strains (2, 23). No natural SIV infection has been found in Asian primates, and cross-species transmission (e.g., African SIV in Asian macaques) leads to the development of AIDS presenting the common markers of HIV-1 infection (20, 36). A cross-species population passage may cause initially rapid genomic evolution during adaptation to the new immune environment. Large numbers of nonsynonymous changes are due to the *env* alterations needed to avoid immune attack and/or to maximize the affinity of the viral *env* protein for efficient binding to and infection of the cells of the new host species. Since natural cross-species transmission occurs via blood contact, rather than by mucosal contact, intravenous serial passage was carried out. Parenteral transmission is one of the major routes of infections with HIV-1 and HIV-2, which is closely related to SIVsm. Recently, it was shown that the route of transmission (mucosal versus blood) does affect virus heterogeneity (4, 34a, 38a), both routes resulting in the transmission of a more homogeneous selection of variants. Similar observations were made by Amadee et al. (using the same strain, Delta B670 SIV, as was used here [3, 4]), who also showed that intravenous, oral, and transplacental transmissions limit virus heterogeneity and select for macrophage tropism, as we found for intravenous passage (39a).

The decrease in the *Ka* values after the first three passages suggests that the adaptation of SIVsm to rhesus macaques is rapid and punctuated. The decreases in the times to death postinfection and *env* adaptation follow similar patterns, with the intrasample sequence variation levelling off once the time to death postinfection is less than 4 months (passages four to six A and six B) (Fig. 4A). The decreased intrasample variation remains stable after the third passage, indicating that *env* variation is restricted, most likely because of the adaptive equilibrium reached. The nearly 50% decrease in *env* variation indicates that the total adaptation rate is decreasing and that there is positive or purifying selection of a narrow assortment of *env* genotypes. The lowest *Ks* value is that for P3S, with almost zero variation of the constant regions and a *Ka* of 10.0 for the variable regions, since an almost clonal population was present.

The low variation in seroconversion samples compared to death samples is evidence of a strong founder effect, since this outgrowth of a dominant homogeneous viral population emerges prior to detectable immune responses. This finding is in accord with the variation seen in *env* of HIV-infected individuals progressing to AIDS, in whom homogeneity at sero-

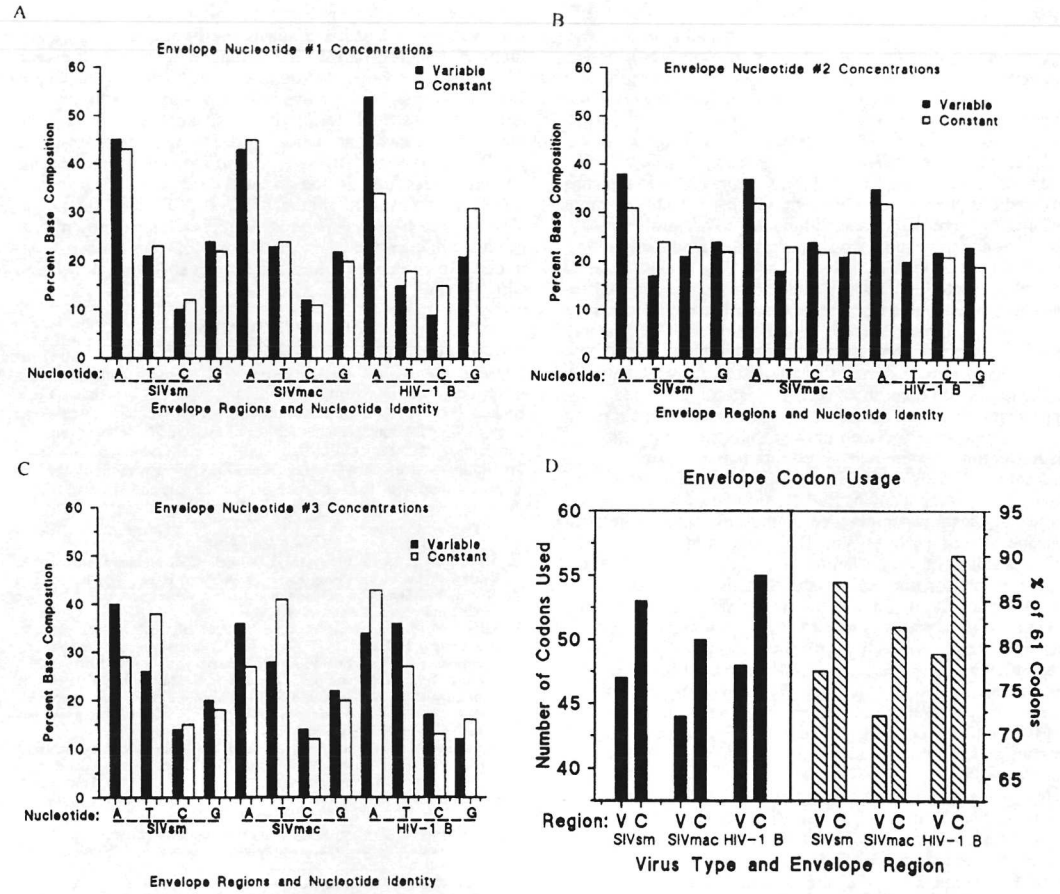


FIG. 5. Envelope nucleotide concentrations and envelope codon usage. Nucleotide concentrations of the variable and constant regions were calculated for the three positions of each codon (A, B, and C) in *env*. The numbers indicate the average for the sums of all clones (SIVmac251 [29] and HIV-1 subtype B data were from the Los Alamos Sequence Database). Variation among clones was less than 0.2%. (D) By use of the MEGA program, the numbers of codons used to represent amino acids in the variable and constant regions were calculated for the above-mentioned sequences. The percentage of 61 codons is the number of codons used divided by 61 (the total number of codons, not including those that represent stop signals).

conversion is higher than that at death (40). The hypothesis that quasispecies homogeneity in HIV infections is caused by single-particle transmissions is not confirmed here, since even after the passage of millions of virus particles, we saw the same patterns of homogeneity at seroconversion and heterogeneity at death. The effect is thus caused by the selective amplification and eventual outgrowth of a dominant viral genome that then diversifies into a quasispecies under the influences of replication competence during adaptation, the immune system, and cell availability.

The fourfold higher *Ka* values for the variable regions than for the constant regions indicate positive selection (32). Similarity in *Ks* values and discontinuity in *Ka* values underscore the difference in the functions of the constant and variable regions and the system governing nucleotide substitutions. If the rate of nucleotide substitutions caused by RT errors is constant over all the nucleotides of a retroviral gene, then there should be equal amounts of variation in all regions. The distinct variations in *Ka* values reported here are presumably the effects of immune pressure, variations in receptor and coreceptor binding sites on the variable regions, and purifying selection on the constant regions. The similarity of the *Ks* values for the constant and variable regions reflects the effects of purifying selection on nucleotide substitutions in conserved regions.

Increasing the concentration of a single nucleotide may affect the number of possible codons that can encode amino acids. Amino acids with more than four representative codons (leucine, serine, and arginine have six) can be represented by codons with various nucleotide concentrations, as opposed to methionine and tryptophan, which have single codons. Thus, the result of an increased concentration of one nucleotide is that fewer codons can encode these amino acids when they are present in the variable regions, where the highest adenine (A)

concentrations are found (Fig. 5D). Although this concentration effect is only possible with these three amino acids, there is a resulting reduction in the number of codons used in encoding the other amino acids as well. As shown in Fig. 5, the redundancy in the codons used in the variable regions is decreased up to 12% compared to that in the constant regions. Reduced codon usage will lower the redundancy of the encoded amino acids. Presumably, the net effect is greater amino acid variation from a given nucleotide substitution in the variable regions than in the constant regions, which are more buffered by increased codon redundancy. This strategy of reduced redundancy would cause more amino acid substitutions in the variable regions as a selective advantage against humoral and cell-mediated immunities while allowing the constant regions to remain stable during the adaptation reported here.

Reaching a plateau in the evolutionary pace in passages three to six suggests that sufficient *env* adaptation has occurred for optimal fitness. Exponential gains in the growth kinetics of RNA viruses (35) occur in vitro during a more prolonged series of passages. The short adaptation phase seen here could result from immune system-driven positive selection, since the previous experiments were performed with tissue culture under no such selection. This short adaptation phase could also be and probably is just as likely caused by selection for viruses with greater replication competence in the new species, and not immune selection alone. With the increase in pathogenicity, the shortening of the asymptomatic period falls below the response time for humoral immunity. This effect may lead to decreased adaptation, or evolutionary stasis, of *env* in the plateau phase from passages four to six. If so, it appears that *env* gene substitution is selected for by immunocompetence; thereafter, selective amplification and purifying selection control the breadth and direction of variation. Because *env* is the target for neutralizing antibodies and cell-mediated immunity, this finding is significant because *env* contains determinants for cell tropism and replication, thus playing a putative role in virulence.

In the first passage, nonsynonymous variations are almost equal to synonymous variations. Since these adaptive variations are not selective, they must be occurring in a very large and heterogeneous quantity of virus to produce the variations in noncoding nucleotides. Decreasing adaptation then continues to the seroconversion in passage three, at which point the lowest nonsynonymous and synonymous variations are seen. According to the competition exclusion principle (17, 19), equilibrium and competition lead to the outgrowth of a very homogeneous population, which we saw at seroconversion in passage three. The P3S quasispecies is dominated by a virus or viruses with a narrowed selection of genomes compared to that seen at passage one. The drop in evolution rate thus signals the end of the adaptation phase and allows competition among the viruses then present, which are of reasonably equal fitnesses. Their competition to achieve the greatest replication competence decides the dominance of the next progeny, not adaptation or immune evasion. This point marks the end of rapid *env* evolution and of large virulence increases as well.

Shpaer and Mullins have shown that immunogenicity and pathogenicity are presumably linked (38) by the correlation of high rates of amino acid change with increased virulence. The cross-species transmission of SIVsm into rhesus macaques would require the adaptation to the new host of the proteins that are involved in cell binding, entry, replication, immune evasion, and escape. The burst of evolution seen in the first three passages is evidence of this positive selection and its enforcement by the immune response to a foreign pathogen (32), as well as the cell-dependent changes needed for entry. The increased intrasample genetic variation seen between seroconversion and death confirms the notion that competent immune responses drive viral evolution to a certain extent. That nonsynonymous variations are greater than synonymous variations represents positive selection, and not drift, and the adaptation phase of the first three passages is exemplary evidence of this fact. The plateau phase of the later passages is illustrative of the purifying selection that takes place during lentiviral infections. In the absence of an antibody response, the evolutionary rate of the adapted replication-competent virus is close to stasis. The results reported here present suggestive evidence for the relationship of particular *env* sequences to virulence in rhesus macaques following inoculation with SIVsm.

ACKNOWLEDGMENTS

We thank H. Niphuis (Biomedical Primate Research Centre) for providing the passage monkey serum. P. J. Spencer Valli thanks his father for encouragement, support, and wisdom. We thank Carla Kuiken (Los Alamos National Laboratory) for crucial discussions; Lucy Phillips for much needed editorial assistance; Debbie Hauer, J. Clements, and C. Zink (Johns Hopkins University School of Medicine) for cloning advice; Chris Contag (Stanford University School of Medicine) for help with the sequence primer design; and Howard S. Fox (Scripps Research Institute) for the SIVmac sequences.

REFERENCES

1. **Allan, J. S., P. Ray, S. Broussard, E. Whitehead, G. Hubbard, T. Butler, K. Brasky, P. Luciw, C. Cheng-Mayer, J. A. Levy, et al.** 1995. Infection of baboons with simian/human immunodeficiency viruses. J. Acquired Immune Defic. Syndr. Hum. Retrovirol. **9**:429–441.
2. **Allan, J. S., M. Short, M. E. Taylor, S. Su, V. M. Hirsch, P. R. Johnson, G. M. Shaw, and B. H. Hahn.** 1991. Species-specific diversity among simian immunodeficiency viruses from African green monkeys. J. Virol. **65**:2816–2828.
3. **Amedee, A. M., N. Lacour, J. L. Gierman, L. N. Martin, J. E. Clements, R. Bohm, Jr., R. M. Harrison, and M. Murphey-Corb.** 1995. Genotypic selection of simian immunodeficiency virus in macaque infants infected transplacentally. J. Virol. **69**:7982–7990.
4. **Amedee, A. M., N. Lacour, L. N. Martin, J. E. Clements, R. B. Bohm, Jr., B. Davison, R. Harrison, and M. Murphey-Corb.** 1996. Genotypic analysis of infant macaques infected transplacentally and orally. J. Med. Primatol. **25**: 225–235.
5. **Back, N. K., L. Smit, M. Schutten, P. L. Nara, M. Tersmette, and J. Goudsmit.** 1993. Mutations in human immunodeficiency virus type 1 gp41 affect sensitivity to neutralization by gp120 antibodies. J. Virol. **67**:6897–6902.
6. **Baskin, G. B., L. N. Martin, M. Murphey-Corb, F. S. Hu, D. Kuebler, and B. Davison.** 1995. Distribution of SIV in lymph nodes of serially sacrificed rhesus monkeys. AIDS Res. Hum. Retroviruses **11**:273–285.
7. **Baskin, G. B., L. N. Martin, S. R. Rangan, B. J. Gormus, M. Murphey-Corb, R. H. Wolf, and K. F. Soike.** 1986. Transmissible lymphoma and simian acquired immunodeficiency syndrome in rhesus monkeys. JNCI **77**:127–139.
8. **Berkhout, B., and F. J. van Hemert.** 1994. The unusual nucleotide content of the HIV RNA genome results in a biased amino acid composition of HIV proteins. Nucleic Acids Res. **22**:1705–1711.
9. **Bohm, R. P., Jr., L. N. Martin, B. Davison-Fairburn, G. B. Baskin, and M. Murphey-Corb.** 1993. Neonatal disease induced by SIV infection of the rhesus monkey (Macaca mulatta). AIDS Res. Hum. Retroviruses **9**:1131–1137.
10. **Boom, R., C. J. Sol, M. M. Salimans, C. L. Jansen, P. M. Wertheim-van Dillen, and J. van der Noordaa.** 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. **28**:495–503.
11. **Bou-Habib, D. C., G. Roderiquez, T. Oravecz, P. W. Berman, P. Lusso, and M. A. Norcross.** 1994. Cryptic nature of envelope V3 region epitopes protects primary monocytotropic human immunodeficiency virus type 1 from antibody neutralization. J. Virol. **68**:6006–6013.
12. **Bronson, E. C., and J. N. Anderson.** 1994. Nucleotide composition as a driving force in the evolution of retroviruses. J. Mol. Evol. **38**:506–532.
13. **Brown, E. W., N. Yuhki, C. Packer, and S. J. O'Brien.** 1994. A lion lentivirus related to feline immunodeficiency virus: epidemiologic and phylogenetic aspects. J. Virol. **68**:5953–5968.
14. **Chen, Z., P. Telfer, A. Gettie, P. Reed, L. Zhang, D. D. Ho, and P. A. Marx.** 1996. Genetic characterization of new West African simian immunodeficiency virus (SIV) SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. J. Virol. **70**:3617–3627.

15. **Chesebro, B., K. Wehrly, J. Nishio, and S. Perryman.** 1992. Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical amino acids involved in cell tropism. J. Virol. 66:6547–6554.
16. **Cicmanec, K., and S. Norley.** 1993. Lack of immune suppression in SIV-infected natural hosts. AIDS 7(Suppl 1.):S25–S35.
17. **Clarke, D. K., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. Holland.** 1994. The red queen reigns in the kingdom of RNA viruses. Proc. Natl. Acad. Sci. USA 91:4821–4824.
18. **Daniel, M. D., N. L. Letvin, N. W. King, M. Kannagi, P. K. Sehgal, R. D. Hunt, P. J. Kanki, M. Essex, and R. C. Desrosiers.** 1985. Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. Science 228:1201–1204.
19. **Drake, J. W.** 1993. Rates of spontaneous mutation among RNA viruses. Proc. Natl. Acad. Sci. USA 90:4171–4175.
20. **Gao, F., L. Yue, A. T. White, P. G. Pappas, J. Barchue, A. P. Hanson, B. M. Greene, P. M. Sharp, G. M. Shaw, and B. H. Hahn.** 1992. Human infection by genetically diverse SIVSM-related HIV-2 in West Africa. Nature 358: 495–499.
21. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244.
22. **Hirsch, V. M., and P. R. Johnson.** 1994. Pathogenic diversity of simian immunodeficiency viruses. Virus Res. 32:183–203.
23. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson.** 1989. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature 339:389–392.
24. **Huet, T., R. Cheynier, A. Meyerhans, G. Roelants, and S. Wain-Hobson.** 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature 345:356–359.
25. **Joag, S. V., E. B. Stephens, R. J. Adams, L. Foresman, and O. Narayan.** 1994. Pathogenesis of SIVmac infection in Chinese and Indian rhesus macaques: effects of splenectomy on virus burden. Virology 200:436–446.
26. **Jurriaans, S., B. van Gemen, G. J. Weverling, D. Van Strijp, P. Nara, R. Coutinho, M. Koot, H. Schuitemaker, and J. Goudsmit.** 1994. The natural history of HIV-1 infection: virus load and virus phenotype independent determinants of clinical course? Virology 204:223–233.
27. **Kumar, S., K. Tamura, and M. Nei.** 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Comput. Appl. Biosci. 10: 189–191.
28. **Lackner, A. A., P. F. Moore, P. A. Marx, R. J. Munn, M. B. Gardner, and L. J. Lowenstine.** 1990. Immunohistochemical localization of type D retrovirus serotype 1 in the digestive tract of rhesus monkeys with simian AIDS. J. Med. Primatol. 19:339–349.
29. **Lane, T. E., M. J. Buchmeier, D. D. Watry, D. B. Jakubowski, and H. S. Fox.** 1995. Serial passage of microglial SIV results in selection of homogeneous env quasispecies in the brain. Virology 212:458–465.
30. **Lowenstine, L. J., N. C. Pedersen, J. Higgins, K. C. Pallis, A. Uyeda, P. Marx, N. W. Lerche, R. J. Munn, and M. B. Gardner.** 1986. Seroepidemiologic survey of captive Old-World primates for antibodies to human and simian retroviruses, and isolation of a lentivirus from sooty mangabeys (Cercocebus atys). Int. J. Cancer 38:563–574.
31. **Lukashov, V. V., C. L. Kuiken, and J. Goudsmit.** 1995. Intrahost human immunodeficiency virus type 1 evolution is related to length of the immunocompetent period. J. Virol. 69:6911–6916.
31a. **Miller, M. A., M. Murphey-Corb, and R. C. Montelaro.** 1992. Identification of broadly reactive continuous antigenic determinants of simian immunodeficiency virus glycoproteins. AIDS Res. Hum. Retroviruses 8:1153–1164.
32. **Mindell, D. P.** 1996. Positive selection and rates of evolution in immunodeficiency viruses from humans and chimpanzees. Proc. Natl. Acad. Sci. USA 93:3284–3288.
33. **Murphey-Corb, M., L. N. Martin, S. R. Rangan, G. B. Baskin, B. J. Gormus, R. H. Wolf, W. A. Andes, M. West, and R. C. Montelaro.** 1986. Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys. Nature 321:435–437.
34. **Nei, M., and T. Gojobori.** 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418–426.
34a. **Neildez, O., R. Le Grand, P. Caufour, B. Vaslin, A. Cheret, F. Matheux, F. Theodoro, P. Roques, and D. Dormont.** 1998. Selective quasispecies transmission after systemic or mucosal exposure of macaques to simian immunodeficiency virus. Virology 243:12–20.
35. **Novella, I. S., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. J. Holland.** 1995. Exponential increases of RNA virus fitness during large population transmissions. Proc. Natl. Acad. Sci. USA 92:5841–5844.
36. **Reimann, K. A., K. Tenner-Racz, P. Racz, D. C. Montefiori, Y. Yasutomi, W. Lin, B. J. Ransil, and N. L. Letvin.** 1994. Immunopathogenic events in acute infection of rhesus monkeys with simian immunodeficiency virus of macaques. J. Virol. 68:2362–2370.
37. **Sellner, L. N., R. J. Coelen, and J. S. Mackenzie.** 1994. Sensitive detection of Ross River virus—a one-tube nested RT-PCR. J. Virol. Methods 49:47–58.
38. **Shpaer, E. G., and J. I. Mullins.** 1993. Rates of amino acid change in the envelope protein correlate with pathogenicity of primate lentiviruses. J. Mol. Evol. 37:57–65.
38a. **Sodora, D. L., F. Lee, P. J. Dailey, and P. A. Marx.** 1998. A genetic and viral load analysis of the simian immunodeficiency virus during the acute phase in macaques inoculated by the vaginal route. AIDS Res. Hum. Retroviruses 14: 171–181.
39. **Tersmette, M., J. M. Lange, R. E. de Goede, F. de Wolf, J. K. Eeftink-Schattenkerk, P. T. Schellekens, R. A. Coutinho, J. G. Huisman, J. Goudsmit, and F. Miedema.** 1989. Association between biological properties of human immunodeficiency virus variants and risk for AIDS and AIDS mortality. Lancet i:983–985.
39a. **Valli, P. J. S., J. Heeney, and J. Goudsmit.** Selection for macrophage tropism during serial passage of SIVsm in rhesus macaques concomitant with increased virulence. Submitted for publication.
40. **Wolfs, T. F., G. Zwart, M. Bakker, and J. Goudsmit.** 1992. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. Virology 189:103–110.
41. **Zhang, J. Y., L. N. Martin, E. A. Watson, R. C. Montelaro, M. West, L. Epstein, and M. Murphey-Corb.** 1988. Simian immunodeficiency virus/delta-induced immunodeficiency disease in rhesus monkeys: relation of antibody response and antigenemia. J. Infect. Dis. 158:1277–1286.

Chapter 5

Structured-Tree Topology and Adaptive Evolution of the Simian Immunodeficiency Virus SIVsm Envelope during Serial Passage in Rhesus Macaques According to Likelihood Mapping and Quartet Puzzling

JOURNAL OF VIROLOGY, May 1998, p. 3673–3683
0022-538X/98/$04.00+0

Vol. 72, No. 5

Structured-Tree Topology and Adaptive Evolution of the Simian Immunodeficiency Virus SIVsm Envelope during Serial Passage in Rhesus Macaques According to Likelihood Mapping and Quartet Puzzling

P. J. SPENCER VALLI AND JAAP GOUDSMIT*

Department of Human Retrovirology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

Received 9 June 1997/Accepted 26 January 1998

Species-specific strains of simian immunodeficiency virus (SIV) are nonpathogenic in African primates. The SIV strain most closely related to human immunodeficiency virus type 2 (HIV-2) is SIVsm, the strain specific to the sooty mangabey (*Cercocebus atys*). Infection of Asian primates with SIV causes AIDS and allows the study of the adaptive evolution of a lentivirus to replicate efficiently in a new host, providing a useful animal model of HIV infection and AIDS in humans. Serial passage of SIVsm from sooty mangabeys in rhesus macaques drastically shortened the time of disease progression from 1.5 years to 1 month as the retrovirus adapted to these Asian hosts. In the present study we analyzed the quasispecies nature of the SIVsm envelope gene (*env*) during serial population passage in rhesus macaques. We asked ourselves if phylogenetic evidence could be provided for the structured topology of the SIVsm *env* tree and subsequently for the adaptive evolution of SIVsm *env*. Likelihood mapping showed that phylogenetic reconstruction of the passage was possible because a high percentage of the sequence data had a "tree-like" form. Subsequently, quartet puzzling was used and produced a phylogeny with a structure parallel to the known infection history. The adaptation of SIVsm to Asian rhesus macaques appears to be an ordered process in which the *env* evolves in a tree-like manner, particularly in its constant regions.

Zoonotic transmission of some lentiviruses causes AIDS. The two principal etiologic agents of AIDS in humans are the human immunodeficiency viruses types 1 and 2 (HIV-1 and HIV-2), whose probable ancestors are found in two species of feral African primates (simian immunodeficiency virus SIVcpz and SIVsm) (17). HIV-2 appears to be descended from the SIVsm lineage (Fig. 1), as shown not only by sequence homology but also by the geographically localized infections (27) of humans and sooty mangabeys (*Cercocebus atys*) with this virus. However, the issue of HIV origins remains open to question (33a). The SIV strains do not induce disease in their natural hosts. The relation between pathogenesis and virus strain is not well understood but appears to be both host and pathogen related. When Asian macaques are experimentally infected with SIVsm, SIVagm, SIVmac, or a related molecular clone, the disease course mimics that seen in human AIDS (41) and thus provides a useful model for vaccine testing.

Although the length of the asymptomatic period of SIV infection varies greatly among individual Asian primates, the numerous SIVsm-derived viruses show a wide spectrum of pathogenicity or virulence. The most pathogenic SIVsm strain is the SIVsmPbJ14 (14, 20, 37), with an average time to death of 14 days. SIVsmH4 and SIVMac239 are both pathogenic during experimental infection but have a mean time to death of 1 year (19, 21–24). The related clones SIVsmm9 and SIV1A11 are relatively nonpathogenic, with times to death of several years. The length of the asymptomatic period during experimental infection with the many SIVsm-derived clones (11, 25, 28, 31, 37, 41, 42, 46) is presumably virus strain dependent, as shown by repeated infections with the same clones.

Viral variation is known to occur within a single cycle of intracellular replication (40), and the progeny virus from an infected cell may vary genetically from the parental virus. This genetic flexibility enables the rapid development of new mutants (3) by means of subtle changes in the *env* surface glycoprotein (3, 9, 12). The *env* genes of HIV-2 and SIVsm have the same functional layout, with five variable and constant domains and discrete (over 98% nonsynonymous [35, 38, 55]) nucleotide changes within the variable regions (V1, V2, V3, V4, and V5). The high nonsynonymous-substitution rate supports positive selection for nonfounder or adapted mutant virus by outgrowth of these new variants from the circulating virus population of founder or primary inoculum genotypes (14, 33, 47). Thought to be driven by the antigenic selection (14, 47) of new variants, constant evolution of new genotypes and the coexistence of distinct strains (54) during infection is believed to be sequence variation dependent, due to the number of different genomic variants that exist at any one time.

The goal of this study was to examine viral sequence variation during adaptation of the V1 to V4 regions of the *env* gene of SIVsm in the serum of immunodeficient rhesus macaques that were experimentally infected. The experiment was carried out by serial population passaging of a known SIVsm strain (SIVsmB670), accentuating the selection for adapted mutants in vivo (10), to elucidate the changes in the *env* gene that influence the marked increases in pathogenicity and virulence of SIVsm infection in the rhesus macaque model. Phylogenetic inference methods have been used to study the origins of HIV (33a, 34, 36) and the epidemiology of HIV infections (17) and to determine definitive transmission patterns (27). The use of phylogeny reconstruction methods with a known infection his-

* Corresponding author. Mailing address: Department of Human Retrovirology, Academic Medical Centre, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands. Phone: (31-20) 566-4853. Fax: (31-20) 691-6531. E-mail: p.j.valli@amc.uva.nl.

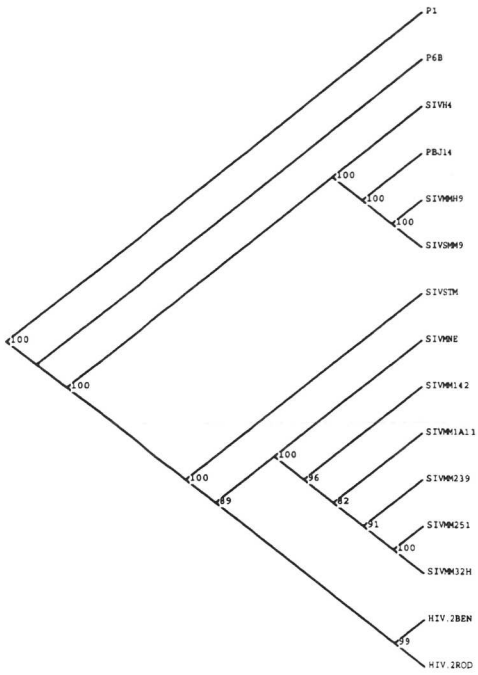


FIG. 1. Phylogeny of the African SIV and HIV-2 strains commonly used in experimental infections. Known pathogenic and less pathogenic SIVs are seen to cluster together (e.g., SIVH4 and SIVSMM9, and SIVMM239 and SIVMM1A11) according to virus strain and not by pathogenicity. The P1 and P6 sequences are more divergent than some of the other clades. It is also noteworthy that the pathogenicity of the P6A and P6B viruses is equal to that of PBJ14 but there is no tendency for these two SIVSMM derived *env* sequences to cluster together. The same results were achieved with different P1 and P6A or P6B sequences as examples. This tree was constructed with PUZZLE 3.0 with the Tamura-Nei distance approximation and 1,000 tree-puzzling steps and was rooted with the serial passage P1 sequence. The number at each internal branch is the percentage of the 1,000 intermediate trees having that branching pattern. The branch lengths are not important in this cladogram; only the reliability values and branching pattern are important.

tory has allowed the study of the evolutionary pattern of SIVsm *env* adaptation during serial passaging. We used likelihood mapping (49) to display the amount of tree-like information contained in the SIVsm *env* sequence set and to infer the phylogeny. The various modes of nucleic acid evolution have led to the development of programs that infer tree-like relationships (18, 51), net-like geometries (4), and start formations (49). The likelihood mapping shows the phylogenetic content in a graphic form, demonstrating whether the data are suitable for phylogenetic reconstruction.

Likelihood mapping determines the best arrangement of the smallest information-containing subset of a tree, the quartet (four taxa [49]). Plotting all of the possible quartets as tree-, net-, or star-like relationships of the seven possible forms of a quartet of sequences (16, 49) allows efficient visualization of the estimated phylogenetic content. Our application of this method to the SIVsm passage alignment showed it to be highly informative compared to HIV alignments, and the reconstruction of the phylogeny was tree-like. Construction of trees by quartet puzzling (QP) (48–50) produced a phylogeny that showed the adaptive process to be highly structured, since the divergence increases with passaging. The ordered divergence from the primary inoculum occurred in the same time, at the same rate, as the shortening of the asymptomatic period or the increase in pathogenicity of the passaged SIVsm strain. The adaptation of SIVsm to Asian rhesus macaques is thus a structured, quickly occurring process in which the *env* gene evolves in a tree-like manner.

MATERIALS AND METHODS

Virus. Delta B670 (2, 5–7, 32, 35, 55) is an SIV strain that was originally discovered in a sooty mangabey presenting with a cutaneous lepromatous lesion. When tissue from this animal was used to inoculate rhesus macaques, it resulted in a syndrome that closely reproduced the effects of AIDS in humans (55). The virus isolate was proven to be type D negative and was extensively characterized (2, 5, 55).

Passage. The passaging was carried out with seven age-matched Asian rhesus macaques, all 2 years old. Their experimental infection with an SIVsm strain was performed in six consecutive steps. The first monkey was infected intravenously with 5×10^2 infectious doses of SIV Delta B670. The next five monkeys were intravenously inoculated in a serial fashion with 2×10^6 peripheral blood mononuclear cells (PBMC) taken at the symptomatic stage of disease from the preceding rhesus macaque. For passage 5, PBMC of the passage 5 monkey (P5) were used to infect two monkeys (P6A and P6B) because little blood was available from monkey P5 due to its rapid progression to AIDS and sudden death. The time to death postinfection (tdpi) and moment of sampling for the reverse transcriptase PCR (RT-PCR) were as follows: P1, tdpi = 18 months, RT-PCR at 18 months; P2, tdpi = 12 months, RT-PCR at 3 and 7 months; P3, tdpi = 9 months, RT-PCR at 2 and 4.1 months; P4, tdpi = 4 months, RT-PCR at 1 and 1.8 months; P5, tdpi = 2 months, RT-PCR not done since no sample was available; P6A, tdpi = 2 months, RT-PCR at 2 months; P6B, tdpi = 2 months, RT-PCR at 2 months. The animals were euthanized upon evidence of undue discomfort. PBMC used for serial passage were not cultured or cocultivated with any other cells.

Viral RNA was harvested with silica in the presence of a chaotropic agent (8) from the serum of the infected macaques for use as a template in the RT-PCR. Viral RNA was isolated from 20 μl of sera and resuspended in 20 μl of RNasin in H_2O (1 U/μl) and used in a reverse transcription reaction (5 μl of viral RNA, 250 μM each deoxynucleoside triphosphate, 2 ng of 3′ RT-PCR primer [SIV4Not1: TTATATGCGGCCGCCTACTTTGTGCCACGTGTTG] per μl, 2.5 mM Mg^{2+}, 1 U of RNasin [Promega] per μl, 10 U of Super Script I [Gibco-BRL], and 1× reaction buffer [45] in a 20-μl volume). The components were assembled at 37°C and incubated at that temperature for 90 min. The PCR mixture consisted of 250 μM each deoxynucleoside triphosphate, 2 ng of 3′ and 5′ (SIV1HIII: GTAGACAAGCTTGGGATAATACAGTCACAGAAC) PCR primers per μl, 1× reaction buffer, 2.25 mM Mg^{2+}, and 1.5 U of *Taq* polymerase (Perkin-Elmer Cetus) in a final volume of 100 μl including 5 μl of the reverse transcription reaction mixture. The PCR mixture was overlaid with paraffin, heated to 95°C for 5 min, and subjected to 35 cycles of 1 min at 95°C, 1 min at 55°C, and 1 min at 72°C, ending with 10 min at 72°C in a Perkin-Elmer Cetus DNA thermocycler. The reverse transcription reaction and PCR were done in duplicate for each sample to prevent mispriming and to ensure the fidelity of the virus genotypes sampled. After PCR, the samples were combined and size selected on 0.8% agarose gels. The 1,151-bp band was excised and isolated from the gel slice, digested with *Not*I and *Hin*dIII, and again isolated from the agarose gel by silica particle (8) fragment isolation. Size-selected, digested, purified RT-PCR product was ligated overnight into plasmid pSP64 (Promega) containing a *Not*I site. Ligated products were electroporated into electrocompetent *Escherichia coli* C600, and double-stranded plasmid DNA for sequencing was isolated on Qiagen columns. Five clones per sampled time point were sequenced, except for the primary inoculum (P1), for which 10 clones were sequenced. The clones were named by the passage number (P), and by the infection point, i.e., seroconversion (S) and death (D).

Sequencing and analysis. The double-stranded plasmid DNA was sequenced by using custom-labelled dye primers (Applied Biosystems Inc.) with an automated sequencer (model 373A; Applied Biosystems) and version 1.2.0 software. Clones were assembled and aligned with the Sequence Navigator program (Applied Biosystems) or Clustal V (18) and further optimized manually. Phylogenetic analysis was conducted with MEGA (26), Clustal V (18), PAUP 3.1.1 (52), and PUZZLE 3.0 (49). Neighbor-joining, maximum-likelihood (ML), and puzzle-based trees were produced by using the Tamura-Nei distance estimation with pairwise comparison and then bootstrapped with 1,000 replications before tree construction with and without preselected outgroups (P1 sequences). The ML mapping and QP were carried out by using the Tamura-Nei distance estimation (49). Briefly, likelihood mapping is the construction of all quartets of sequences possible (e.g., A, B, C, and D, and then A, B, C, and E, etc., until all of the possible quartets of sequences in the analysis have been compared) and calcu-

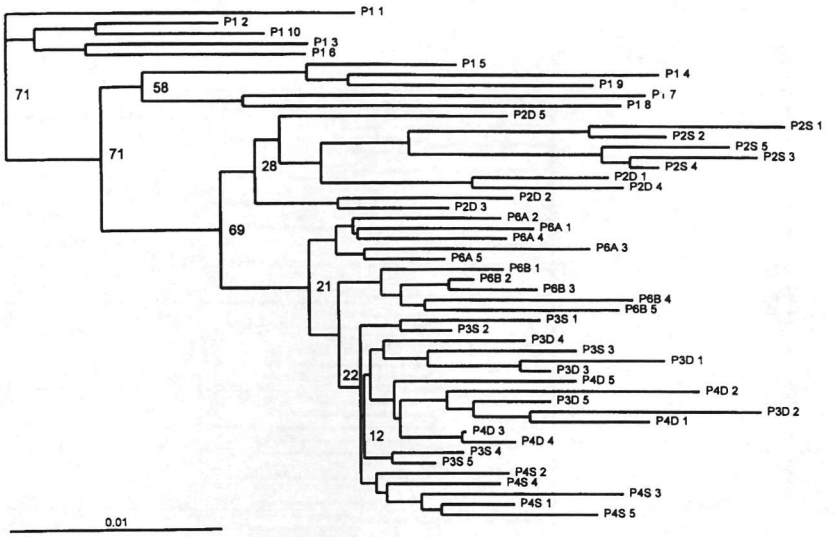


FIG. 2. SIVsmB670 passage full-length clone NJ tree constructed with ClustalV, allowing for positions with gaps. The numbers on the major branches are the bootstrap values (percentages) after 500 resamplings. The tree is rooted with the P1-1 clone. The tree had similar topology to trees generated by MP. The scale bar is in the bottom left-hand corner.

lation of the relative weights of the probabilities for all three resolved tree structures (1: A is like B but not C or D; 2: A is like C but not like B or D; and 3: A is like D but not like B or C) (48–50) and four unresolved structures (partially resolved 1: A is like B or C but not D; partially resolved 2: A is like B or D but not C; partially resolved 3: A is like C or D but not B; and unresolved: A is not like B, C, or D). Thus, for the first four sequences analyzed by likelihood mapping, the seven possible locations of the probability vectors (49) include three with a completely resolved placement of the four taxa, three with one taxon split between attraction to two others and not like the fourth taxon (partially resolved), and one of star-like form or no favored relation between the four taxa (unresolved). These seven relations can be plotted in a triangle to show the phylogenetic signal (see Fig. 3, 5, and 7). The corners represent areas of completely resolved (tree-like) quartets (A is most like one of B, C or D, resolved 1, 2, and 3); the sides represent quartets with a split between equal attraction of one taxon for two other taxa (A is like B and C, B and D, or C and D, partially resolved); and the center of the triangle represents an area of equal likelihood or distance between the four taxa (A is not like B, C, or D, unresolved, giving a star-like formation). If each possible quartet vector (resolved, partially resolved, and unresolved quartet comparison) is plotted as a point in the triangle, it gives a rapid visualization of the phylogenetic content of the alignment data (robustness of the tree, or lack thereof). Tree construction was carried out with the ML-based PUZZLE 3.0 (48). The ML analysis is used for determination of a tree and its corresponding branch lengths that have the greatest likelihood of reconstructing the correct phylogeny. The numbers of possible topologies increases exponentially as the number of taxa increases, making heuristic searches very slow. Reconstructing a tree from all of the possible sets of four taxa, or quartets, allows the heuristic search to proceed efficiently in an ML procedure (50). QP reconstructs the ML tree for all possible quartets. The total set of quartet trees is subsequently combined to form a complete tree in the QP step. This procedure was carried out 1,000 times. The QP tree is a majority-rule consensus (30) of the set number of puzzling steps. There are three main steps to the QP method; ML determination of all possible quartet trees, combining all of the quartet trees 1,000 times to form a complete tree (puzzling step), and the final majority-rule consensus computation of all intermediate trees (49). The HIV sequences from 44 progressors and nonprogressors at seroconversion and after 5 years of infection (29) are available upon request.

RESULTS

Construction of molecular phylogenetic trees. Our study included seven Asian rhesus macaques whose serial infection involved a well-characterized viral isolate (SIVsm B670) (2, 5, 6, 35), and known times of infection, seroconversion, and direction of transmission. It thus relates a serial population passage and known phylogeny of an SIV. The use of phylogenetic analysis on the quasispecies sequences allows evaluation of the inference methods and insight into the pattern of viral evolution. The inference of the correct molecular phylogeny has

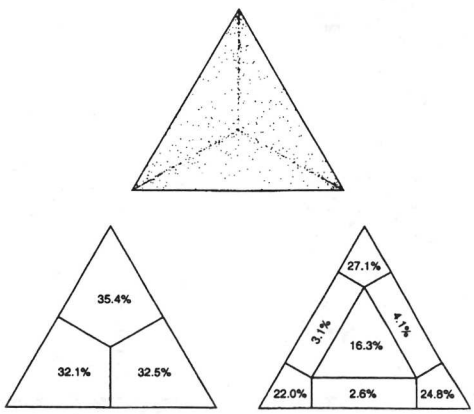


FIG. 3. HIV-1B likelihood-mapping analysis for the 44 HIV-1 infected progressors and nonprogressors (29). The top triangle shows the distribution pattern of the quartet analysis, and the bottom triangles show the distribution of the data in the seven basins of attraction or quartet possibilities. As shown, the majority of the data clusters in the center of the triangle, where unresolved or star-like trees with little phylogenetic content are found. The three corners of the triangles, where completely resolved trees are found, are virtually devoid of content.

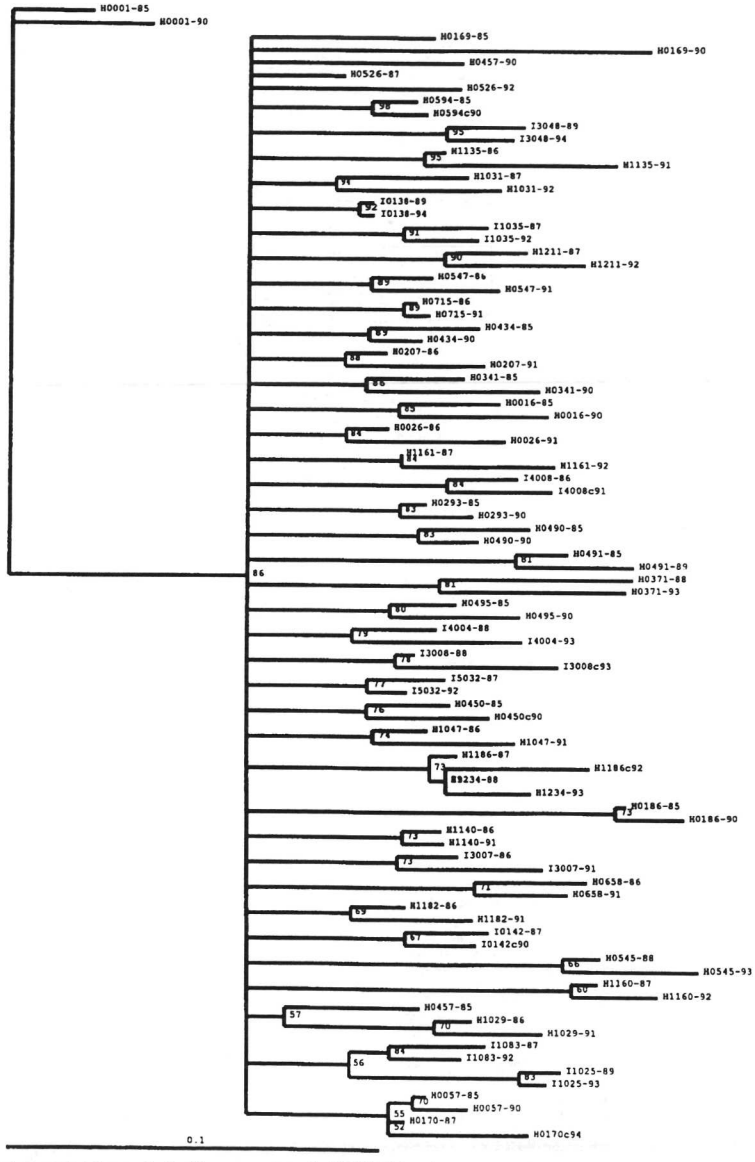


FIG. 4. HIV-1B phylogenetic inference for the 88 sequences from the 44 HIV-1-infected progressors and nonprogressors, 5 years apart, (29) by ML determination (PUZZLE 3.0) with the Tamura-Nei distance approximation and 1,000 tree-puzzling steps. The QP percentages of the important branches are in bold numbers. The scale bar is shown in the bottom left-hand corner. The topology of the HIV-1B tree is star-like as predicted by likelihood mapping and shows little relation between the sequences, except for the pairs of sequences from individual patients, which group together.

important implications for our understanding of transmission patterns and quasispecies evolution. We used four different tree-building methods: neighbor-joining, ML, maximum parsimony (MP), and QP. The results of the four methods were similar only if the number of taxa used in the analysis was reduced. The neighbor-joining, ML, and MP methods accommodated all 50 clones and maintained a recognizable portrayal of the known phylogeny (Fig. 2), albeit with discrepant branch placement. Inordinate amounts of time were needed with MP for branch swapping to prevent the irreversible misplacement of branches early on in the stepwise addition. The QP reconstruction produced the most true phylogeny, with knowledge of the relation of the viruses present at the time points sampled during the serial passage. The more closely related sequences (less heterogeneity between the sampled time points; e.g., the genetic distance between P1 to P2S is more than from P3D to P4S) after the third passage present a resolution problem for all of the known methods. The P3S, P3D, and P4D sequences clustered consistently. The discrepancy was the branching order of the P6A and P6B sequences, which were closer to the founder P1 and P2 sequences than to their precursor P3 and P4 sequences. This problem was not resolved by the use of other distance estimation methods or increased bootstrap resamplings.

The problem with tree construction with a large number of taxa is the inability to calculate all of the possible tree combinations. It is further complicated by the long-standing problems in phylogenetic inference. The evolutionary history of retroviruses is difficult to reconstruct because of recombination between similar but unlike RNA strands, selection of progeny during rapid turnover of viral genomes, differences in the rates of mutation among strains, unbalanced nucleotide frequencies, and differences in the individual substitution rates of the already skewed nucleotide concentrations. The use of quartet analysis, with its focus on the smallest informative subtree (27, 48–50), has been applied to an inference method that takes advantage of the determination of all possible quartets, or heuristic search, which is unavailable to the other methods, too time-consuming, or available only as a statistical test of best fit (27). PUZZLE 3.0 (48–50) incorporates a first-step plot of all the possible four taxon quartets as a means of weighting the best placement of them all within a phylogeny. With seven possible organizations for each quartet, calculation of all possibilities is feasible. The ML of the grouping of a quartet can also be plotted graphically to give an idea of the "tree likeness" of the data known as likelihood mapping (49) (see Fig. 3, 5, and 7). The four taxa are compared by grouping three and comparing the fourth with them by placing it in one of the seven possible ML areas as a dot: with taxon D most like taxon A, B, or C (the three corners: tree-like), split between A and B, A and C, or B and C (the three sides: unresolved between the two), and, lastly, unlike A, B, or C and placed in the middle as a neutral attractor (star-like formation). The centers of the ML triangles are inhabited by star-like phylogenies lacking tree-like content, i.e., having no attraction to cluster with the other taxa (Fig. 3 and 4). The percentages of the contents of the complete set of quartets (≥300,000 for 50 taxa) are displayed in the numbered triangles. Each of the seven areas has a finite number of the ML dots, with the total of the seven being 100%.

The utility of likelihood mapping is shown by analysis of an extensive set of HIV-1B sequences (Fig. 3) (29) from 44 infected progressors and nonprogressors sampled at seroconversion and 5 years thereafter. The three corners contain 74% of the quartets, and fully 26% of the quartets are unresolved, with points in the center of the triangle (16.3%) being neutral attractors or star forms. The phylogeny of this alignment (Fig. 4)

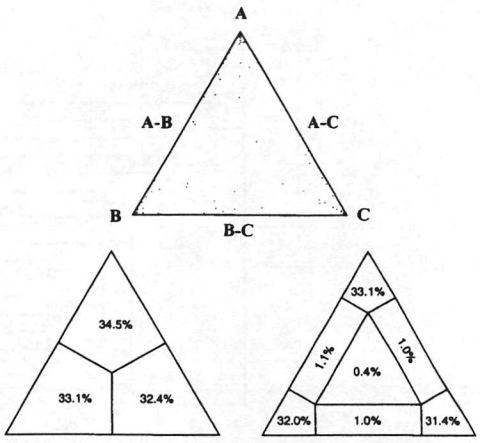


FIG. 5. Likelihood-mapping analysis for the SIVsmB670 passage complete clone alignment. The top triangle shows the distribution pattern of the quartet analysis, and the bottom triangles show the distribution of the data in the seven basins of attraction or quartet possibilities. As shown, the majority of the data clusters in the three corners, where completely resolved robust trees with large amounts of phylogenetic content are found. The three corners of the triangles, where completely resolved trees cluster, have almost the entire content.

shows intrapatient sequence clustering and no definition of evolution of the HIV-1 B *env* gene over time. Evolution of HIV-1 during infection is not as structured as the cross-species passage because several, or many, infections have already taken place between the progenitor virus (presumably a relation of SIV from chimpanzees) and the modern HIV-1. The lack of sequences from the "original" founder of HIV-1 leaves a gap in the phylogeny. Without the founder and early infection sequences, the variation is seen to be nonstructured, since there is no directed evolution or divergence over time.

Figure 5 shows the tree likeness of all 50 taxa of the SIVsm passage. The sequence alignment contains a very high 96.5% of the phylogeny in the three corners, or "in" the tree. There are 241 variable and 135 parsimonious sites of 1,119 nucleotides in this alignment. The split-between-partners quartets or sides of the triangle include 3.1% that cluster between two other taxa and only 0.4% that are unresolved ("out" of the tree) in the "star" area where HIV-1 phylogenies exist (Fig. 4).

The phylogeny derived from generation of all possible quartets, QP (Fig. 6), tells the complete story of the serial passage in graphic form. As the infection history is known, the tree is rooted in the P1-1 clone as the founder sequence. The heterogeneity of the P1 sequences is evidenced by the long branches that contain only pairs of taxa. The initial adaptative effect is seen in the large differences in the lengths of the P1 branches, varying from P1-6 to P1-8 and P1-9 (which are the same distance from P1-1). P2S and P2D are not completely separate, with two clusters forming, but four of the five P2-D taxa are distinctly divergent from the P2-S cluster, and P2D-2 is a likely founder for the next passage, although P2S-4 is actually closer to the P3 to P6 sequences. The branch lengths and the clustering numbers place the P2 taxa between the P1 and the P3 to P6 taxa. This pattern follows the changes in pathogenicity during passage and holds for the P1 to P2 clustering as well. The P3 to P6 taxa are placed on the same branch, with a

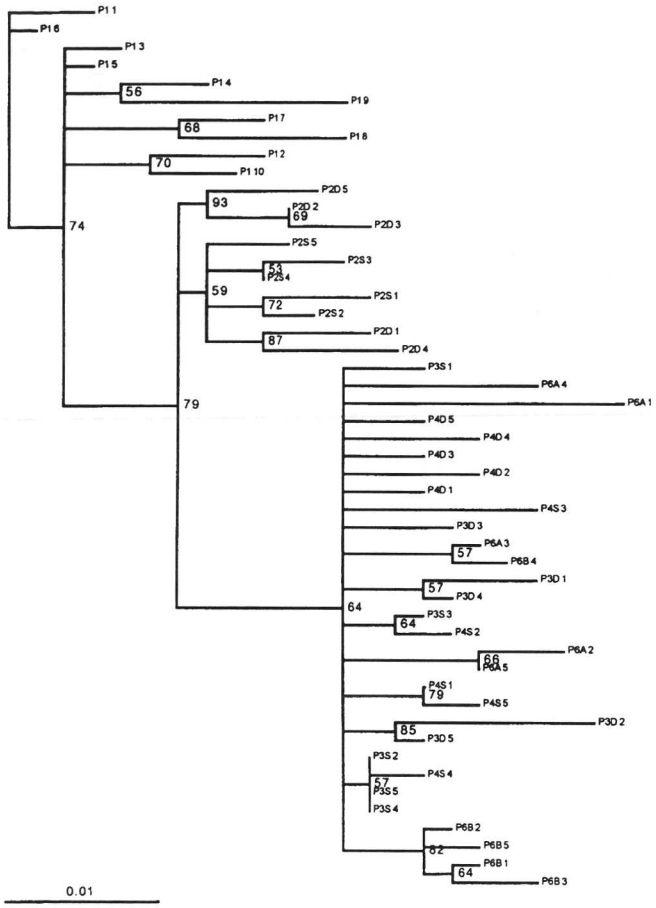


FIG. 6. SIVsmB670 passage complete clone phylogenetic inference for the passage alignment. The tree was constructed by ML determination (PUZZLE 3.0) with the Tamura-Nei distance approximation and 1,000 tree-puzzling steps. The QP percentages of the important branches are shown by boldface numbers. The scale bar is in the lower left-hand corner. The tree shows the pattern of evolution during adaptation of SIVsm to a new species of nonhuman primate and the effect of the serial passages upon *env* during the increases of pathogenicity.

clustering value of 64, since the P3 and P4 taxa do not carry much information that could cluster them separately as distinct clades. The differences in branch lengths are indistinct because of minimal sequence variation between the P3 and P6 clones. There is an almost quantum-like effect, with the branch lengths of the P3S-1 and seven other taxa all being equal. The P4D-4 and six other clones also have equal branch lengths, seemingly twice as long as those of the P3S-1 length. The tree continues to grow outward from the root, with the P6 sequences having longer branch length than those of the P3 and P4 sequences. The P6B sequences form a distinct branch, except for the P6B-4 clone, which has seven amino acid substitutions not found in the other P6B sequences.

ML mapping was used to display the amount of tree-like information held by a separation of the sequences (P1 and P2 as one cluster and the P3 to P6A and P6B sequences as another). The robustness of the branching separation of the two clusters (Fig. 7) is seen as the absence of any shared quartets which would appear in the center of the triangles. The mapping shows the correlation of the branching order and distinctly altered pathogenicities of these quasispecies. The taxa were set in two distinct clusters, with 97.6% of the phylogenetic information confirming the separation of these subsets of the passage clones as evolutionarily distinct.

We divided variable and constant regions into two separate phylogenies to assess the pattern of evolution of these two functional types of *env* domains during adaptation and increase in pathogenicity (Fig. 8 and 9). The variable-region phylogeny

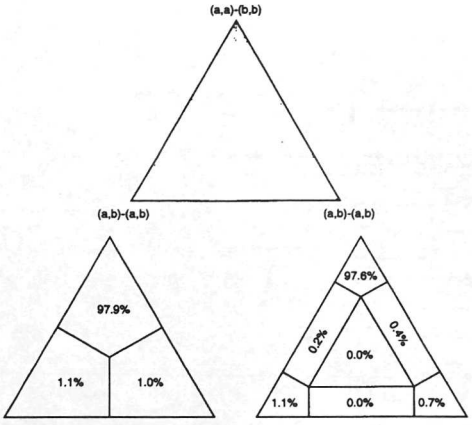


FIG. 7. SIVsmB670 passage complete clone two-cluster likelihood-mapping analysis. The sequences were split into two disjoint groups: a (P1 to P2) and b (P3 to P6). The corners of the triangles are labelled with the corresponding tree topologies. The top triangle shows the distribution pattern of the quartet analysis, and the bottom triangles show the distribution of the data in the seven basins of attraction or quartet possibilities. The likelihood-mapping analysis can also be applied to the testing of an internal edge of a tree. The clusters chosen separate the early, less pathogenic *env* clones of the virus from the later, more pathogenic *env* clones.

shows the large heterogeneity within the P1 sequences that was observed above. With 65% of the variable sites being found within the variable regions, the P1 sequences are split into three separate branches. Cluster definition decreases after seroconversion of passage 3 (P3S clones), except for the P6B sequences. The heterogeneity appears to be very large, but, as the distance bar shows, the distances are far greater than those encountered in the constant regions. Misclustering was found due to the enormous variation, and not the similarity, of the sequences. In the variable regions, the average diversity was limited after the P3D sequences.

The constant-region tree structure places the different passages in order (Fig. 9). The P1 heterogeneity is still large, but not large enough to reduce the cluster definition of the whole tree, as in the variable regions. The P3S sequences are the most clustered and mark the greatest increase in pathogenicity, with viral diversity increasing as the passage progresses toward the P6 sequences.

DISCUSSION

Serial passage of SIVsmB670 from sooty mangabeys in rhesus macaques has drastically shortened the time of disease progression from 1.5 years to 1 month during adaptation of the retrovirus to the Asian rhesus macaque hosts. The increase in pathogenicity has been correlated with higher antigen levels in plasma, persistent plasma antigenemia, and a more rapid loss of T-helper/memory cells (53a). The more rapid disease progression of the P6 inoculum produces clinical symptoms similar to those due to the slower progressing P1 inoculum. The SIVsm-related strain PBJ14 (Fig. 1) is pathogenic within 2 weeks in pig-tailed, cynomolgus, and some rhesus macaques (43) but does not produce the immunodeficiency seen in HIV-1 infections. It caused massive T-cell proliferation, increased cytokine release, and mucoid diarrhea, and its *env* glycoprotein has some superantigen-like properties (44). The P6 inoculum reproduced the relevant immunodeficiency symptoms to model AIDS in humans, except for lack of seroconversion.

The present study was used to analyze the genetic adaptation of the SIVsm *env* gene during serial population passage in rhesus macaques. The clones sequenced were all unique, with decreased heterogeneity or viral diversity in the P6 sequences compared to those of P1. The heterogeneity or divergence seen may not be representative of genotypes present in tissue, but, for the purpose of identifying the evolutionary path during the adaptation of the *env* gene between the P1 and P6 sequences, this single compartment sampling suffices. The major areas of variability were observed in the V1 and V4, as is the norm in SIV infections (10, 19, 39), but mutations were seen also within the V2 and V3 regions and the C3 region just C-terminal to the V3 loop.

Phylogenetic inference methods were used to analyze the evolution of the *env* gene of SIVsm during adaptation in rhesus macaques. The QP method was used to allow the rendering of phylogenetic content of the data set and for the two-cluster mapping of the variable versus the constant domains of the *env*. Four-taxon trees may, in some purposely selected cases, be difficult to resolve (18a), but a one-tree example is not a defining study (54). Sequence length is always a crucial factor in tree building, but the simultaneous analysis of many lineages tends to improve phylogenetic estimation considerably (18b). Quartets can be hard, but extra information helps. If all that is available consists of data on species A, B, C, and D, it might be relatively difficult to find the correct tree for them. However, if additional data are available (species E, F, G, ...) and an attempt is made to find a tree for all the species, that part of the tree relating A, B, C, and D will be expected to be more accurate than if just the data for A, B, C, and D were available. There are many examples of subsets of four species which in themselves might be hard to resolve correctly but which are correctly resolved thanks to additional data (e.g., inordinately long sequences or many lineages). PUZZLE 3.0 gains advantage from extra data in the same way. Its "understanding" or resolution of the quartet A, B, C, and D may be incorrect, but the information on the relationships of A, B, C, and D implicit in its treatment of A, B, C, and E, of A, B, D, and E, of A, C, D, and E, of B, C, D, and E, of A, B, C, and F, of A, B, D, and F, of A, C, D, and F, of B, C, D, and F, of A, B, C, and G, etc., should overcome this problem (17a). Using this quartet-based heuristic search algorithm (50) for the best tree fit and 50 sequences of 1,119 bp each, we found that the sampled sequences diverged from that of the primary inoculum in a time-dependent fashion, following the path of the serial passage. The topology of the tree placed the *env* sequences in clusters roughly according to the pathogenicity of the virus and the history of the passage infections.

The relationship between infection history, pathogenicity, and the hierarchy of the inferred trees indicates that *env* plays a role in the progression to AIDS during SIV infection. The sequence divergence and its correlation with enhanced virulence suggest that our recorded amino acid substitutions are important to the shortening of the asymptomatic period of progression to AIDS in the Asian rhesus macaques.

Viral diversity decreased during passaging (Fig. 6, 8, and 9) (53), in particular in the constant regions, following the initial adaptation to the new host. The phenomenon of viral diversity is pivotal to the production of vaccines. We have shown here that this diversity is related to the level of adaptation of the virus to its new environment and that, after several infections, an equilibrium is reached in amino acid substitution (unpub-

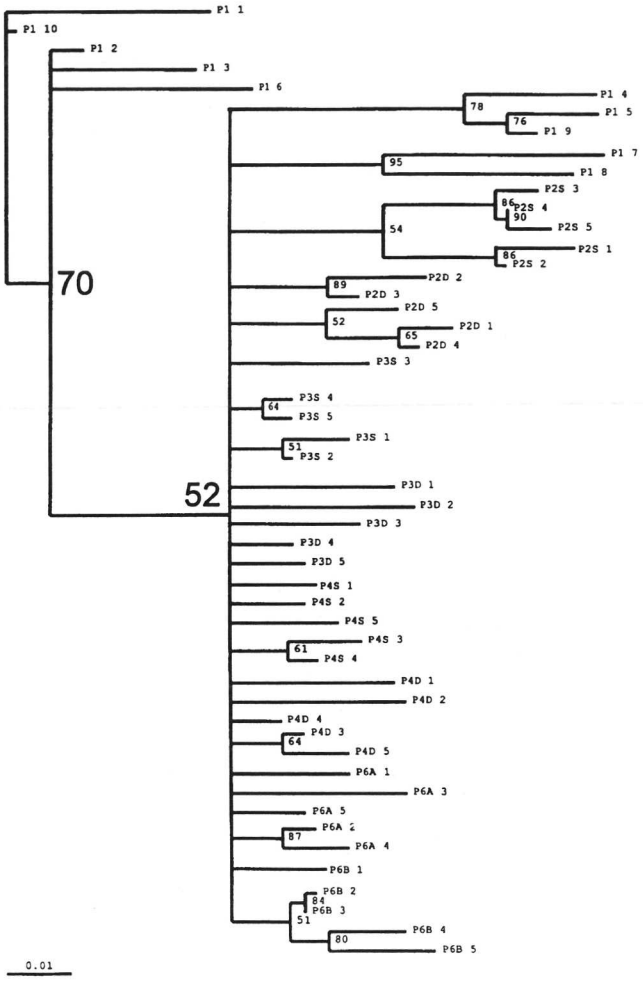


FIG. 8. Phylogenetic inference for the variable regions of the SIVsmB670 passage alignment. The tree was constructed by ML determination (PUZZLE 3.0) with the Tamura-Nei distance approximation and 1,000 tree-puzzling steps. The QP percentages of the important branches are shown by boldface numbers. The scale bar is in the lower left-hand corner. The pattern of evolution of the variable regions is less structured than that of the constant regions (Fig. 9). The P1 to P6 sequences cluster together on one branch, except for the root (P1-1) and four P1 clones (P1-2, P1-3, P1-6, and P1-10).

lished data), K_s/K_a ratio (53), and branch length variation at which the virus is optimally tailored to the surrounding conditions. Variation in the amino acids of the *env* product was greatest in the V1 and V4 regions followed by V2 and V3. The difference in receptor sequence between the CD4 of sooty mangabeys and Asian rhesus macaques is not known, nor is the difference in relatedness of the accessory receptors for this virus in these two hosts (1, 13, 15). The adaptive process may hinge on the viral need for the most efficient *env* conformation to bind the CD4 T cells and begin the infective process. Increased pathogenicity would imply increased viral adaptation and increased ability to bind and enter the target cells of the new host. The most variable, or most adaptable, regions were V1 and the V4. The precise role of V1 (in HIV or SIV infections) is not yet known. V3 and C4 are known to play a role in the binding of CD4 (23); although not binding directly, they play a conformational role. Because V3 and C4 span V4, V4 (along with other regions in the *env* gene) may encode amino acids that are part of the CD4 binding domain. If the binding of the CD4 and accessory receptor are relevant to the development of lentivirus pathogenicity, it would follow that the fastest-adapting areas of the *env* gene would be those need-

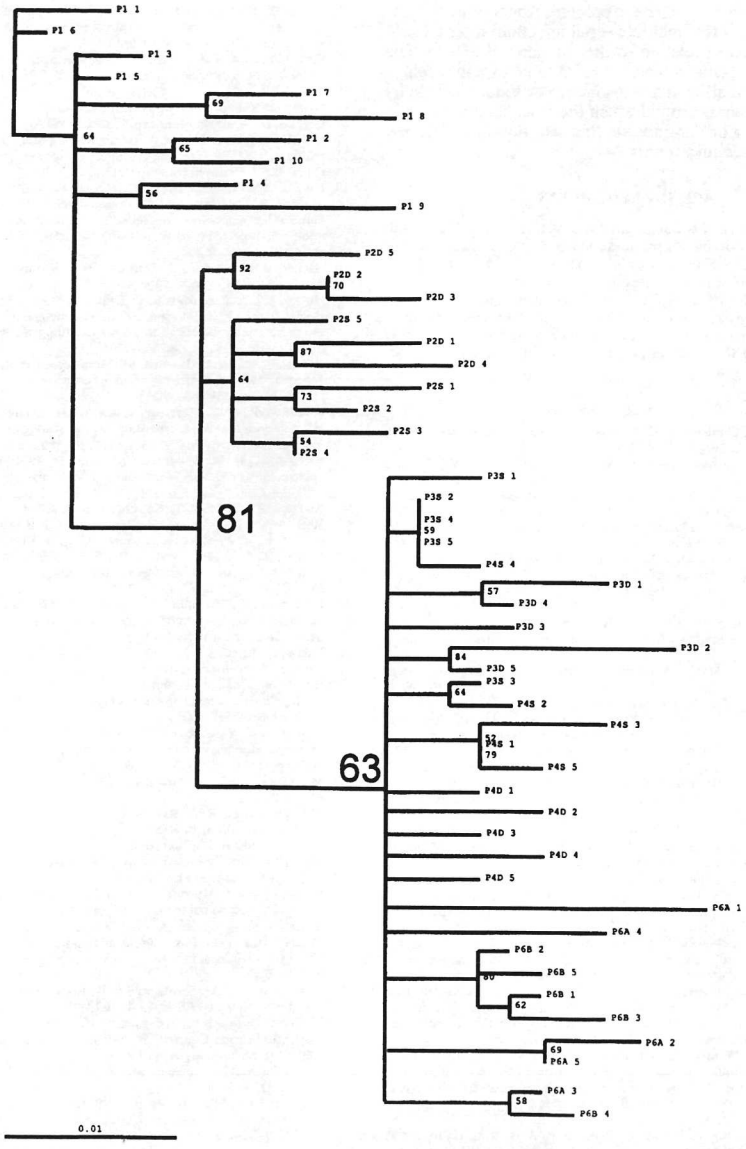


FIG. 9. Phylogenetic inference for the constant regions of the SIVsmB670 passage alignment. The tree was constructed by ML determination (PUZZLE 3.0) with the Tamura-Nei distance approximation and 1,000 tree-puzzling steps. The QP percentages of the important branches are in bold numbers. The scale bar is in the lower left-hand corner.

ed to contact the CD4 and accessory receptor for SIVsm (CCKR-5 is the known accessory receptor for HIV-1 and SIV (15). If so, V1 and V4 are in conformational contact and are the major *env* structures involved in binding and entry.

The relative clustering of the *env* sequences by virulence and by sequence divergence shows the impact of genetic variation of *env* upon pathogenicity. The difficulty of nonrobust phylogenetic separation of the P3 to P6 sequences shows that *env*

evolution is not a simple linear process. Apparently, convergence can occur during multiple serial infections after adaptation has occurred in relation to the *env* gene of SIVsm. The decreasing heterogeneity seen in the P3 to P6 taxa may reflect the reduced adaptation and convergence or sequence stability of the new consensus around which these sequences evolve in a decreasing area or may indicate that selection has effectively removed unsuccessful variants.

ACKNOWLEDGMENTS

We thank Korbinian Strimmer and Arndt von Haeseler (Universität München) for providing a prerelease version of PUZZLE 3.0 and helpful advice, Carla Kuiken (Los Alamos) and David Mindell (University of Michigan) for critical reading and revision of the manuscript, Nick Goldman (University of Cambridge) and David Hillis (University of Texas) for explaining the basics of quartet resolution, and Lucy Phillips for editorial wisdom. An anonymous reviewer was exceedingly helpful in coaxing the manuscript into a more informative and readable form. Thank you.

REFERENCES

1. **Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and E. A. Berger.** 1996. CC-CKR-5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. Science **272**:1955–1958.
2. **Amedee, A. M., N. Lacour, J. L. Gierman, L. N. Martin, J. E. Clements, R. Bohm, Jr., R. M. Harrison, and M. Murphey-Corb.** 1995. Genotypic selection of simian immunodeficiency virus in macaque infants infected transplacentally. J. Virol. **69**:7982–7990.
3. **Back, N. K., L. Smit, M. Schutten, P. L. Nara, M. Tersmette, and J. Goudsmit.** 1993. Mutations in human immunodeficiency virus type 1 gp41 affect sensitivity to neutralization by gp120 antibodies. J. Virol. **67**:6897–6902.
4. **Bandelt, H. J., and A. W. Dress.** 1992. Split decomposition: a new and useful approach to phylogenetic analysis of distance data. Mol. Phylogenet. Evol. **1**: 242–252.
5. **Baskin, G. B., L. N. Martin, M. Murphey-Corb, F. S. Hu, D. Kuebler, and B. Davison.** 1995. Distribution of SIV in lymph nodes of serially sacrificed rhesus monkeys. AIDS Res. Hum. Retroviruses **11**:273–285.
6. **Baskin, G. B., L. N. Martin, S. R. Rangan, B. J. Gormus, M. Murphey-Corb, R. H. Wolf, and K. F. Soike.** 1986. Transmissible lymphoma and simian acquired immunodeficiency syndrome in rhesus monkeys. JNCI **77**:127–139.
7. **Bohm, R. P., Jr., L. N. Martin, B. Davison-Fairburn, G. B. Baskin, and M. Murphey-Corb.** 1993. Neonatal disease induced by SIV infection of the rhesus monkey (Macaca mulatta). AIDS Res. Hum. Retroviruses **9**:1131–1137.
8. **Boom, R., C. J. Sol, M. M. Salimans, C. L. Jansen, P. M. Wertheim-van Dillen, and J. van der Noordaa.** 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. **28**:495–503.
9. **Bou-Habib, D. C., G. Roderiquez, T. Oravecz, P. W. Berman, P. Lusso, and M. A. Norcross.** 1994. Cryptic nature of envelope V3 region epitopes protects primary monocytotropic human immunodeficiency virus type 1 from antibody neutralization. J. Virol. **68**:6006–6013.
10. **Burns, D. P., C. Collignon, and R. C. Desrosiers.** 1993. Simian immunodeficiency virus mutants resistant to serum neutralization arise during persistent infection of rhesus monkeys. J. Virol. **67**:4104–4113.
11. **Chakrabarti, L., V. Baptiste, E. Khatissian, M. C. Cumont, A. M. Aubertin, L. Montagnier, and B. Hurtrel.** 1995. Limited viral spread and rapid immune response in lymph nodes of macaques inoculated with attenuated simian immunodeficiency virus. Virology **213**:535–548.
12. **Chesebro, B., K. Wehrly, J. Nishio, and S. Perryman.** 1992. Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical amino acids involved in cell tropism. J. Virol. **66**:6547–6554.
13. **Cocchi, F., A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, and P. Lusso.** 1995. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. Science **270**: 1811–1815.
14. **Courgnaud, V., F. Laure, P. N. Fultz, L. Montagnier, C. Brechot, and P. Sonigo.** 1992. Genetic differences accounting for evolution and pathogenicity of simian immunodeficiency virus from a sooty mangabey monkey after cross-species transmission to a pig-tailed macaque. J. Virol. **66**:414–419.
15. **Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton.** 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature **381**:667–673.
16. **Eigen, M., R. Winkler-Oswatitsch, and A. Dress.** 1988. Statistical geometry in sequence space: a method of quantitative comparative sequence analysis. Proc. Natl. Acad. Sci. USA **85**:5913–5917.
17. **Franchini, G., and M. S. Reitz, Jr.** 1994. Phylogenesis and genetic complexity of the nonhuman primate retroviridae. AIDS Res. Hum. Retroviruses **10**: 1047–1060.

17a. **Goldman, N.** Personal communication.

18. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73**:237–244.

18a. **Hillis, D. M.** 1996. Inferring complex phylogenies. Nature **383**:130–131.

18b. **Hillis, D. M.** Personal communication.

19. **Hirsch, V. M., G. Dapolito, P. R. Johnson, W. R. Elkins, W. T. London, R. J. Montali, S. Goldstein, and C. Brown.** 1995. Induction of AIDS by simian immunodeficiency virus from an African green monkey: species-specific variation in pathogenicity correlates with the extent of in vivo replication. J. Virol. **69**:955–967.
20. **Hirsch, V. M., and P. R. Johnson.** 1994. Pathogenic diversity of simian immunodeficiency viruses. Virus Res. **32**:183–203.
21. **Hirsch, V. M., J. E. Martin, G. Dapolito, W. R. Elkins, W. T. London, S. Goldstein, and P. R. Johnson.** 1994. Spontaneous substitutions in the vicinity of the V3 analog affect cell tropism and pathogenicity of simian immunodeficiency virus. J. Virol. **68**:2649–2661.
22. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson.** 1989. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature **339**:389–392.
23. **Kirchhoff, F., H. G. Morrison, and R. C. Desrosiers.** 1995. Identification of V3 mutations that can compensate for inactivating mutations in C4 of simian immunodeficiency virus. Virology **213**:179–189.
24. **Kirchoff, F., K. Mori, and R. C. Desrosiers.** 1994. The "V3" domain is a determinant of simian immunodeficiency virus cell tropism. J. Virol. **68**: 3682–3692.
25. **Kodama, T., K. Mori, T. Kawahara, D. J. Ringler, and R. C. Desrosiers.** 1993. Analysis of simian immunodeficiency virus sequence variation in tissues of rhesus macaques with simian AIDS. J. Virol. **67**:6522–6534.
26. **Kumar, S., K. Tamura, and M. Nei.** 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Comput. Applic. Biosci. **10**: 189–191.
27. **Leitner, T., D. Escanilla, C. Franzev, U. Mathias, and J. Albert.** 1996. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. Proc. Natl. Acad. Sci. USA **93**:10864–10869.
28. **Luciw, P. A., K. E. Shaw, R. E. Unger, V. Planelles, M. W. Stout, J. E. Lackner, E. Pratt-Lowe, N. J. Leung, B. Banapour, and M. L. Marthas.** 1992. Genetic and biological comparisons of pathogenic and nonpathogenic molecular clones of simian immunodeficiency virus (SIVmac). AIDS Res. Hum. Retroviruses **8**:395–402.
29. **Lukashov, V. V., C. L. Kuiken, and J. Goudsmit.** 1995. Intrahost human immunodeficiency virus type 1 evolution is related to length of the immunocompetent period. J. Virol. **69**:6911–6916.
30. **Margush, T., and F. R. McMorris.** 1981. Consensus n-trees. Bull. Math. Biol. **43**:239–244.
31. **Marthas, M. L., R. A. Ramos, B. L. Lohman, K. K. Van Rompay, R. E. Unger, C. J. Miller, B. Banapour, N. C. Pedersen, and P. A. Luciw.** 1993. Viral determinants of simian immunodeficiency virus (SIV) virulence in rhesus macaques assessed by using attenuated and pathogenic molecular clones of SIVmac. J. Virol. **67**:6047–6055.
32. **Miller, M. A., M. Murphey-Corb, and R. C. Montelaro.** 1992. Identification of broadly reactive continuous antigenic determinants of simian immunodeficiency virus glycoproteins. AIDS Res. Hum. Retroviruses **8**:1153–1164.
33. **Mindell, D. P.** 1996. Positive selection and rates of evolution in immunodeficiency viruses from humans and chimpanzees. Proc. Natl. Acad. Sci. USA **93**:3284–3288.

33a. **Mindell, D. P., J. W. Shultz, and P. W. Ewald.** 1995. The AIDS Pandemic is new, but is HIV new? Syst. Biol. **44**:77–92.

34. **Muller-Trutwin, M. C., S. Corbet, M. D. Tavares, V. M. A. Herve, E. Nerrienet, M. Georges-Courbot, W. Saurin, P. Sonigo, and F. Barre-Sinoussi.** 1996. The evolutionary rate of nonpathogenic simian immunodeficiency virus (SIVagm) is in agreement with a rapid and continuous replication in vivo. Virology **223**:89–102.
35. **Murphey-Corb, M., L. N. Martin, S. R. Rangan, G. B. Baskin, B. J. Gormus, R. H. Wolf, W. A. Andes, M. West, and R. C. Montelaro.** 1986. Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys. Nature **321**:435–437.
36. **Myers, G., K. MacInnes, and B. Korber.** 1992. The emergence of simian/human immunodeficiency viruses. AIDS Res. Hum. Retroviruses **8**:373–386.
37. **Novembre, F. J., P. R. Johnson, M. G. Lewis, D. C. Anderson, S. Klumpp, H. M. McClure, and V. M. Hirsch.** 1993. Multiple viral determinants contribute to pathogenicity of the acutely lethal simian immunodeficiency virus SIVsmmPBj variant. J. Virol. **67**:2466–2474.
38. **Overbaugh, J., and L. M. Rudensey.** 1992. Alterations in potential sites for glycosylation predominate during evolution of the simian immunodeficiency virus envelope gene in macaques. J. Virol. **66**:5937–5948.
39. **Overbaugh, J., L. M. Rudensey, M. D. Papenhausen, R. E. Benveniste, and**

W. R. Morton. 1993. Variation in simian immunodeficiency virus env is confined to V1 and V4 during progression to simian AIDS. J. Virol. 65: 7025–7031.
40. Pathak, V. K., and H. Temin. 1990. Broad spectrum of in vivo forward mutations, hypermutations and mutational hot spots in a retroviral vector after a single replication cycle: substitutions, frameshifts and hypermutations. Proc. Natl. Acad. Sci. USA. 87:6019–6023.
41. Reimann, K. A., K. Tenner-Racz, P. Racz, D. C. Montefiori, Y. Yasutomi, W. Lin, B. J. Ransil, and N. L. Letvin. 1994. Immunopathogenic events in acute infection of rhesus monkeys with simian immunodeficiency virus of macaques. J. Virol. 68:2362–2370.
42. Rudensey, L. M., J. T. Kimata, R. E. Benveniste, and J. Overbaugh. 1995. Progression to AIDS in macaques is associated with changes in the replication, tropism, and cytopathic properties of the simian immunodeficiency virus variant population. Virology 207:528–542.
43. Sakuragi, S., J. Sakuragi, and A. Adachi. 1995. Both SU and TM env proteins are responsible for monkey cell tropism of simian immunodeficiency virus SIV mac. Arch. Virol. 140:2255–2260.
44. Schwiebert, R., and P. N. Fultz. 1994. Immune activation and viral burden in acute disease induced by simian immunodeficiency virus SIVsmmPBj14: correlation between in vitro and in vivo events. J. Virol. 68:5538–5547.
45. Sellner, L. N., R. J. Coelen, and J. S. Mackenzie. 1994. Sensitive detection of Ross River virus-a one-tube nested RT-PCR. J. Virol. Methods 49:47–58.
46. Sharma, D. P., M. C. Zink, M. Anderson, R. Adams, J. E. Clements, S. V. Joag, and O. Narayan. 1992. Derivation of neurotropic simian immunodeficiency virus from exclusively lymphocyte tropic parental virus: pathogenesis of infection in macaques. J. Virol. 66:3550–3556.
47. Shpaer, E. G., and J. I. Mullins. 1993. Rates of amino acid change in the envelope protein correlate with pathogenicity of primate lentiviruses. J. Mol. Evol. 37:57–65.
48. Strimmer, K., N. Goldman, and A. von Haeseler. 1997. Bayesian probabilities and quartet puzzling. Mol. Biol. Evol. 14:210–211.
49. Strimmer, K., and A. von Haeseler. 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc. Natl. Acad. Sci. USA 94:6815–6819.
50. Strimmer, K., and A. von Haeseler. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. J. Mol. Evol. 13: 964–969.
51. Swofford, D. L. 1991. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.
52. Terrassette, M., R. A. Gruters, F. de Wolf, R. E. Y. de Goede, J. M. A. Lange, P. T. A. Schellekens, J. Goudsmit, H. G. Huisman, and F. Miedema. 1989. Evidence for a role of virulent human immunodeficiency (HIV) in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential HIV isolates. J. Virol. 63:2118–2125.
53. Valli, P. J. S., V. V. Lukashov, and J. Goudsmit. Shortening of the symptom-free period in rhesus macaques related to decreasing non-synonymous variation in the *env* variable regions of SIVsm during passage. Submitted for publication.
53a. Valli, P. J. S., J. L. Heeney, and J. Goudsmit. Selection for macrophage tropism during serial passage of SIVsm in rhesus macaques concomitant with increased virulence. Submitted for publication.
54. Wain-Hobson, S. 1996. Running the gamut of retroviral variation. Trends Microbiol. 4:135–141.
54a. Yang, Z., and N. Goldman. 1997. Are big trees indeed easy? Trends Ecol. Evol. 12:357.
55. Zhang, J. Y., L. N. Martin, E. A. Watson, R. C. Montelaro, M. West, L. Epstein, and M. Murphey-Corb. 1988. Simian immunodeficiency virus/delta-induced immunodeficiency disease in rhesus monkeys: relation of antibody response and antigenemia. J. Infect. Dis. 158:1277–1286.

Chapter 6

Increased heterogeneity of SIVsm *env* Associated with Low Viral Load in Macaca mulatta

Increased Heterogeneity of SIVsm *env* Associated with Low Viral Load in *Macaca mulatta*

P.J. Spencer Valli, Jonathan Heeney[1], and Jaap Goudsmit

Department of Human Retrovirology, Academisch Medisch Centrum, Amsterdam, [1]BPRC, P.O.Box 5815, 2280HV Rijswijk, The Netherlands

Department of Human Retrovirology
Academisch Medisch Centrum
Universiteit van Amsterdam
Meibergdreef 15, 1105 AZ Amsterdam
The Netherlands
telephone: (31-20) 566-4853
fax: (31-20) 691-6531
e-mail: j.goudsmit@amc.uva.nl

Abstract: Lentiviruses in primates susceptible to AIDS swarm as a quasispecies of genetic and biological variants which evolve as the infection persists. The development of AIDS is associated with the evolution of highly fit and highly virulent end-stage variants. To determine when virulent end-stage variants emerge in a susceptible host we correlated the virus load (prognostic for disease progression) with the degree of variation of the C1 to CD4 binding regions of the *env*. This was studied during the sequential in vivo passage of SIVsm DeltaB670 in a uniform group of age and sex matched rhesus monkeys (*Macaca mulatta*). Interestingly, the most marked sequence variation was observed in animals participating in earlier passages with low viral loads. Certain genotypic motifs found in the highly fit end-stage variants could be traced back to animals with low virus loads and a long survival history. These data suggest that the evolution of highly-fit end-stage variants begins during periods of low viral loads. They may be in part be driven by host immune responses early in infection until they become highly fit and capable of immune escape.

Introduction: The use of SIV in macaques allows for the analysis of virus of known sequence, and parameters or determinants of disease progression. The various strains of SIV differ in their infective and disease causing capacities (1,24) and in their various natural hosts as well as in experimental infections in other primate species (25,38). Patterns of disease have been widely investigated with studies on prior immunization and analysis of the progressors and slow progressors after infection to show that early replication patterns correlate with disease outcome (20,42). The route of transmission has also been shown to influence early viral selection and diversity (6,40). Virus load is a measure of disease progression in susceptible hosts (9,20,22,31,38,42). In disease resistant hosts viral loads as high as 10^5 to 10^7 copies per millilitre have been reported (38,42).

The ability of lentiviruses to absorb remarkable variations in their genomes and remain replication competent or fit is due to their quasispecies nature (17). The various copies of the SIV genome present during infection are all slightly different from one another due to the infidelity of their own reproductive machinery, and thus the virus may be more or less fit because of these variations. The size of the infectious inoculum is crucial as the variability of the virus can be detrimental if the inoculum is too small and only incompetent or less fit variants are transmitted alone (8,15). Increases in the size of the infectious inoculum have been shown to exponentially increase the virulence or pathogenicity of the virus during infection (8,35). Whereas dilute passages of virus, or lowered inoculum size promotes the expression of genetic and phenotypic variants (39) there is a finite limit to the extent to which the transmitted quasispecies, or inoculum can be reduced (8), limited reduction in the size of the transmitted quasispecies gives less prevalent or less replication competent variants a chance at the limited number of target cells. Experimental infection with a pathogenic clone can result in rapid progression as the initial burst of replication of a fit variant can overcome the effect of dose (23). Although quasispecies support the replication of the most fit variants by virtue of their quicker replication capacity (better able to compete to be the founder of the next round of replication or infective party to a new target cell) the sequence space within which they operate (17) is prone to disruption by the stochastic action of newly derived mutants of already fit variants (7). Although the derivation of new more fit variants is the basis for genetic variation during adaptation and infection the antigenic variation is not necessarily dependent upon immune selection as has long been proposed (13), but also to the action of the stochastic behaviour of large fit quasispecies (7). The combination of viral variation and immune response leads to a change in the biological properties of the during

infection and frequently a change in the tropism of the target cells (10), and the evolution of more virulent variants during progression to AIDS (19,28). The stabilization levels of viremia are predictive of the outcome of the infection (19,43), and increased viral load is a marker for a more rapid progression to AIDS.

To study the molecular evolution of a pathogen during cross species transfer the sooty mangabey strain of SIVsm was sequentially passaged six times (23) in Asian rhesus macaques. Pathogenesis was studied by following the fixation of *env* sequences, plasma viral load, disease characteristics, and length of the asymptomatic period during adaptation to a new host. The highest rates of evolution as measured by intrasample nonsynonymous variation were found in the initial passages and correlated with low virus load. These infections demonstrated that large amounts of nonsynonymous variation, or antigenic diversity are related to decreased replication kinetics, or less fit early passage variants of SIVsm. Eventually the selection and fixation of these nonsynonymous variations led to increased loads and high virulence.

Materials and Methods:

Virus. The SIV strain used in the multiple passage is of the SIVsm family of lentiviruses (SIVmac, SIVsmmH4 and SIVPBj14) originally discovered in sooty mangabeys or in captive Asian macaques infected accidentally with SIVsm (11,12,21). The SIVsm DeltaB670 strain used is well characterised and has been used in previous experimental infections of rhesus macaques because it accurately reproduces the clinical manifestations of AIDS seen in humans (2-4,18,26,32,34,37,41,49). The virus stock has been thoroughly analysed as to genotypic and phenotype variation during experimental infection of non-human primates and during in vitro infection of rhesus macaque and human PBMCs (2,30).

Passage. As reported elsewhere (23), the passage was carried out using 16 age matched Asian rhesus macaques (Macaca mulatta) of two years of age. The passage with SIVsmDeltaB670 was carried out using intravenous inoculation, with the first rhesus macaque receiving $5x10^2$ infectious doses of cell free virus. The following rhesus macaques (for P2 to P6A and P6B, there were two infections at the P6 level) were inoculated with $2x10^6$ uncultured cryopreserved peripheral blood mononuclear cells (PBMC) taken at the symptomatic stage of disease from the preceding rhesus macaque. Routine clinical biochemistry and haematological analysis was performed at two week intervals, and following two months, at four week intervals. The CD4+ T cell changes in concentration were monitored by FACScan (Becton Dickinson) using double labelling with Leu3a and 4B4, as described earlier (33). The concentrations of SIV p27 antigen were measured with an SIV p27 antigen capture assay (Coulter Corporation).

Clinical Findings During Serial Passage: The levels of plasma antigen (p27) and CD4+ T cells were measured at all available time points. Shortening of the asymptomatic period coincided with decreased viral diversity, greater plasma antigen levels, and a more rapid CD4+ T cell loss. The concentrations of SIVp27 and relative percentages of CD4+ T cells were monitored throughout the passage, no data was available for passage one (P1), except for the last serum sample, as used for RT-PCR amplification, cloning, and load determination. The P5 resulted in unexpectedly quick progression to AIDS, and no samples were available for haematological assay or for sequence, or load determination. The passage of SIVsm

DeltaB670 in juvenile rhesus monkeys resulted in the decrease of the asymptomatic period from eighteen months to one and two months, in the fifth and sixth passage respectively. Disease development was characterised by rapid loss of CD4+ T cells persistent and persistently high level plasma viremia (23). The pathological and clinical manifestations of disease did not change during the decrease in the clinical latency. The rapid reduction of the asymptomatic period occurs within three consecutive passages, dropping to weeks after P3 (23).

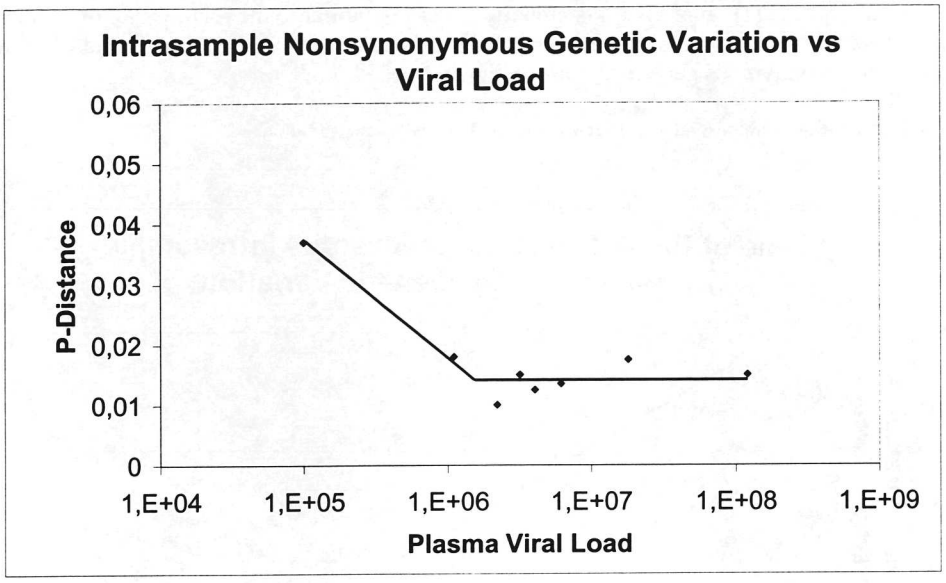
Viral load determination: Determination of SIV RNA levels in plasma of infected macaques was carried out using a highly sensitive and reproducible quantitative competitive (QC) RT-PCR assay (43). Briefly, 200 pl of plasma was added to 600 pl of guanidine-isothiocyanate-based lysis solution containing 300 copies of internal standard RNA. The RNA was precipitated by propanol-2 and was reversed transcribed and amplified with rTth DNA polymerase. The amplification products were hybridised in six five-fold dilutions to a capture probe that was that was covalently bound to microwell plates. The amplification products were detected by a streptavidin-horseradish-peroxidase-mediated calorimetric reaction. The amplified internal standard was hybridised to a rearranged 26-bp capture probe in separate microwells. The number of RNA copies in the plasma sample was calculated from the optical density of the sample wells compared to that of the corresponding internal standard. The number of RNA copies in the plasma sample was calculated from the optical density of the sample wells compared to that of the corresponding internal standard well. Data was plotted for each individual animal involved in the sequential *in vivo* passage as the concentration of RNA per ml of plasma (23).

***Env* RT-PCR, Cloning Strategy, and Sequencing:** Genomic viral RNA was extracted from macaque serum, amplified using RT-PCR and 60 *env* genes were sequenced form the V1 to V4, and analysed using phylogenetic analysis (44), and Ks/Ka variation (46). Briefly, viral RNA was extracted using silica beads and a chaotropic agent according to the Boom method (5), and used as template in a reverse transcriptase reaction followed by PCR amplification. The 3' PCR primer was also used to prime the reverse transcriptase reaction, followed by a single PCR reaction and cloning (44). The time points for sampling by sequencing of serum viral RNA were seroconversion and death. The sequenced clones were named by passage number (P1 to P6A and B) and start (S) or death (D). The sequences were aligned using CLUSTAL W(18) and adjusted by eye. Synonymous and nonsynonymous nucleotide variations (P-Distances) were calculated by using the MEGA program (29,46). Intrasample variation calculations are the result of comparisons of clones from the same time-point (within a sample).

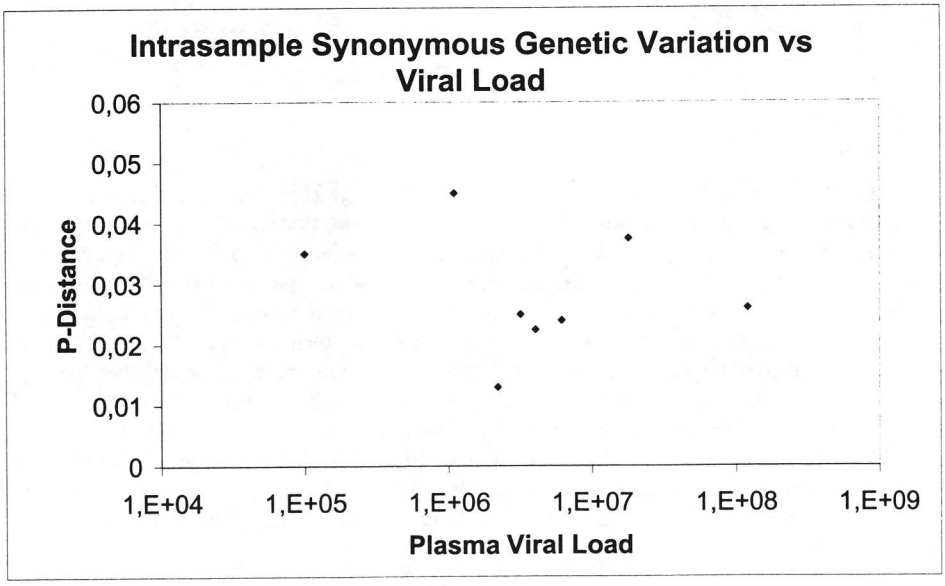
Results: Increasingly rapid progression to AIDS caused by the large inoculum size occurred in the first three passages. Macaques infected with only 50 MID_{50} of another nonpassaged SIVsm strain died between twelve and thirty weeks with loads of up to 10^9 per millilitre of plasma (20). During the first passage the virus at time of death showed a large amount of heterogeneity (46). During the second (and later) passages there was a very rapid progression to AIDS and death. As has been shown (23) the monkeys with extended TDPI (time of death post infection) showed a lower load and some containment of viremia before death, whereas the fast progressing monkeys showed no containment of viremia reaching a plateau at a high early peak load. The data were analysed according to TDPI, synonymous and nonsynonymous intrasample genetic variation, and virus load at peak, start (first available sample), and death.

Figure 1: Nonsynonymous (A), and Synonymous (B) P-Distances within the V1 to V4 sequenced *env* clones versus the virus viral load in the serum of the serially passaged SIVsm. P-Distances are the numbers of nucleotide substitutions between two strands of nucleic acid divided by the number of nucleotides they comprise in length, and gives an estimate of the average genetic heterogeneity.

A

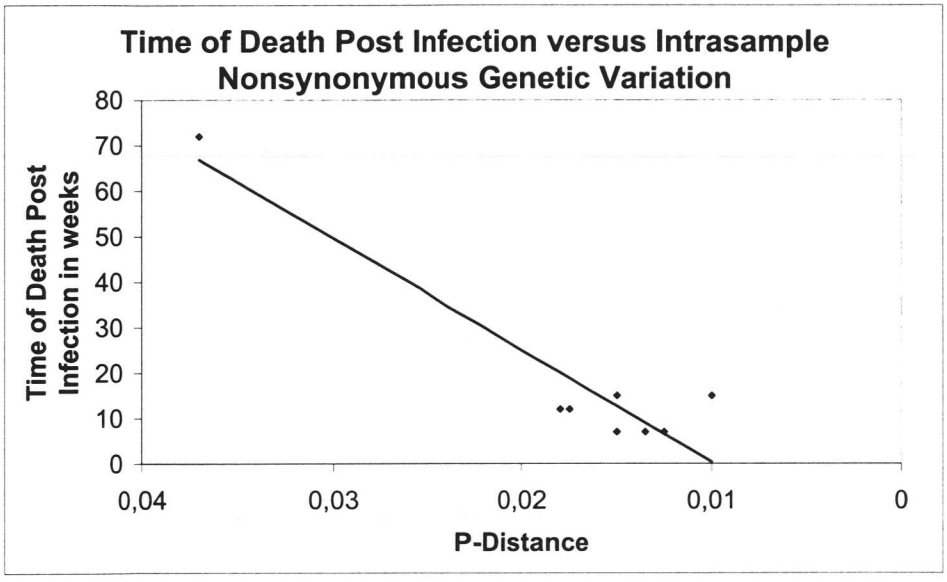


B



The nonsynonymous and synonymous genetic variation versus viral load (Figure 1A & B) follow different patterns, with the nonsynonymous variation being biphasic with a steep decline from the initially heterogenous primary inoculum at the lowest viral load. The synonymous variation follows no particular pattern when compared to either viral load (Figure 1A), or TDPI (similar pattern, not shown). The nonsynonymous genetic heterogeneity decreased almost three fold in the first passage and then remained at a low level thereafter when compared to virus load (Figure 1B); this biphasic pattern is similar to the relationship of TDPI and viral load (23). The increasing peak viral load follows a consistent decline in the TDPI and also in nonsynonymous genetic variation (Figure 2).

Figure 2: Time of death post infection versus nonsynonymous genetic variation.



Discussion: Lentiviral evolution during sequential passage is under the influence of population passage (8,35), purifying selection and the steady state variability or quasispecies of an adapted virus population (7,14,17), and antigenic variation which is not dependent upon immune selection (13,47). Sequential passage of a sooty mangabey derived SIVsm in rhesus macaques resulted in marked increases in the virulence of the virus. A genetic quasispecies with a wide range around the master sequence was found during adaptation and was initially associated with lower viral loads. The peak loads increased from 10^5 in the first passage to 10^9 in the end of the fifth passage. The genetic variation of the evolving virus decreased from the first passage to the third passage and remained at a basal level with a nominal amount of nonsynonymous variation (Figure 1A). Heterogeneity of the passaged virus was always present and displayed a quasispecies nature. The TDPI of the passages decreased in a similar pattern as the genetic variation and remained at an exceedingly short average after the fourth passage (Figure 2).

Rapid progression to AIDS occurred in the abscence of an effective immune response (viremia was controlled to some extent only in one monkey in passage two and three). Antigenic variation is decreased by the rapid adaptation and the selection for the most rapidly replicating viral variant. Rapid replication after the initial selection and dissemination preclude the immune system of producing even a weak or shortened suppression of viral replication.

A large number of seedling viruses used as the inoculum increases the chances that a better adapted or more rapidly replicating virus will exist in the inoculum population. Progeny will be selected as founders by the preceding rounds of replication and infection thereby raising the ability to replicate of the entire virus population. During decreasing TDPI and genetic variation of the quasispecies its phenotype narrows (15,16,36). The large number of genomes present overcomes the stochastic events of reverse transcriptase errors, sampling errors during fluctuating viral loads and immune or tissue specific selection as the quasi-species swarm evolves through sequence space towards elevated terrain (17). Nonsynonymous genetic variation is the hallmark of lentiviral infection and over time the heterogeneity or breadth of the quasispecies will alter according to selection and escape event, as will the master or consensus sequence encoding the aggregate of the quasispecies existing at any time. A virus is in constant competition due to the competitive exclusion principle (7). The complex mixture of unique antigenically distinct viral genomes compete as a most fit variant is selected over its nearest relatives (derived by errors of the reverse transcriptase) which may have some advantage, be neutral, or have some disadvantage, as compared to the other viruses present. Selection of fitness occurs by several processes of replication kinetics, cell tropisms, and immune escape ability. The sequence most fit (replicating to the highest levels, tropically advantaged, or more immune resistant) in a given host is at a peak of efficiency, and variation is of a limited value, as the virus would then shift to a lower terrain in the fitness landscape or lesser ability to replicate (17). The life cycle of SIV is dependent upon acclimatization to produce increased virulence (27) as is seen by the rapid and directed direction of the nonsynonymous variation.

The events that take place early in HIV/SIV infection have a distinct role in the outcome of the infection. Diverse profiles of plasma viremia occur during infection and progression and the early containment is a prognostic indicator of a long asymptomatic period, and early viremia is predictive of a more rapid disease progression (20,31,48). The emergence of cytopathic and antigenic variants influences the progression to disease (19,28). The selection early in the infection of antigenically distinct *env* fixations (nonsynonymous variations) may also be an indicator of progression (45). The viruses containing advantageous mutations early in infection are selected during the heterogeneic phase of the infection and fixed in a sequential fashion leading to a more fit quickly replicating variant (Figure 1A). The consensus sequence or intrasample nonsynonymous variation is seen to follow a biphasic pattern after the needed variations present in the primary inoculum are found, selected and fixed (45). These variants correlate directly with the viral load, whereas the synonymous or antigenically indistinct variations have no correlation (Figure 1B). This implies that the low virus load correlated *env* variations explored by the virus and fixed could be the cause of the increased load and virulence. The existence of the variations early in infection would lead one to believe that the development of pathogenesis is caused by the rapid replication (even in the earliest infection the load is 10^5 copies per ml of plasma), and continuous adaptation of the virus in selecting the virulence inducing substitutions in the *env* of the early passages until

they are all present in one genome. The study of these early infection selections could lead to the finding of *env* variations that could be used as a prognostic indicator. As demonstrated in this study the ability to replicate efficiently with decreasing genetic heterogeneity are key predictors of a rapid progression to AIDS.

References:

1. **Allan, J. S., M. Short, M. E. Taylor, S. Su, V. M. Hirsch, P. R. Johnson, G. M. Shaw, and B. H. Hahn.** 1991. Species-specific diversity among simian immunodeficiency viruses from African green monkeys. Journal of Virology **65**:2816-2828.

2. **Amedee, A. M., N. Lacour, J. L. Gierman, L. N. Martin, J. E. Clements, R. Bohm,Jr., R. M. Harrison, and M. Murphey-Corb.** 1995. Genotypic selection of simian immunodeficiency virus in macaque infants infected transplacentally. Journal of Virology **69**:7982-7990.

3. **Baskin, G. B., L. N. Martin, M. Murphey-Corb, F. S. Hu, D. Kuebler, and B. Davison.** 1995. Distribution of SIV in lymph nodes of serially sacrificed rhesus monkeys. AIDS Research & Human Retroviruses **11**:273-285.

4. **Bohm, R. P., Jr., L. N. Martin, B. Davison-Fairburn, G. B. Baskin, and M. Murphey-Corb.** 1993. Neonatal disease induced by SIV infection of the rhesus monkey (Macaca mulatta). AIDS Research & Human Retroviruses **9**:1131-1137.

5. **Boom, R., C. J. Sol, M. M. Salimans, C. L. Jansen, P. M. Wertheim-van Dillen, and J. van der Noordaa.** 1990. Rapid and simple method for purification of nucleic acids. Journal of Clinical Microbiology **28**:495-503.

6. **Bosch, M. L., A. Schmidt, M. B. Agy, L. E. Kimball, and W. R. Morton.** 1997. Infection of M. nemestrina neonates with HIV-1 via different routes of inoculation. AIDS **11**:1555-1563

7. **Clarke, D. K., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. Holland.** 1994. The red queen reigns in the kingdom of RNA viruses. Proceedings of the National Academy of Sciences of the United States of America **91**:4821-4824.

8. **Clarke, D. K., E. A. Duarte, A. Moya, S. F. Elena, E. Domingo, and J. Holland.** 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. Journal of Virology **67**:222-228.

9. **Coffin, J. M.** 1996. HIV viral dynamics. AIDS **10 Suppl 3:S75-84**:S75-84.

10. **Connor, R. I., K. E. Sheridan, D. Ceradini, S. Choe, and N. R. Landau.** 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1-Infected Individuals. Journal of Experimental Medicine **185**:621-628.

11. **Courgnaud, V., F. Laure, P. N. Fultz, L. Montagnier, C. Brechot, and P. Sonigo.** 1992. Genetic differences accounting for evolution and pathogenicity of simian immunodeficiency virus from a sooty mangabey monkey after cross-species transmission to a pig-tailed macaque. Journal of Virology **66**:414-419.

12. **Daniel, M. D., N. L. Letvin, N. W. King, M. Kannagi, P. K. Sehgal, R. D. Hunt, P. J. Kanki, M. Essex, and R. C. Desrosiers.** 1985. Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. Science **228**:1201-1204.

13. **Domingo, E., J. Diez, M. A. Martinez, J. Hernandez, A. Holguin, B. Borrego, and M. G. Mateu.** 1993. New observations on antigenic diversification of RNA viruses. Antigenic variation is not dependent on immune selection. Journal of General Virology **74**:2039-2045.

14. **Drake, J. W.** 1993. Rates of spontaneous mutation among RNA viruses. Proceedings of the National Academy of Sciences of the United States of America **90**:4171-4175.

15. **Duarte, E. A., D. K. Clarke, A. Moya, S. F. Elena, E. Domingo, and J. Holland.** 1993. Many-trillionfold amplification of single RNA virus particles fails to overcome the Muller's ratchet effect. Journal of Virology **67**:3620-3623.

16. **Duarte, E. A., I. S. Novella, S. Ledesma, D. K. Clarke, A. Moya, S. F. Elena, E. Domingo, and J. J. Holland.** 1994. Subclonal components of consensus fitness in an RNA virus clone. Journal of Virology **68**:4295-4301.

17. **Eigen, M.** 1993. Viral quasispecies. Scientific American **269**:42-49.

18. **Higgins, D. G. and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73**:237-244.

19. **Hirsch, V. M.** 1999. Evolution of the fittest ends in tragedy. Nature Medicine **5**:488-489.

20. **Hirsch, V. M., T. R. Fuerst, G. Sutter, M. W. Carroll, L. C. Yang, S. Goldstein, M. Piatak,Jr., W. R. Elkins, W. G. Alvord, D. C. Montefiori, B. Moss, and J. D. Lifson.** 1996. Patterns of viral replication correlate with outcome in simian immunodeficiency virus (SIV)-infected macaques: effect of prior immunization with a trivalent SIV vaccine in modified vaccinia virus Ankara. Journal of Virology **70**:3741-3752.

21. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson.** 1989. An African

primate lentivirus (SIVsm) closely related to HIV-2. Nature **339**:389-392.
22. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz.** 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature **373**:123-126.
23. **Holterman, L., H. Niphuis, P. ten Haaft, J. Goudsmit, G. B. Baskin, and J. L. Heeney.** 1999. Specific Passage of Simian Immunodeficiency Virus from End-Stage Disease Results in Accelerated Progression to AIDS in Rhesus Macaques. Journal of General Virology 3089-3097.
24. **Johnson, P. R. and V. M. Hirsch.** 1992. Genetic variation of simian immunodeficiency viruses in nonhuman primates. AIDS Research & Human Retroviruses **8**:367-372.
25. **Kaur, A., R. M. Grant, R. E. Means, H. McClure, M. Feinberg, and R. P. Johnson.** 1998. Diverse host responses and outcomes following simian immunodeficiency virus SIVmac239 infection in sooty mangabeys and rhesus macaques. Journal of Virology **72**:9597-9611.
26. **Khabbaz, R. F., T. Rowe, M. Murphey-Corb, W. M. Heneine, C. A. Schable, J. R. George, C. P. Pau, B. S. Parekh, M. D. Lairmore, and J. W. Curran.** 1992. Simian immunodeficiency virus needlestick accident in a laboratory worker. Lancet **340**:271-273.
27. **Kimata, J. T., J. J. Gosink, V. N. KewalRamani, L. M. Rudensey, D. R. Littman, and J. Overbaugh.** 1999. Coreceptor specificity of temporal variants of simian immunodeficiency virus Mne. Journal of Virology **73**:1655-1660.
28. **Kimata, J. T., L. Kuller, D. B. Anderson, P. Dailey, and J. Overbaugh.** 1999. Emerging cytopathic and antigenic simian immunodeficiency virus variants influence AIDS progression. Nature Medicine **5**:535-541.
29. **Kumar, S., K. Tamura, and M. Nei.** 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Computer Applications in the Biosciences **10**:189-191.
30. **Martin Amedee, A., N. Lacour, L. N. Martin, J. E. Clements, R. B. Bohm,Jr., B. Davison, R. Harrison, and M. Murphey-Corb.** 1996. Genotypic analysis of infant macaques infected transplacentally and orally. Journal of Medical Primatology **25**:225-235.
31. **Mellors, J. W., C. R. Rinaldo,Jr., P. Gupta, R. M. White, J. A. Todd, and L. A. Kingsley.** 1996. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science **272**:1167-1170.
32. **Miller, M. A., M. Murphey-Corb, and R. C. Montelaro.** 1992. Identification of broadly reactive continuous antigenic determinants of simian immunodeficiency virus glycoproteins. AIDS Research & Human Retroviruses **8**:1153-1164.
33. **Murphey-Corb, M., L. N. Martin, B. Davison-Fairburn, and et.al.** 1989. A formalin-inactivated whole SIV vaccine confers protection in macaques. Science **246**:1293-1297.
34. **Murphey-Corb, M., L. N. Martin, S. R. Rangan, G. B. Baskin, B. J. Gormus, R. H. Wolf, W. A. Andes, M. West, and R. C. Montelaro.** 1986. Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys. Nature **321**:435-437.
35. **Novella, I. S., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. J. Holland.** 1995. Exponential increases of RNA virus fitness during large population transmissions. Proceedings of the National Academy of Sciences of the United States of America **92**:5841-5844.
36. **Novella, I. S., S. F. Elena, A. Moya, E. Domingo, and J. J. Holland.** 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. Journal of Virology **69**:2869-2872.
37. **Novembre, F. J., P. R. Johnson, M. G. Lewis, D. C. Anderson, S. Klumpp, H. M. McClure, and V. M. Hirsch.** 1993. Multiple viral determinants contribute to pathogenicity of the acutely lethal simian immunodeficiency virus SIVsmmPBj variant. Journal of Virology **67**:2466-2474.
38. **Rey-Cuille, M. A., J. L. Berthier, M. C. Bomsel-Demontoy, Y. Chaduc, L. Montagnier, A. G. Hovanessian, and L. A. Chakrabarti.** 1998. Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. Journal of Virology **72**:3872-3886.
39. **Sanchez-Palomino, S., J. M. Rojas, M. A. Martinez, E. M. Fenyo, R. Najera, E. Domingo, and C. Lopez-Galindez.** 1993. Dilute passage promotes expression of genetic and phenotypic variants of human immunodeficiency virus type 1 in cell culture. Journal of Virology **67**:2938-2943.
40. **Sodora, D. L., F. Lee, P. J. Dailey, and P. A. Marx.** 1998. A genetic and viral load analysis of the simian immunodeficiency virus during the acute phase in macaques inoculated by the vaginal route. AIDS Research & Human Retroviruses **14**:171-181.
41. **Spira, A. I., P. A. Marx, B. K. Patterson, J. Mahoney, R. A. Koup, S. M. Wolinsky, and D. D. Ho.** 1996. Cellular targets of infection and route of viral dissemination after an intravaginal inoculation of simian immunodeficiency virus into rhesus macaques. Journal of Experimental Medicine **183**:215-225.
42. **Staprans, S. I., P. J. Dailey, A. Rosenthal, C. Horton, R. M. Grant, N. Lerche, and M. B. Feinberg.** 1999. Simian immunodeficiency virus disease course is predicted by the extent of virus replication during primary infection. J. Virol. **73**:4829-4839.
43. **ten Haaft, P., B. Verstrepen, K. Uberla, B. Rosenwirth, and J. Heeney.** 1998. A pathogenic threshold of

virus load defined in simian immunodeficiency virus- or simian-human immunodeficiency virus- infected macaques. Journal of Virology **72**:10281-10285.
44. **Valli, P. J. S. and J. Goudsmit**. 1998. Structured Tree Topology and Adaptive Evolution of the SIVsm Envelope During Serial Passage in Rhesus Macaques According to Likelihood Mapping and Quartet Puzzling. Journal of Virology **72**:3673-3683.
45. **Valli, P. J. S., J. L. Heeney, and J. Goudsmit**. 2000. Genotypic Variation in the SIVsm DeltaB670 *Env* During the Shift to High Virulence. submitted
46. **Valli, P. J. S., V. V. Lukashov, and J. Goudsmit**. 1998. Shortening of the symptom-free period in rhesus macaques related to decreasing non-synymous variation in the env variable regions of SIVsm during passage. Journal of Virology **72**:7494-7500.
47. **Wain-Hobson, S.** 1996. Running the gamut of retroviral variation. Trends in Microbiology **4**:135-141.
48. **Watson, A., J. Ranchalis, B. Travis, J. McClure, W. Sutton, P. R. Johnson, S. L. Hu, and N. L. Haigwood**. 1997. Plasma viremia in macaques infected with simian immunodeficiency virus: plasma viral load early in infection predicts survival. Journal of Virology **71**:284-290.
49. **Zhang, J. Y., L. N. Martin, E. A. Watson, R. C. Montelaro, M. West, L. Epstein, and M. Murphey-Corb**. 1988. Simian immunodeficiency virus/delta-induced immunodeficiency disease in rhesus monkeys: relation of antibody response and antigenemia. Journal of Infectious Diseases **158**:1277-1286.

Chapter 7

General Disscussion

General Discussion:

The structure and function of viral genes and their interaction with cellular factors greatly influence the events during infection. Viral variation of the *env* in human infections is governed by both functional and epidemiological variables. It is difficult to trace due to the usually unknown infectious inoculum, and the detection of infection well past seroconversion. In order to study the human infection we have used an analogous animal model. SIV is African in origin, and in Asian non-human primates it induces AIDS, where as in its "natural host" species of African origin it is nonpathogenic. The infection in the African nonhuman primates is species specific and there are now eight separate strains of SIV which have been characterized (11,12,14). The SIV strain most closely related to human immunodeficiency virus type 2 (HIV-2) is SIVsm, the strain specific to the sooty mangabey (Cercocebus atys (8)). The experimental infection of Asian primates with SIV allows the study of the viral infection with a known strain of the virus from the moment of infection. The sequencing of large numbers of clones of the *env* reveals the genetic imprint made on the virus in time.

A preponderonce of nonsynonymous nucleic acid changes is seen as evidence of adaptive evolution (classically known as Darwinian evolution, or the opposite form of selection to the neutral theory (13)). The high rate of amino acid change belies a biological pressure upon the proteins translated from the nucleic acid of the organism, and from there a selection from the pool of available genomes (which is very large in SIV infection). The reference to the competing models (Darwinian and neutral) is important as the evolution seen during infection of SIVsm in a new species is rapid and punctuated (20). Evolution of the *env* takes place at an extremely rapid rate during the initial infections but then moves towards a steady state of low variation. The two modes of variation (synonymous and nonsynonymous) are shown to vary dramatically according to the time after cross-species transmission. The decrease in variation coincides with the increase in virulence, and the markedly shortened asymptomatic period before progression to AIDS. Variation in codon usage between the variable (external), and the constant (internal) regions allows the virus to adjust at different rates in different regions to changes caused by mistakes in reverse transcription, and may be a possible mechanism of increasing amino acid variation to escape immune selection.

The clustering of *env* sequences according to passage and pathogenicity shows that variation found during passage takes place in a structured manner. Likelihood mapping (17,19) showed that phylogenetic reconstruction of the passage was possible because a high percentage of the *env* nucleic acids sequenced contained tree-like data. The use of quartet puzzling (18) to produce the phylogenies showed a correlation between the *env* sequences, number of passages, and increased pathogenicity.

Originally the T cell CD4 receptor was thought to be the only mediator of binding and entry of HIV and SIV. Inconclusive inhibition of binding and entry by Abs led to the belief that there must be other cellular receptors in use by these viruses, and the cloning of the first and recognition of the other "orphan" coreceptors was the beginning of the understanding of tropism during infection. What were termed coreceptors are now known to be able to support binding and infection without the use of CD4 in some cases. The genetic variation in the *env* and its biological significance was shown to differ with the situation in HIV-1 in humans (5), where the *env* follows similar patterns of variation but the biological pattern of infection

varies with the switch from CCR5 use to CXCR4 use late in infection. Switches in coreceptor use from CCR5 (macrophage tropic virus) to CXCR4 (T cell tropic virus) herald the progression to AIDS in some cases (5,7). In the SIVsm transmission there was no change in coreceptor usage throughout the passage. The variation in the *env* would lead one to believe that it played a role in the increased pathogenicity. The lack of variation of coreceptor usage may mean the changes seen in the *env* are for increased binding of the receptor, and or coreceptors used. The charted changes in the *env* show that in later passages some clones are found that were not seen in the earlier passages, but were present in the primary inoculum (1). Large quasispecies can dominate, or suppress, fitter variants (6) simply by occupying the available target cells. In *in vitro* studies decreasing the size of the quasispecies, or dilute passage, can bring out variants not seen earlier in infection (16). The decreasing heterogeneity of the quasispecies during passage did not obscure the earlier hidden (or minority and thus not sampled) variants. In passage three (P3S-3, P3D-1 & 3) variants known to exist in the primary inoculum were present at a percentage high enough to be found during isolation from the serum.

Viral load is a clinical marker of the health of the HIV infected, especially when under drug, or multi-drug treatment, a rise in viral load can mean the beginning of the failure of intervention as mutants arise able to replicate while under treatment. During the serial passage the virus load increased until the third passage. The virus load increased as the genetic heterogeneity of the quasispecies decreased. Population passages cause vast gains in growth rates during each large quasispecies infection (4,15). The large number of genomes present increases the chances that there is a very fit variant in the inoculum, the stochastic nature of virus growth (21) leads to a competition between the most fit variants. The increased homogeneity seen during passage may be an indicator of greater intrinsic differences seen between variants present than in HIV infection (21), which allow for the outgrowth of the major variant in spite of the complex relationship of replicative ability, host cell, and cellular stimulation.

The wild populations of primates represent an unexplored reservoir of evolving and emerging infectious diseases that can spread to man (3,9,10,14,22). This reservoir can be used to serve as an indicator of coming pathogens and provide models of natural transmission dynamics. The low frequency of seropositivity in the wild population may be due to predation effects. The seemingly nonpathogenic infection may possibly make the infected non-human primates more opportune to other pathogens, or predators, thus removing them from the population before the virus could spread to the remaining troupe members. It may be that the feral populations have already been through the selection of an "ancient infection" (2). The genetic pool of African nonhuman primates surviving today may be the descendants of ancestors who were long term nonprogressors, or had resistance genes able to ensure asymptomatic infection and survival. There is a discontinuity between the numbers of infected feral monkeys found and the nonmonogamous mating habits of these species which should lead to a higher seroprevalence. Elucidation of the transmission patterns and serosurvey of the feral species is seemingly impossible now with the large scale destruction of their territories, and decreasing population size. Any advances made in understanding the infection among the nonhuman primates, and its natural history, might help to elucidate its spread to man. Wild primate populations remain a source of emerging human virus infection.

References:

1. **Amedee, A. M., N. Lacour, J. L. Gierman, L. N. Martin, J. E. Clements, R. Bohm,Jr., R. M. Harrison, and M. Murphey-Corb.** 1995. Genotypic selection of simian immunodeficiency virus in macaque infants infected transplacentally. Journal of Virology **69**:7982-7990.

2. **Brown, E. W., N. Yuhki, C. Packer, and S. J. O'Brien.** 1994. A lion lentivirus related to feline immunodeficiency virus: epidemiologic and phylogenetic aspects. Journal of Virology **68**:5953-5968.

3. **Chen, Z., P. Telfer, A. Gettie, P. Reed, L. Zhang, D. D. Ho, and P. A. Marx.** 1996. Genetic Characterization of New West African Simian Immunodeficiency Virus SIVsm: Geographic Clustering of Household-Derived SIV Strains with Human Immunodeficiency Virus Type 2 Subtypes and Genetically Diverse Viruses from a Single Feral Sooty Mangabey Troop. Journal of Virology **70**:3617-3627.

4. **Clarke, D. K., E. A. Duarte, A. Moya, S. F. Elena, E. Domingo, and J. Holland.** 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. Journal of Virology **67**:222-228.

5. **Connor, R. I., K. E. Sheridan, D. Ceradini, S. Choe, and N. R. Landau.** 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1-Infected Individuals. Journal of Experimental Medicine **185**:621-628.

6. **de la Torre, J. C. and J. J. Holland.** 1990. RNA virus quasispecies populations can suppress vastly superior mutant progeny. J. Virol. **64**:6278-6281.

7. **de Roda Husman, A. M., R. P. van Rij, H. Blaak, S. Broersen, and H. Schuitemaker.** 1999. Adaptation to Promiscuous Usage of Chemokine Receptors Is Not a Prerequisite for Human Immunodeficiency Virus Type 1 Disease Progression. J. Infect. Dis. **180**:1106-1115.

8. **Desrosiers, R. C., M. D. Daniel, N. L. Letvin, N. W. King, and R. D. Hunt.** 1987. Origins of HTLV-4. Nature **327**:107

9. **Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature **397**:436-441.

10. **Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp.** 2000. AIDS as a zoonosis: scientific and public health implications. Science 2000. Jan. 28;287(5453):607-14. **287**:607-614.

11. **Hirsch, V. M., G. A. Dapolito, S. Goldstein, H. McClure, P. Emau, P. N. Fultz, M. Isahakia, R. Lenroot, G. Myers, and P. R. Johnson.** 1993. A distinct African lentivirus from Sykes' monkeys. Journal of Virology **67**:1517-1528.

12. **Hirsch, V. M., R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, and P. R. Johnson.** 1989. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature **339**:389-392.

13. **Kimura, M.** 1977. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature **267**:275-276.

14. **Marx, P. A., Y. Li, N. W. Lerche, S. Sutjipto, A. Gettie, J. A. Yee, B. H. Brotman, A. M. Prince, A. Hanson, and R. G. Webster.** 1991. Isolation of a simian immunodeficiency virus related to human immunodeficiency virus type 2 from a west African pet sooty mangabey. Journal of Virology **65**:4480-4485.

15. **Novella, I. S., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. J. Holland.** 1995. Exponential increases of RNA virus fitness during large population transmissions. Proceedings of the National Academy of Sciences of the United States of America **92**:5841-5844.

16. **Sanchez-Palomino, S., J. M. Rojas, M. A. Martinez, E. M. Fenyo, R. Najera, E. Domingo, and C. Lopez-Galindez.** 1993. Dilute passage promotes expression of genetic and phenotypic variants of human immunodeficiency virus type 1 in cell culture. Journal of Virology **67**:2938-2943.

17. **Strimmer, K., N. Goldman, and A. von Haeseler.** 1997. Bayesian Probabilities and Quartet Puzzling. Molecular Biology & Evolution **14**:258-259.

18. **Strimmer, K. and A. von Haeseler.** 1996. Quartet Puzzling: A Quartet Maximum-Likelihood Method for Reconstructing Tree Topologies. Journal of Molecular Evolution **13**:964-969.

19. **Strimmer, K. and A. von Haeseler.** 1997. Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. Proceedings of the National Academy of Sciences **94**:6815-6819.

20. **Valli, P. J. S., V. V. Lukashov, and J. Goudsmit.** 1998. Shortening of the symptom-free period in rhesus macaques related to decreasing non-synymous variation in the *env* variable regions of SIVsm during passage. Journal of Virology **72**:7494-7500.

21. **Wain-Hobson, S.** 1996. Running the gamut of retroviral variation. Trends in Microbiology **4**:135-141.

22. **Wolfe, N. D., A. A. Escalante, W. B. Karesh, A. Kilbourn, A. Spielman, and A. A. Lal.** 1998. Wild primate populations in emerging infectious disease research: the missing link? Emerg. Infect. Dis. **4**:149-158.

Chapter 8

Summary

Samenvatting

Summary:

The introduction of this thesis describes the history of the AIDS pandemic and the viruses discovered almost simultaneously in man and in non-human primates in two primate colonies on opposite coasts of the United States. The discovery of the causative agent followed the advancing "syndrome" by several years and the link between HIV and SIV was slowly brought to light after much confusion as to how a human epidemic could have spread in North America, while the virus reservoir was in African non-human primates. Originally thought to be a disease of homosexuals, it's spread to haemophiliacs, blood bank recipients and the heterosexual population showed how incipient it was. The development of an animal model for use in vaccine studies was complicated by the apparent large genetic differences between HIV-1 in man and the SIV found in the nonhuman primate colonies. The discovery of an HIV-1-like virus in chimpanzees and the realization that HIV-2 originated in sooty mangabeys brought about the change in research direction that led to today's studies.

Chapter two is a paper about the natural infection in nonhuman primates. We used the known strains of SIV as well as newly cloned and analysed SIVrcm, L'hoest, Sun, and CpzUS to gain insight into the phylogeny of their *env* gp160 and *gag* genes to study the patterns of evolution of the immune reactive outer proteins and the more core proteins. The L'hoest virus is seen to be the most likely progenitor of the present pandemic, and all of the HIV/SIV species.

Chapter three is a study of the role of the *env* adaptations in cellular tropism, and virulence. No changes in glycosylation, charge, hydrophobicity or secondary structure prediction were seen between the *env* clones of the early versus the late passages. Motifs of macrophage tropic B670 within the *env* V1 were selected and fixed. Present in the primary inoculum they became the majority of the virus sampled in the later passages. During the increase in virulence and decrease in *env* evolution there were no alterations in coreceptor CCR5 being used exclusively. These results show that the increased pathogenicity and *env* variation are not related to coreceptor switching in SIVsm infection.

Chapter four describes the serial passage of SIVsm from sooty mangabeys in rhesus macaques, which drastically shortened the time of disease progression from 1.5 years to 1 month as the virus adapted to its new Asian host. The use of a well studied virus isolate in *in vivo* experimental infection allows the comparison of the disease characteristics and the genetic variation during transmission to a new species and passage in that species. To study SIVsm *env* evolution during passage in rhesus macaques, multiple *env* clones were sequenced at seroconversion and again at death. The *env* variation found during adaptation was almost completely confined to the variable or exposed regions. Shortening of the asymptomatic period coincided with a rise in the Ks/Ka ratio (ratio between the number of synonymous [Ks] and the number of nonsynonymous [Ka] substitutions). The Ks/Ka ratio followed the changes in the disease characteristics with most of the adaptation occurring in the first three passages followed by near stasis. Analysis of codon usage revealed decreased codon redundancy in the variable regions.

The quasispecies nature of the SIVsm *env* was studied in chapter five. We obtained phylogenetic evidence provided by the structured topology of the SIVsm *env* tree. Likelihood mapping showed that phylogenetic reconstruction of the passage was possible because a high percentage of the sequence data had a "tree-like" form. Subsequently, quartet puzzling was used and produced a phylogeny with a structure parallel to the known infection history. The adaptation of

SIVsm to Asian rhesus macaques appears to be an ordered process in which the *env* evolves in a tree-like manner, particularly in its constant regions.

In the sixth chapter we compare the measured virus load during increased virus virulence. In feral sooty mangabeys SIVsm is nonpathogenic while replicating to high copy numbers (as does SIVagm in African Green monkeys). We studied the effect of multiple passages on virus load and genetic variation. The virus load from multiple monkeys at each passage was measured to assess the relationship between pathogenicity and virus load. Nonsynonymous mutations were fixed in the sequential passages and present in the early passages. The selected and fixed mutations related to increased virulence existed in the primary inoculum, but not in one viral genome. Shortening of the disease free period was shown to be correlated with increased viral load and decreasing genetic heterogeneity.

Finally the conclusion to this thesis is the discussion of the findings and a chance to philosophize about the possible ramifications.

Samenvatting:

In hoofdstuk één wordt een overzicht gegeven van de geschiedenis van de pandemische HIV-infectie en de gelijktijdige ontdekking van een virus (SIV) bij apen in twee apen kolonies in de Verenigde Staten. De ontdekking van het HIV-virus kwam een aantal jaren na het beschrijven van het in toenemende mate voorkomende AIDS syndroom. De relatie tussen HIV en SIV kwam langzaam aan het daglicht nadat veel verwarring was ontstaan over hoe een virus van een Afrikaanse aap een epidemie bij de mens in de VS had kunnen veroorzaken. Terwijl men aanvankelijk dacht dat het een infectie was die alleen voorkwam bij homosexuelen, bleek dat de verspreiding onder patiënten die behandeld waren met bloedproducten aantoonde dat HIV ook met bloed overdraagbaar was en in Afrika werd verspreid door middel van heterosexuele overdracht. De ontwikkeling van een diermodel werd gecompliceerd door de grote genetische verschillen tussen het HIV-1 in mensen en het SIV gevonden in bovengenoemde apenkolonies. De ontdekking van een HIV-1achtig virus in chimpansees (SIVcpz) en vervolgens de ontdekking dat HIV-2 afkomstig was van sooty mangabeys (SIVsm) was de doorbraak die leidde tot het hedendaagse onderzoek.

In hoofdstuk twee wordt een vergelijkende studie van het SIV beschreven. Wij hebben gebruik gemaakt van bekende SIV stammen en de in 1999 ontdekte virussen L'hoest, Sun en CpzUS om de stamboom te analyseren van *env* gp160 en *gag* structurele genen met als doel het bestuderen van patronen van evolutie van het virus. Uit dit onderzoek bleek dat het L'hoest virus naar alle waarschijnlijkheid de voorouder van de hedendaagse HIV- en SIV-soorten is.

In het derde hoofdstuk is een studie beschreven over het gebruik van co-receptoren door het virus in rhesusmakaken tijdens een infectie met SIVsm. Er werden geen veranderingen in glycosylering, lading, hydrofobiciteit of de voorspellende secundaire structuur gezien tussen de verschillende passages van het virus. Bepaalde macrofaag-tropische B670 motieven binnen het eerste variabele gebied (V1) van het *env* gen, aanwezig als een minderheid in het primaire inoculum, bleken het leeuwedeel van de motieven te vormen na meerdere passages. Tijdens de toename in virulentie en afname van genetische variatie in het *env* gen gedurende de eerste drie passages van dit virus werden er geen veranderingen in CCR5 co-receptor gebruik gevonden. Deze resultaten wijzen erop dat een toename van de pathogeniciteit niet van invloed is op het co-receptor gebruik tijdens een infectie met SIVsm.

Hoofdstuk vier van dit proefschrift beschrijft het effect van seriële passage van SIVsm - afkomstig uit sooty mangabeys - in rhesusmakaken waardoor de incubatietijd van de ziekte werd verkort van 1.5 jaar tot 1 maand als uiting van de aanpassing van het virus aan de nieuwe - uit Azië afkomstige - gastheer. Het gebruik van een goed bestudeerd virus voor *in vivo* infecties maakt het mogelijk de genetische variatie tijdens transmissie naar een nieuwe soort te bestuderen in relatie tot de pathogeniciteit van het virus. Om de evolutie van het SIVsm omhulsel bij resus makaken te bestuderen werd van een aantal viruspreparaten de RNA sequentie bepaald op het moment van seroconversie en op het moment van overlijden. De variatie binnen het *env* gen tijdens het adaptatieproces was vrijwel volledig beperkt tot de gebieden die zich aan de buitenzijde van het envelop eiwit bevinden. Verkorting van de asymptomatische periode viel samen met met een verhoging van de Ks/Ka ratio (de verhouding tussen het aantal synonieme [Ks] en niet synonieme [Ka] substituties). De Ks/Ka ratio daalde en ging gepaard met een versneld ziekte proces gedurende de eerste drie passages. Tijdens deze passages vond er een verandering plaats van een heterogene tot een

vrijwel homogene virus populatie. Analyse van het codongebruik liet zien dat het aantal codons dat werd gebruikt voor het coderen van bepaalde aminozuren daalde na de passage.

Het kwasi-soortspecifieke karakter van het *env* gen van SIVsm is beschreven in hoofdstuk vijf. De vraagstelling was of er phylogenetisch een verband gelegd kon worden tussen de variatie in env en veranderingen in de ziekteprogressie. Likelihood-mapping en quartet-puzzling toonden vervolgens aan dat de viruspopulatie binnen elke passage correleerde met een specifieke plek op de phylogenetische boom en het daarbij behorende ziekteverloop. De adaptatie van SIVsm in Aziatische rhesusmakaken bleek een gestructureerd proces te zijn waarin het geconserveerde gebied van het *env* gen veranderde in phylogenetische clusters die correleerden met de virulentie van de viruspopulatie tijdens desbetreffende passage.

Bij in het wild levende sooty mangabeys is SIVsm niet pathogeen terwijl zeer hoge hoeveelheden virus voorkomen in het bloed zoals ook gevonden is in SIVagm geïnfecteerde African green monkeys. In hoofdstuk zes werd het effekt van meerdere virus passages op de hoeveelheid virus en de genetische variatie bestudeerd. De gemeten hoeveelheid virus in het bloed steeg in het algemeen tijdens de passages met een maximale waarde van 10^{10} copieën SIVsm RNA per milliliter serum terwijl de virus variatie daalde. Een verkorting van de levensduur na infectie hing samen met een afname van genetische variatie en een toename van de hoeveelheid virus in het bloed.

In hoofdstuk zeven, de algemene discussie, worden de bevindingen van de in dit proefschrift beschreven experimenten bediscussieerd.

Dankwoord:

My promotor Prof. Dr. J. Goudsmit and copromotor J.L. Heeney made this thesis possible.

I started on this thesis project with Margreet Bakker and we quickly found that it was a daunting endeavour. Rene Boom and Cees Sol took over from there and taught me everything I needed to know to complete the project molecularly. Your enthusiasm for science is contagious even if you weren't thrilled with the project itself. Once on the way it became clear as a foreigner that wading through the immigration and contractual legalities was more than a handful and Jantien Manders and Mieke de Liefde were always there with an open door. Carla Kuiken taught me all I know about the process of creating a phylogeny and it's pitfalls and together with Lucas Bowler you both led to my first paper which won the Kiem prize from the Netherlands Vereniging van Microbiologie. Thank you for your patience. I didn't always get it right the first time.

Margreet, Rene and Maarten, Jeroen, Bep, John, Margreeth, Miriam, Lia, Karel, and Raymond, you made the lab run, the sequencers sequence, and the computers compute. Which wasn't always easy. The needed scientific discussions were shared often with Ben, Ineke, Koen, Tonja, Patrizia, Bill, Charles, and Jacques, you all contributed to this work with your ideas. Michiel Huijberts and Rob Wolfram need applause for all the mainframe time and for tutoring me in UNIX.

Thanks to the members of the "promotie commissie" who took part in the assessment of the thesis.

Wouter Leene, and Lennart Bouman thanks for the guidance of yours needed to complete this thesis as it will never be forgotten, As a foreigner the system in the AMC, and at the UvA was not an easy one to get through without your help.

Wim van Est was superb in his patience with seemingly never ending changes in figures and diagrams. Lucy D. Phillips skill in editing is enormous and I enjoyed all the jokes we enjoyed about the Canadiana of my writing.

The foreign scientists I have not yet met but who were crucial in turning my manuscripts into scientific papers were Vanessa Hirsch, and Bette Korber. Preston Marx, thanks for the noodle dinner and a very enthusiastic day in the Big Apple and lots of helpful conversations. Chris Zink, Janice Clements, and Debbie Hauer for solving a difficult cloning problem, and also for being there.

Michiel Tjebbes helped greatly when things became pear-shaped with the Vreemdeling Politie, you were always patient in explaining how to sail the waters of the Dutch system.

Tina, and Angela our landlords who gave us shelter.

My paranimfen Marcel Beld, and Neil Renwick who were the best friends and comarades a guy could ever have.

Jeroen Vermeulen you're a great friend and thanks for introducing me to your sister against Dinko's advice. Jessica Vermeulen you were also a great landlord, sorry I was a lousy tenant. Time with you in Experimental Cardiology was illustrative and helpful. Aart and Beatrice Vermeulen you are heroes in the greates sense of the word. You taught me so many things and became my "nep-ouders". Your contribution to this thesis is without a doubt that it happened at all. Dad who supported me from the beginning to the end even in the darkest of hours. And most of all my loving wife Dorota who found out that marrying a PhD student certainly has its drawbacks but was unfailing in making our apartment a home. Here's to the future.