ADDIS ABEBA UNIVERSITY SCHOOL OF GRADUATE STUDIES

Molecular Epidemiology of HIV-1 Infection in a Cohort of Ethiopian Factory Workers

Submitted to the Biology Department in Partial Fulfilment for the Master of Science in Biology

By ASEFAW BERHE

JUNE,1999

TABLE OF CONTENTS

	PAGE
1. ABBREVIATIONS	3
2. REAGENTS	4
3. ABSTRACT	5
4. INTRODUCTION	6
5. METHODS AND MATERIALS	11
6. RESULTS	13
7. DISCUSSION AND CONCLUSION	32
8. REFERENCES	36
9. TABLES and ANNEXES	44

1. ABBREVIATIONS

AIDS ==	Acquired immunodeficiency syndrome
AMV-RT ==	Avian myeloblastosis virus reverse transcriptase
APS ==	Ammonium persulphate
cDNA ==	complementary deoxyribonucleic acid
dATP ==	2′- Deoxyadenosine ′5 tri phosphate
dCTP ==	2′-Deoxytyrosine 5′- tri phosphate
dGTP ==	2′-Deoxyguanosine 5′ tri phosphate
dH_2O ==	distilled water
DTP ==	2′-Deoxythymidine 5′ tri phosphate
dNTP ==	2′ Deoxinucleosine 5′ tri phosphate
ddNTP ==	2′ 3′ Dideoxynocleosine 5′ tri phosphate
EDTA ==	ethylene diamine tetra acetic acid
env ==	Envelope
gag ==	Group specific antigen
GUSCN ==	Guanidium Isothiocynate
HIV ==	Human Immunodeficiency Virus
NSI ==	Non-Syncytium-Inducing
PBMC ==	Peripheral Blood Mononuclear Cells
PCR ==	Polymerase Chain Reaction
SI ==	Syncytium-Inducing
Taq ==	*Thermophylus aquaticus* DNA polymerase
tat ==	Transactivator of transcription
TEMED ==	N-N-N-tetramethylethylenediamine
	$\{ (CH_3)_2 NCH_2CH_2 N(CH_3)_2 \}$

2. REAGENTS

. oligonucleotide primers:

ED31 ==	5′ CCTCAGTCATTACACAGGCCTGTCCAAAG (at 6816-6844)
ED33 ==	5′ TTAGGCATCTCCTATGGCAGGAAGCGC (at 7359-7380)
5′ V3 ==	5′-ATAAGCTTCAATGTACACATGGAATT-3′ (at 6506-6523)
3′ V3 ==	5′-ATGAATTCATTACAGTAGAAAAATTCC-3′ (at 6909-6928)

. 50x TAE (Tris-acetate) == 242 g Tris was dissolved in 800 ml of dH_2O, 57.1 ml glacial acetic acid and 100 ml EDTA .5M (pH8.0) was added and dH_2O was added to 1000 ml.

. 10x TBE (Tris-borate) == 108 g Tris base was dissolved in 800 ml of dH_2O, 55 g boric acid and 40 ml 0.5M EDTA (pH 8.0) was added and stirred well by a magnetic stirrer for one hour. dH_2O was added to 1000 ml.

. 6x Loading buffer == 1.5 ml Ficoll (Type 400; Pharmacia)
1 ml 0.5M EDTA (pH 8.0)0.025 g bromophenol
0.025 g Xylene cyanol FF and dH_2O to 10 ml.

. 20,000 x EtBr 1 g of EtBr was dissolved well by shaking gently for two hours in 10ml of dH_2O.

. DNA sequencing kit RPN 2436 Thermos Sequenase fluorescent labelled primer cycle sequencing Kit
Pro. No. 27.1492-01

. all other reagents used are described in the ENARP laboratory protocols, which can be found in the appendices.

ACKNOWLEDGMENT

I wolud_like to_ thank my Advisor Dr. Tobias Rinke de Wit for his all rounded assistance from the time of inception to accomplishment, in working extra hours to edit the manuscript. Also my internal Advisor Dr. Mekuria Lakew for his edifying counselling and editing the manuscript and fascilitating the program. I thank also Dr. Beyene Petros for helping me to identify the research topic and editing the Project Proposal.

I like to thank my Lab. Instructor Ms. Aletta Kliphuis Lab Technician, for her unreserved assistance in introducing me to the ALF automated Sequencer and other molecular techniques. Mr. Rutger Nugteren, the Data manager of ENARP also for assisting me in using the excel program to produce graphs. Ato Tesfaye Tilahun, a senior lab technician, and W/rit Almaz Abebe who _helped me in isolation and sequencing also deserve my sencere thanks. Last but not least, I would like to thank W/ro Aster Tseggai, all the lab technicians and staff of ENARP for creating a good working atmosphere in the laboratory.

This study is part of the Ethio-Netherlands AIDS Research program (ENARP), a collaborative effort of the Ethiopian Health and Nutrition research Institute (EHNRI), the Amsterdam Municipal Health Service (GG/GD), the Central Laboratory of the Netherlands Red Cross Blood Transfusion Service (CLB) and the Academic Medical Center of the University of Amsterdam (AMC). ENARP is financially supported by the Netherlands Ministry of Foreign Affairs and the Ethiopian Ministry of Health (MOH) as bilateral project.

3. ABSTRACT

In this study, to investigate the molecular epidemiology of the HIV-1 virus in Ethiopia, the HIV-1 gp120 V3 loop and flanking nucleotides were amplified from viruses isolated from the plasma samples of 23 HIV positive participants in the ongoing EHNRI-ENARP cohort project in Akaki and Wonji factories. A nested PCR DNA amplification protocol yielded correct size products, as determined by agarose gel electrophoresis in 61 out of 75 (81%)plasma samples collected between 1996-1999. Out of these 61 samples, 53 (87%) were sequenced successfully and analysed.

Phylogenetic analysis determined that all V3 sequences are of the HIV-1 subtype C, the most predominant in Ethiopia. Also, the V3 sequences derived showed lower intra-person mutation frequencies than inter-person mutation frequencies. In addition, it could be demonstrated that the Wonji and Akaki epidemics have a common origin and do not originate from separate introductions. Finally, the proportions of non-synonymous (dn) versus synonymous (ds) mutations are variable per subject and are not dependent on CD4 counts or viral load.

4. INTRODUCTION

A. HIV-1 GENERAL

Twenty years after the beginning of the AIDS epidemic and more than 14 years after the discovery of its etiological agent, the human immunodeficiency virus (HIV), the world is facing a global pandemic of vast proportions. Despite the intense national and international efforts to prevent new infections, about 35 million people are estimated to be living with HIV/AIDS, and about 11.7 million AIDS related deaths have been recorded up to the end of 1998 (UNAIDS and WHO Report, 1998). In Ethiopia, according to the Ministry of Health, approximately 2.5 million people are infected with HIV-1, amongst whom 100.000 are children (MOH, update on HIV situation in Ethiopia).

ORIGIN OF HIV

Human immunodeficient virus (HIV), is a retrovirus that belongs to the subfamily of lentiviruses. It is a retrovirus because it can make DNA from RNA by using reverse transcriptase (RT) enzyme. HIV includes at least 2 different families of viruses: HIV-1 and HIV-2. Both of these viruses cause slow chronic diseases that ultimately lead to degeneration of the immune system of the host. Although the rate of the disease progression varies substantially, the median time from initial infection to the development of AIDS is approximately ten years, in the industrialised world. The accepted predictors of HIV disease progression are decline of CD4, increase in viral load and change from NSI to SI phenotype of the virus.

As regards the origin of HIV-1 it has been speculated that transmission took place between Chimpanzees and humans (3). This is based on DNA analyses of the genomes of lentiviruses isolated from two chimpanzees (*Pantroglodites troglodites*), originating from Gabon ($SIV_{-cpz\ gab}$) and ($SIV_{-cpz\ ant}$) from Zaire, which were demonstrated to be structurally related to the genome of HIV-1. On the other hand, several area's of the Chimpanzee genomes had less than 50% homology with the corresponding regions in the HIV-1 genome. In contrast, it has been clearly demonstrated that HIV-2 and simian lentiviruses isolated from wild sooty mangabeys (SIV_{sm}) have a high degree of genetic and phenotypic homology (3), making the assumption of a zoonotic transmission of this virus a very likely one.

HIV-1 SUBTYPES

HIV-1 is an extremely variable virus, due to lack of proof-reading capacity of its reverse transcriptase. The error rate for HIV-1 is about 10^{-4} per base, or about one mis-incorporation per genome per cycle. As a result, HIV continually produces 'escape mutants', which eventually evade the surveillance of the immune system and chemotherapeutic agents. Thus, this variability of HIV-1 poses a major challenge to the immune system of the host, as well as to design of appropriate and effective vaccines.
As a consequence of this variability, HIV-1 hasevolved in various area's of the world as different groups, with different subtypes. First of all, HIV-1 is divided into three groups: the major (M), the outliers (O) and the new (N), which differ >35% from each other in terms of their complete envelop encoding genes (1). HIV-1 subtypes are defined as viruses which are 25-35% genetically different in their gp120 genes. Thus, the HIV-1 group M is divided into at least 10 different (A-J) subtypes, whereas the group O consists of a pool of highly divergent, genetically related strains with at least 4 defined subtypes. The N subgroup is genetically placed between the two groups M and O. (2).
In Africa most of the ten subtypes of group M, as well as subtypes O and N are currently circulating. The C subtype is the one predominantly found in East and South Africa, India, China and South East Asia. According to WHO, of all the HIV-1 infected people in the world, 48% carry the C subtype. In Europe and America, where most of the HIV research is carried out, the most prevalent subtype is the B subtype.

HIV-1 STRUCTURE

HIV is an enveloped virus i.e. the virion is surrounded by a lipid membrane derived from the host cell (4). Two viral proteins encode the envelope: the outer surface glycoprotein gp120 and the transmembrane glycoprotein, gp41 (See fig. 1. A and C). The HIV-1 virion has a diameter of ~100-120 nm. Inside the viron, the matrix protein, p18 (fig 1. B), is intimately associated with the inner surface of the viral membrane and the capsid protein, p24, forms the cone-shaped viral core (5). The core capsid protects two identical molecules of genomic viral RNA that are associated with the nucleocapsid protein, p15. The combination of viral nucleic acid and the nucleocapsid make up the ribonucleoprotein complex. A single HIV-1 viron is estimated to contain 1200 Gag molecules, 80 Gag-Pol and upto 280 molecules of Env (6). There are also tRNA molecules found in HIV-1 virons

that are used for the initiation of reverse transcription; HIV-1 uses a specific tRNA species (tRNA ^{Lys3}) for initiation of reverse transcription (6).

HIV-1 GENOME

The genome of HIV is ~9.5 kilo bases in length (7). In common with other lentiviruses, it contains two long terminal repeats (LTR) and three major genes (gag, pol, and env). These genes encode the major structural proteins of the virus, via polypeptide precursors Pr 55^{gag}, Pr 160 $^{gag\text{-}pol}$, and gp160 ^{env} (8). The genome also encodes three regulatory proteins; Rev, Tat, and Nef. Finally, three other accessory proteins that are probably involved in virus maturation and release are encoded: Vif, Vpu, and Vpr.

HIV-1 LIFE CYCLE

The life cycle of HIV-1 can be arbitrarily divided into early and late phases (see fig. 2) (10). The early phase consists of: (1) the virion binding to the membrane using CD4 and a co-receptor, (2) fusion with the membrane and (3) viral entry into the host cell. The virus soon after it enters the host cytoplasm releases the RNA genome (4) and reverse transcription takes place (5). The resulting double stranded DNA molecule is transported to the nucleus as provirus DNA (6), where it is integrated into the host genome. Here, the virus can remain silent for a long period. The late phase starts upon immune activation of the host cell, mediated by Tat (7). Transcription of the integrated viral genome produces full-length mRNA transcripts that are transported intact to cytoplasm or spliced and then (8) transported. Translation of doubly spliced mRNA in the cytoplasm produces control proteins Rev (regulator of virion proteins) and Tat (transactivator of transcription). Both Rev and Tat migrate back to the nucleus. Tat binds to the TAR (Transactivation response) element, a sequence found on nascent transcripts, causing transactivation of viral transcription (11). Rev accumulates in the nucleus until it reaches a threshold concentration. It then binds to the RRE (Rev-response element), an RNA element on full-length and singly spliced transcripts, inhibiting their splicing and facilitating their export from the nucleus to the cytoplasm (12). In the cytoplasm, the full-length mRNA is (9) translated to form the Gag (group specific antigen) and Gag-pol proteins, and the singly spliced mRNA is translated to form Env (envelope) protein and gp41 (10). The viral

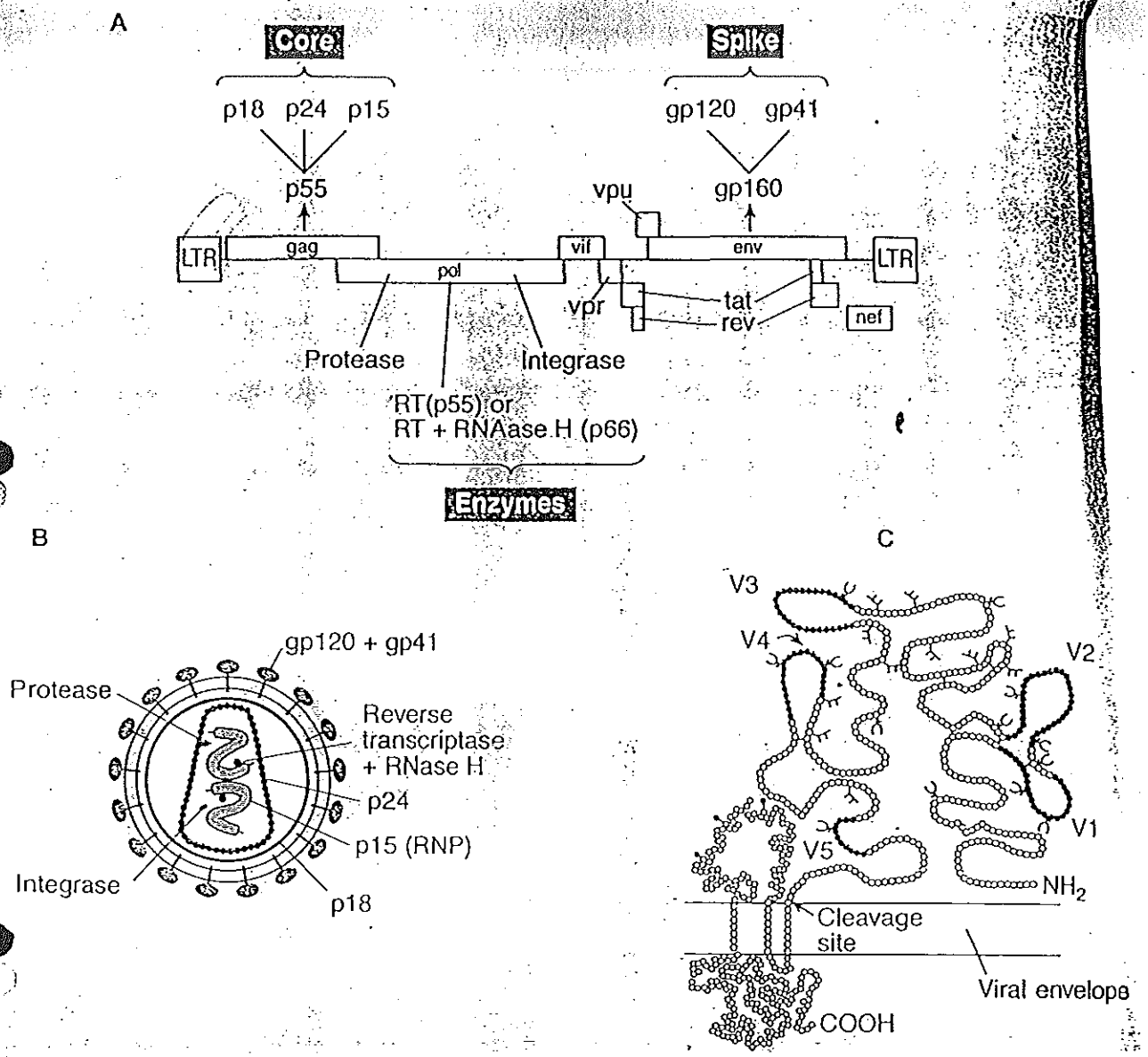


Fig. 1 (A) Production of the structural proteins of HIV-1" p15" indicates a polypeptide that has M_r of 15,000 etc.

(B) Structure of the viron.

C) The tetrameric envelope spike gp 120 and gp 41 transmembrane glycoprotein. The variable regions (V1-V5) and the conserved regions inbetween are shown.

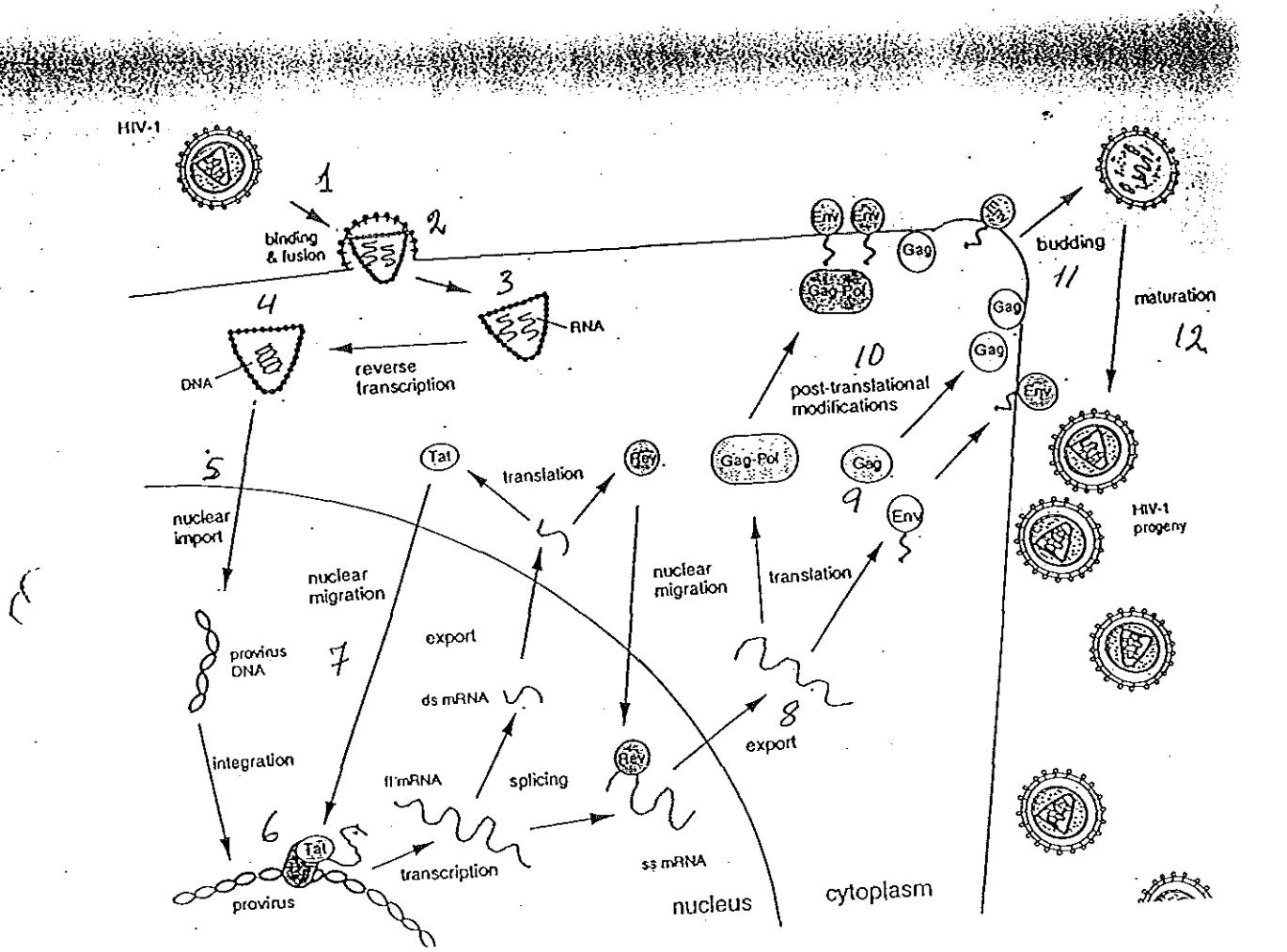


Fig. 2 Overview of the HIV-1 growth cycle.

proteins undergo post-translational modifications and are transported to the plasma membrane (13). When sufficient quantities of viral proteins have accumulated near the plasma membrane, budding (11) takes place and immature virions are released into the extracelluar environment. Virion maturation which is initiated by self-cleavage of the protease, occurs during and after budding (14). Thus, HIV-1 genetic material exists as RNA in the plasma and proviral DNA in the PBMC.

HIV-1 AND HOST CELL MEMBRANE INTERACTION

HIV-1 infection is initiated by interaction of the virion envelope glycoproteins with at least two cellular receptors: the CD4 molecule and a seven transmembrane domain G-protein coupled chemokine (CC) receptor.

Macrophage-tropic strains of HIV-1 replicate in macrophages and CD4 T cells and use the CC chemokine receptor CCR5 for entry; these viruses are classified as R5. In most cases R5 viruses have a low titer, are slowly replicating, and do not induce *in vitro* syncytia when co-cultured with the MT-2 cell line. The CCR5 co-receptor is used by almost all primary HIV-1 isolates regardless of viral genetic subtype, and by the related lentiviruses HIV-2 and SIV. T-tropic isolates of HIV-1 replicate in CD4 T cells, established CD4 T cell lines and macrophages. All of these viruses use the chemokine receptor CXCR4, but on top of this co-receptor many of them also use other co-receptors, like CCR2b CCR3, Bonzo, Bob, CCR5 (16). Those that use CXCR4 are called X4, are the high titer, fast replicating and usually of the SI phenotype. Because of higher viral loads (fast transcription) and broader co-receptor use (tropism), SI viruses are much more dangerous than NSI. Usually, the emergence of SI variants in a host predict increased destruction of CD4 T-cells resulting in accelerated progression towards AIDS.

THE GP 120 V3 REGION

The HIV-1 envelope gene is characterised by a high potential for variation without loss of functionality. The gene encodes the viral envelope, including five variable domains (V1-V5), separated by relatively conserved regions (see fig.1C). The third variable domain (V3) has been shown to contain determinants of cell tropism (18), infectivity (19), and cytopathycity (20). It is also claimed to be a target for neutralising antibodies (21), and a recognition site for cytotoxic T-cells.

Specific mutations on either side of the V3 loop can influence cellular tropism and SI capabilities of viruses in culture. Other regions in Env are also partly responsible for features of the viral V3-phenotype (24). In particular, mutations in other regions of Env can affect ligand binding to the V3 loop and, similarly, mutations in V3 loop can affect ligand binding to other regions (25). For HIV-1 A, B, D, and E subtypes, T-cell-tropic strains have been demonstrated to carry positively charged amino acids adjacent to the tip of the loop (23) in positions 11 and 27 of the V3 loop. The situation for other subtypes, like subtype C is not so clear. It has been speculated that positions 8 and 29 play a role for subtype C (47), but this needs further substantiation.

A growing body of evidence argues that NSI, M-tropic forms of the virus are the predominant forms detected immediately post-infection and that SI strains emerge later during the course of the disease (26). However, it should be kept in mind that at least 40% of subtype B infected individuals die of AIDS, without the emergence of any SI variant (cohort studies Amsterdam, 1984-1999).

HIV-1 IN ETHIOPIA

In Ethiopia, since the first report of HIV-1 sera from 1984 and the first AIDS cases in 1986 (31), the epidemic has reached a prevalence of 8-16% amongst sexually active adults in urban areas by the end of 1997. Various studies have indicated the predominant presence of HIV-1 subtype C in infected Ethiopians. (32-37). A national sub-typing effort is being carried out to shed light on this mater. The results indicate the predominant presence of subtype C (38).

EHNRI-ENARP

The Ethiopian Health and Nutrition Research Institute-Ethiopian Netherlands AIDS Research Project (EHNRI-ENARP), has established a cohort study on HIV-1 infection progression in factories in Akaki and Wonji in February 1997. Participants were informed about the project, in groups sessions as well as individually. After signing an informed consent, subjects could enrol into the prospective cohort study. Participants are entitled to withdraw at any moment. As an incentive, free medical care is provided to the participants and their direct dependants (5-7 individuals). Between February 1997 and April 1999, 765 factory workers in Akaki and 511 in Wonji, a total of 1276 individuals enrolled. Out of these 87 (11.4%) in Akaki and 39 (7.6%) in Wonji were HIV-1 positive at intake. Blood

samples have been collected of those subjects who returned (60-65%), every 6 months. In this process, 12 HIV-1 seroconverters, 8 in Akaki and 4 in Wonji, were detected. Detailed information is being collected on those people, including immunological, virological epidemiological and clinical data (40). Table 1 summarises some clinical information on the HIV-1 positive cohort participants.

AIM OF PRESENT STUDY

The aim of the present study was to assess the molecular characteristics of HIV-1 virus isolates from HIV-positive Ethiopian factory workers, to get an insight in the dynamics of mutation and finally to try to relate these molecular characteristics to well established prognostic markers for HIV infection progression, like CD4 T-cells and viral load.

5. METHODS AND MATERIALS

STUDY POPULATION

Study subjects were Ethiopian factory workers at all stages of HIV-1 infection, who participated in the EHNRI-ENARP cohort studies on a regular basis. Out of the 125 sero-positive cohort participants, 23 individuals were selected (17 from Akaki and 6 from Wonji). These subjects include male and female participants with a mean age of 37 (range 18-52).

VIRUS ISOLATION AND RNA AND DNA EXTRACION

The collection of plasma was done from heparinized veinous blood and stored at -80 ° C until use.

PLASMA SAMPLES

RNA was extracted form stored plasma samples by using the method described by BOOM et al (40). For that purpose 100 µl of plasma was taken and mixed with 900 µl of L6, lysis buffer. This was well mixed and 40 µl of silica added to each sample and well vortexed. After incubating for 10 minutes at room temperature, shaking after every 2 minutes, was centrifuged for 35 seconds. at 12000 girs. The supernatant was sacked and the silica was resuspended by adding 1 ml of L2, washing buffer, and vortexing for 5 seconds. After centrifuging the process was repeated once again. Then the same process was repeated twice with 70% ethanol, and once with acetone.

The silica was allowed to dry for 10 minutes at 56 °C, after which 50 µl of low TE-4 buffer was added, mixed well and incubated for 15 minutes shaking gently at 56°C. The mixture was taken out and centrifuged for 2 minutes at 12,000 girs. 10 µl of the supernatant RNA, was collected and added to 10µl of RT mix.

REVERSE TRANSCRIPTION

RNA IS EXTREMELY UNSTABLE (STORAGE –80 °C FOR 3 DAYS ONLY)

THEREFOR WE NEED TO MAKE **DNA** (MORE **STABLE, EASY** TO **AMPLIFY**)

IN RT STEP WE MAKE **SINGLE STRANDED cDNA**

MOST IMPORTANT ITEMS IN RT MIX ARE:

1. 3'END PRIMER (ED33), it has the COMPLEMANTARY SEQUENCE of the virus at the 3' end (right side in nearly all drawings)

2.AMV-RT == Avian myeloblastosis virus reverse transcriptase . This enzyme converts the RNA in to complementary DNA.
3. RNAsin == a reagent that inhibits the activity of the RNAse enzyme.

*** REVERSE TRANSCRIPTION MIX**

RT-MIX	1x (µl)	11x (µl)	Stock conc.	Final conc. (20 µl)
$MgCl_2$	4.0	44.0	25mM	5mM
RT buffer	2.0	22.0	10 x	1 x
DNTPs	2.0	22.0	10 mM	1 mM
RNAsin	0.5	5.5	40 U/µl	1 U/µl
AMV-RT	0.5	5.5	10 U/µl	0.25 U/µl
3' primer ED33	1.0	11.0	10 pmol./µl	.5 pmol/µl
Total	**10**	**110.0**		

HIV 1 RNA AFTER ISOLATION

+ 5'- |_____| -3' HIV 1 RNA

CCUCAAAUUCGGGUACGAAUGGCUAACCGUUGAATCGCG

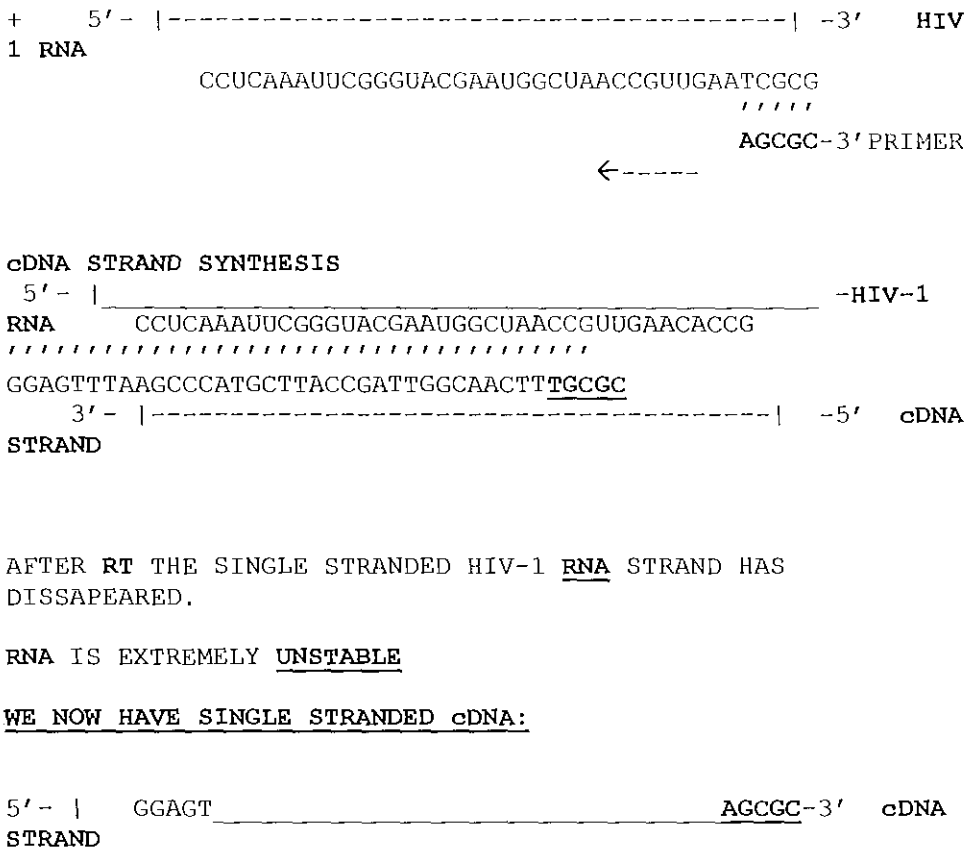
RT STEP REVERSE TRANSCRIPTION MAKING *cDNA* (COMPLEMENTARY) OF *RNA*

5'CCUCA_____TCGCG'3 HIV 1 RNA

+

CCTCA 3' PRIMER (ED33) (COMPLEMENTARY TO *3'END RNA*)
←-----

ANNEALING 3'PRIMER



AFTER **RT** THE SINGLE STRANDED HIV-1 **RNA** STRAND HAS DISSAPEARED.

RNA IS EXTREMELY **UNSTABLE**

WE NOW HAVE SINGLE STRANDED cDNA:

5'- | GGAGT AGCGC-3' **cDNA STRAND**

N. B. The mixture was prepared for ten samples and then aliquoted in 10 µl PCR tubes.

The mixture of 10 µl RNA and 10 µl RT-Mix were incubated for 45 minutes at 42°C. During this time the RNA is converted to cDNA by the AMV-RT enzyme. After the process the mixture was incubated at 95°C for 5 minutes inorder to inactivate the enzyme and stop the process of reverse transcription. The mixture was added to 80 µl of PCR I mix and was covered with two drops of mineral oil and put into a Thermal Cycler for the polymerase chain reaction.

AFTER RT STEP (45 MINUTES AT 42 C°) WE HAVE: **SINGLE STRANDED cDNA**

PCR I. (25) POLYMERASE CHAIN REACTION

In RT step there is no amplification only cDNA is made.

In order to amplify we need to add another **PRIMER:** 5'END PRIMER (ED31)

This PRIMER has the complementary sequence of the **cDNA** on the

5'END (left side in nearly all drawings)

In fact this is the region of interest in sequence of the virus

BESIDES THE 5'END PRIMER, WE NEED:

10*PCR BUFFER

dNTP'S (THE BUILDING STONES A,C,G,T OF DNA)

TAQ POLYMERASE (ENZYME THAT ADDS THE COMPLEMANTRY SEQUENCE)

$MgCl_2$ (IMPORTANT FOR **PRIMER ANNEALING**)

PCR I IS ALWAYS **35** CYCLES: 1 CYCLE:
1 MIN. 95 C° (**DNA DENATURATION**)

1 MIN. 55 C°

(PRIMER ANNEALING)

2 MIN. 72 C°

(EXTENTION)

AFTER 35 CYCLES:

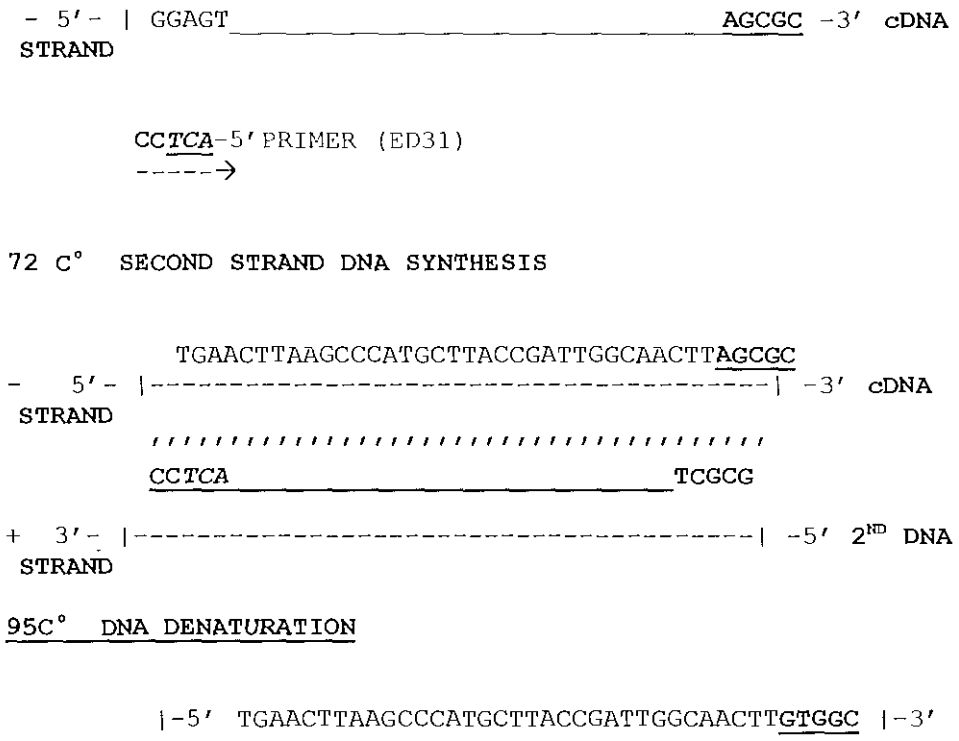
WE CHECK LENGTH OF 1 PCR PRODUCTS ON AGAROSE GEL:

- LOAD A Commercially prepared of known length DNA **MARKER** NEXT TO PCR I PRODUCTS

- RUN FOR **1** HOUR AT 100 mAs.

- CHECK WITH UV LIGHT THE MARKER AND LENGTH OF OUR PRODUCT

- TAKE POLAROID PICTURE OF THE GEL WITH THE HELP OF THE UV LIGHT. USUALLY THE PRODUCT IS NOT VISIBLE, BUT IF IT IS VISIBLE IT IS 564 bps LONG



55 C° PRIMER ANNEALING

72 C° DNA STRAND EXTENTION

95 C° DNA DENATURATION

THEORETICLY AFTER APPOX. 25 CYCLES WE HAVE **1.000.000** COPIES FROM 1 ORIGINAL HIV-1 RNA PARTICLE.

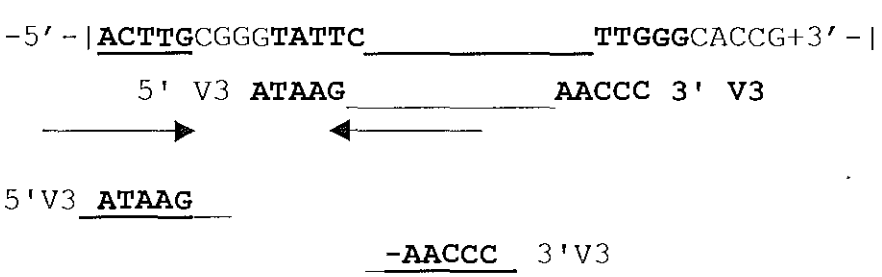
PCR I Conditions : (file (94/ 95/ 98/ 99)

94°C	1 min		
55°C	1 min	5x	(file: 94)
72°C	1 min		

94°C	15 sec		
55°C	45 sec	30x	(file: 95)
72°C	1 min		
72°C	10 min	1x	(file : 98)
4°C	soak file		(file: 99)

From the products of PCR I 10 µl was taken and added into ten PCR tubes of 90 µl each of PCR II* mix, covered with two drops of mineral oil and was put into the thermal cycler for a nested PCR.

2nd PCR MIX	1x (µl)	11x (µl)	Stock conc.	Final conc. (100 µl)
$MgCl_2$	16.0	176.0	25 Mm	4 mm
PCR Buffer	10.0	110.0	10x	1 x
dNTPs	2.0	22.0	10 mM	0.2 mM each
Taq. Pol.	0.5	5.5	5 U/ µl	0.025 U/ µl
5′ primer V3-M13F	1.0	11.0	10 pmol./µl	0.1 pmol/ µl
3′ primer V3- M13R	1.0	11.0	10 pmol/ µl	0.1 pmol/µl
dH_2O	59.5	654.5		
TOTAL	**90.0**	**990.0**		



NESTED PCR CONDITIONS: (file: 96/97/98/99)

94°C	1 min		

94°C	1 min		
55°C	1 min	5x	(file; 96)
72°C	1 min		
94°C	15 sec		
55°C	45 sec	25x	(file: 97)
72°C	1 min		
72°C	10 min	1x	(file: 98)
4°C	soak file		(file; 99)

The first and nested PCR products were run in 1% Agarose Gel Electrophoresis at 100mA for 1 hour and Polaroid Pictures were taken.

AGAROSE GEL ELECTROPHORESIS

To make 1% agarosegel we took 1.5 g of agarose powder and dissolved it in 150 ml of 2% TAE. This was heated at microwave heater for 2 ½ minutes at 360 V. After it was allowed minutes at 360 V. After it was allowed to cool for about 10 minutes at room temperature, 7.5 µl of EtBr was added, and it was cast at the electrophoresis plate and was left for 30 minutes to polymerise. Meanwhile in a microplate wells 10 □l of the PCR product, 5 µl of dH2O, and 5µl of loading dye buffer was mixed and added into the wells in the gel. The DNA which is negative runs towards the anode after the system was connected to an electric source of 100 mAs and 95 Vs. After an hour the gel was taken to the dark room and a Polaroid picture of the gel taken. (see fig. 3)

AUTOMATIC SEQUENCING GEL RUNNING

Sequence running was done according to the Manufacturers' Manual.

PROCESSING OF AUTOMATIC SEQUENCING DATA

Data processing was done by using Microsoft Word and the MEGA soft ware program. Alignments of sequences were performed manually.

PHYLOGENETIC ANALYSES

Phylogenetic analysis was carried out on DNA sequences aligned to sets of previously classified non-recombinant references of subtypes A, B, C, D, E, F, G and H from Brazil, Britain, India, Malawi, Thailand, Uganda and Zambia (41), by using the Mega and PHYLIP soft ware packages (42). Gaps were introduced manually for optimal alignment. Gaps were excluded from any phylogenetic algorithms. The algorithms of the PHYLIP software package including DNADIST, NEIGHBOR, CONSENSE, and DRAW TREE were employed. The maximum Likelihood algorithm was used for bootstrapping.

6. RESULTS

Table 2 summarises some of the most important virological, immunological and other data collected from the 23 selected HIV-1 positive cohort participants (17 from Akaki and 6 from Wonji). The HIV-1 positive participants were 16 males and 7 females with a mean age in 1999 of 38 years for males (range 29-47 years) and 33 years for females (range 24-45 years). Over the past 3 years, some individuals presented already 6 times at the cohort site (e.g. E), others only reported once (e.g. C). The dates of visit are indicated and show that time intervals between visits ranged between 6 months and 2 years. Mean follow up time of the subjects from Akaki was: 1,040 days = 2 years and 10 months; mean follow up time of subjects from Wonji was: 830 days = 2 years and 3 months.

Absolute CD4 T cell counts varied between 36 (subject J) and 943 (subject G) $x10^6$/l whole blood. Absolute CD8 T cell counts varied between 267 (subject O) and 2677 (subject N) $x10^6$/l whole blood. CD4/CD8 ratios varied between (0.1) (e.g. subjects E and F) and 0.9 (subject A). Six subjects were found to consistently have < 200 CD4 T-cells per µl whole blood (B, E, F, J, M, O). These people are in the AIDS stage according to the WHO staging system of HIV-1 infection progression (1). All other subjects were in stages 1-3, as assessed by both CD4 T-cell counts and clinical examinations (not shown).

Viral loads varied between undetectable (<1.90 ^{10}log RNA molecules per ml of plasma, e.g. for subject X) and very high: 5.64 ^{10}log RNA load (subjects C and N). There was no clear correlation between CD4 T-cell counts and viral loads of individual cohort participants (not shown).

Figures 8A and 8B (see annexes) demonstrate the dynamics of CD4 T-cell counts and viral loads over time in those cohort participants from Akaki who had >3 follow up visits. In general, both $CD4^+$ T-cell counts and viral loads tended to be stable over the follow up time of on the average 21/2 years.

From the above 23 subjects, HIV-1 RNA was isolated from a total of 75 plasma samples (see Table 2) and converted into cDNA using the reverse transcription system described in Materials and Methods. The cDNA was subjected to a nested PCR which finally resulted in 61 positive products (61/75 = 81%, Table 2).

Figure 3 shows some representative agarose gel electrophoresis experiments performed on nested PCR products using primers ED31/33 followed by M13 5'F/3'R. The expected length of first PCR products (produced by the ED31/33 primer combination) was 564 bp. However, agarose gels hardly ever showed these products, because of usually a too low sensitivity of this technique. In contrast, when the nested PCR was performed (with primer combination M13 5'F and M13 3'R), the resulting 350 bp product was most of the times clearly visible. Figure 3 displays some examples of PCR products with the correct sizes of 564 and 350 bp, as compared to the molecular weight marker used.

The nested PCR products were subjected to cycle sequencing, either using forward or reverse sequencing primers and the resulting DNA sequences are depicted in Figure 4, aligned with a consensus subtype C sequence from Los Alamos, 1998. Out of the 61 PCR products, 53 (87%) resulted in an interpretable DNA sequence. As can be seen, most of the DNA sequences aligned well with the Los Alamos consensus HIV-1 subtype C sequence. The samples were initially sequenced by using the M13 5'F direct primer. This primer gave good results for many of the samples, up to the first 250 bp. However, towards the end of the sequence (positions 250-279), quite often ambiguities were observed. Those sequences with more than 10 ambiguities after editing, were re-sequenced using the M13 3'R reverse primer. Thus, in most cases ambiguities were reduced or even eliminated. The uncertain positions remaining are indicated by question marks in Figures 4 and 5.

Figure 4. Aligned V3 DNA sequences of Akaki and Wonji HIV+ cohort participants. Legend: Sequences are aligned relative to a consensus subtype C sequence (CONSENSC), which was obtained from the Los Alamos 1998 databases. A dot (.) represents the same nucleotide as the consensus, a dash (-) represents a deletion relative to the consensus, a question mark (?) represents a sequencing ambiguity. Suffix A-P + X.= cohort participants from Akaki; suffix Q-V = cohort participant froms Wonji . The numbers in front of these codes represent Lab ID's, indicating the order of visits of the pertinent subjeet ID's. In **bold** are the V3 loop encoding sequences. Sequences marked by a * have a high number of ambiguous positions.

```
CONSENSC  ATA ATA ATT AGA TCT GAA AAT CTG ACA AAC AAT GTC AAA ACA ATA ATA
00099A     ... G.C ... ... ... ... ... T.. ... ... ... .C. ... .T. ... ...
00101B     ... G.. G.. ... G.. ... ... ... ... ... ... .C. ... .T. T.. ...
01500B    ... G.. ... ... G.. ... ... ... GA. ... ... .C. ... .T. C.. ...
00143C    ... G.. ... ... ... ... ... A.A ... ... ... ... ... ... ... ...
00153D    ... G.. ... ... ... ... ... ... ... ... ... A.. ... ... ... ...
00707D    ?.. G.. ... ... ... ... ... ... ... ... ... A.. ... ... ... ...
01706D    ... G.. ... ... ... ... ... ... G.. .G. G.. ... ... ... ... ...
0724E     ... G.. ... ... ... ... ... ... ... G.. ... ... ... .AT ... ...
00156E    ... G.. ... ... ... ... ... ... ... G.. ... ... ... .AT ... ...
00713E    ... G.. ... ... ... ... ... ... ... G.. ... ... ... .AT ... ...
01625E    ?.. G.. ... ... ... ... ... ... ... G.. ... ... ... .AT ... ...
01105F    ... G.. ... ... ... C.. ... ... ... ... ... .C. ... .T. ... ...
00258G    ... G.. ... ... ... ... ... ..A ... ... ... ... ... .T. ... ...
00849G    ... G.. ... ... ... ... ... ..A ... ... ... ... ... .T. ... ...
01687G    ... G.. ... ... ... ... ... ..A G.. ... ... ... ... .T. ... ...
0688H*    ??? ??? ??? ??. ... A.. ... ?.. ... ... ... .?? ... .T. ... ...
00965H   G.T G.. ... ... ... ... ... T.. ... ... ... .C. ... .T. ... ...
02357H    ... G.. ... ... ... ... ... ... ... ... ... .C. ... .T. ... ...
00647I    ... ... ... ... ... ... ... ... ... ... ... .CG ... .T. ... ...
00330J    ... ... ... ... ... ... ... ... ... ..T ... .C. ... GT. ... ...
00979J    ... ... ... ... ... ... ... ... ... ..T ... .C. ... GT. ... ...
01748J    ?.. ... ... ... ... ... ... ... .?. ..T ... .C. ... GT. ... ...
00363K    ??. ... ... .?. ... ... ... ... ... ... ..A .C. ... .T. ... ...
01169K    .C. ... ... ... ... ... ... ... ... ... ... .C. ... .T. ... ...
01763K    ?C. ... ... ... ... ... ... ... ... ... ... .C. ... .T. ... ...
0592L     ?.. G.. ... .A. ... ... ... ... ... ... ... .C. ... .T. ... ...
00403L    ... G.. ... .A. ... ... ... ... ... ... ... .C. ... .T. ... ...
01195L    ... G.. ... .A. ... ... ... ... ... ... ... .C. ... .T. ... ...
01909L    ?.. G.. ... .A. ... A.G ... ... T.. G.. ... .C. ... .T. ... ...
02776L    ... G.. ... .A. ... ... ... ... ... ... ... .C. ... .T. ... ...
0808M     ... G.. ... ... ... ... ... ... ... ... ... ... ... ... ... ...
00409M    ... G.. ... ... ... ... ... ... ... ?.. ... ... ... ... ... ...
01752M    ... G.. ... ... ... ... ... .?. ... ... ... ... ... ... ... ...
00448N    ... ... ... ... ... ... ... ... ..G ... ... .C. ... .T. ... ...
01266N    ... ... ... ... ... ... ... ... ..G ... ... .C. ... .T. ... ...
02009N    ... ... ... ... ... ... ... ... ..G ... ... .C. ... .T. ... ...
02794N    ... ... ... ... ... ... ... ... ..G ... ... .C. ... .T. ... ...
00499X    ..T ... .?. ... ... ... ... ... ... ... ... .CT ... ..C ... ...
01296X*   ... ... .A. ... ... .?. ... ... ... ... ... .C. C.. ..C ..T ...
01893X    ... ... ... ... ... ... ... ... ... ... ... .C. ... ... ... ...
01094O    ... G.. ... ... ... ... G.. ... ... G.. ... .C. ... .T. ... ...
0764P     ... G.. ... ..G ... A.. ... ... ... ... ... .C. ... .T. ... ...
```

```
02050P  ... G.. ... ..G ... ... ... ... ... ... ... ... ... .T. ... ...
7202Q   ... G.. ... ... ... ... ... ... ... ... G.. .C. ... .T. ... ...
7194R   ... ... G.. ... ... ... ... A.A ... ... ... .C. ... ... ... ...
01732R  ... ... G.. ... ... ... ... A.A ... ... ... .C. ... ... ... ...
7050S   ... G.. ... ..G ... ... ... ... ... G.. ..A ... ... .T. ... ...
01100S* ... ... ... ... ... A.. ..C ... ... G.. ... AC. ... CT. ... ...
01746S  ... ... ... ... ... A.. ..C ... ... G.. ... AC. ... CT. ... ...
7233T*  .C. ... ... ..C ... ... .C. ..C C.C C.. ... .C. ... .T. ... ...
7043U   ?.. G.. ... ..G ... ... ... ... ... G.. ... ... ... .T. ... ...
01842U  ?.. G.. ... ..G ... ... ... ... ... G.. ... ... ... .T. ... ...
02191V  ... ... ... ... ... ... ... A.A ... ... ... .C. ... ... ... ...

CONSENSC  GTA CAT CTT AAT GAA TCT GTA GAA ATT GTG TGT ACA AGA CCC AAC AAT

00099A  ... ..G ... ... ... A.. ... ... ... ... ... ... ..G ... ... ...
00101B  ... ..G ... ..C ... ... ... ... ... ... ... ... ..G ... ... ...
01500B  ... ..G ... ..C ... ... ... ... ... ... ... ... ..G ... ... ...
00143C  ... ... ... ..A A.G C.. ... ... ... ... ... ... ..G ... ... ...
00153D  ..G ..G ... C.A ..C C.. ... ... ... ... ... ... ..G ..A .G. ...
00707D  ..G ..G ... C.A ..C C.. ... A.. ... ... ... ... ..G ..A .G. ...
01706D  ..G ..G ... C.A ..C C.. ... ... ... ... ... ... ..G ..A GG. ...
0724E   ... ... ... C.A A.C C.. ... ... ... .A. ... ... ..G ... ... ...
00156E  ... ..G ... ..A ... C.. ... ... ... ... ... ... ... ... ... ...
00713E  ... ..G ... ..A ... C.. ... C.. ... ... ... ... ..G ... ... ...
01625E  ... ..G ... ..A ... C.. ... C.. ... ... ... ... ... ... ... ...
01105F  ... ..G ... ..A A.C C.. ... ... ... ... ... G.. ..G ... ... ...
00258G  ... ... ... C.A A.? C.. ... ... ... ... ... ... ..G ... ... ...
00849G  ... ... ... C.A A.. C.. ... ... ... ... ... ... ..G ... ... ...
01687G  ... ..G ... C.A ... C.. ... ... ... ... ... ... ..G ... ... ...
0688H*  ... ..G ... ..A ... C.. ... ... ... ... ... ... ..G ... .C. ...
00965H  ... ..G ... ..A ... C.. ... ... ... ... ... ... ..G ... ... ...
02357H  ... ..G ... ..A ... C.. ... ... ... ... ... ... ..G ... ... ...
00647I  ... ..G ... ... ... A.. ... .G. ... AAT ... ..G ... ... ..T ...
00330J  ... ..G ... .GG C.. C.. ... ... ... AAT ... ... ... ... ... ...
00979J  ... ..G ... G.G C.. C.. ... ... ... AAT ... ... ... ... ... ...
01748J  ..G ..G ... ..G C.. C.. ... ... ... AAT ... ... ... ... ... ...
00363K  ..G ..G ... ..G ... C.. ... ... ... ... ... ... ..G ... .G. ...
01169K  ..G ..G ... ..G ... C.. ... ... ... ... ... ... ..G ... GG. ...
01763K  ..G ..G ... ..G ... C.. ... ... ... ... ... ... ..G ... GG. ...
0592L   ... ... ... ..A A.. C.. ... ... ... ... ... ... ..G ... .G. ...
00403L  ... ... ... ..A A.C C.. ... ... ... ... ... ... ..G ... .G. ...
01195L  ... ... ... ..A A.C C.. ... ... ... ... ... ... ..G ..T .G. ...
01909L  ... ... ... ..A A.. C.. ... ... ... ... ... ... ..G ..T ... ...
02776L  ... ... ... ..A A.C C.. ... ... ... ... ... ... ..G ..T .G. ...
0808M   ... ..G ... ..A ACC C.. ... ... ... ... ... ... ..G ... ... ...
00409M  ... ..G ... ..A ACC C.. ... ... ... ... ... ... ..G ... ... ...
01752M  ... ..G ... ..A ACC C.. ... ... ... ... ... ... ..G ... ... ...
00448N  ... ..G ... ..G ..C C.. ... A.. ... .AT ... .T. ... ... ... ...
01266N  ... ..G ... ..G ..C C.. ... C.. ... AAT ... ..G ... ... ... ...
02009N  ... ..G ... ..G ..C C.. ... C.. ... AAT ... ..G ... ... ... ...
02794N  ... ..G ... ..G A.C C.. ... C.. ... AAT ... ..G ... ... ... ...
00499X  ... ..G ... ... .C. C.. ... ... ... AAT ... ..G ... ... ... ..C
01296X* ... ..G ... ..A .C. ?.. ... ... ... AAT ... ..G ... ... ... ..C
01893X  ... ..G ... ..G .C. C.. ... ... ... AAT ... ..G ... ... ... ..C
01094O  ... ..G ... ..A ... C.. ... ... ... ... ... ... ..G ... .G. ...
0764P   ... ... ... ..A A.C C.. ... ... ... ... ... ... ..G ... ... ...
02050P  ... ... ... ..A ..C C.. ... ... ... ... ... ... ..G ... .C. ...
7202Q   ... .GG ... ... A.. ... ... ..C ... ... ... ... ..G ... GG. ...
```

```
7194R      ... ..A ... ... ... A.. ... ... ... AAT ... ... ... G.. GG. ...
01732R     ... ..A ... ... ... A.. ... ... ... AAT ... ... ... ... GG. ...
7050S      ... ... ... ..A ..C C.? ... A.. ... .A. ... ... ..G ... GG. ...
01100S*    ... ..G ... ..A ..C C.. ... C.. ... ... ... ... ..G ... ... ...
01746S     ... ..G ... ..A ..C C.. ... ... ... ... ... ... ..G ... ... ...
7233T*     ... ..G ... ..A ACG C.. ... ... ... ACC ... TTG CCC ... .G. ...
7043U      ... ... ... ..A ..C C.G ... ... ... .A. ... .T. ..G ... GG. ...
01842U     ... ... ... ..A ..C C.G ... A.. ... .A. ... ... ..G ... GG. ...
02191V     ... ..G ... ... A.. A.. ... ..C ... AAT ... ... ... ..T GG. ...
```

```
CONSENSC  AAT ACA AGA AAA AGT ATA AGG ATA GGA CCA GGA CAA --- --- ACA TTC

00099A    ... ..C C.G G.. ..C ..G ... ... ... ... ... ... ... ... ... ...
00101B    .?. C.C G.. ... ... G.G ... ... ... ... ... ... ... ... ... ...
01500B    ... ..C C.. ... ... ..G ... ... ... ... ... ... ... ... ... ...
00143C    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00153D    ... G.G ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00707D   ... ..? ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
01706D   ... ?.. ..G G.. ... ..G ... ... ... ... ... ... ... ... ... ...
0724E     ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00156E    ... ... C.G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00713E    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
01625E    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
01105F    ... ..C G.G GGG G.. ..G ... ... ... ... ... ... ... ... ... ...
00258G    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00849G    ... ... ?.G ... ... ..G ... ... ... ... ... ... ... ... ... ...
01687G    ... ... C.G ... ... ..G ... ... ... ... ... ... ... ... ... ...
0688H*    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00965H    ... ... ..G ... ... ..G ... ... ... ... ... ..- ... ..A ... ...
02357H    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00647I    ... ... ... ... ... ... ... ... ... ... ... ... ... ... G.. ...
00330J    ... ... .T. .G. .A. ..G ... ... ... ... ... ... ... ... ... ...
00979J    ... ... .T. .G. ... ..G ... ... ... ... ... ... ... ... G.. ...
01748J    ... ... .T. .G. .A. ..G ... ... ... ... ... ... ... ... ... ...
00363K    ... ... ..C ... ..C ..G ... ... ... ... ..G ... ... ... ... ..T
01169K    ... ... ..G G.. ..C ..G ... ... ... ... ... ... ... ... ... ..T
01763K    ... ... ..G G.. ..C ..G ... ... ... ... ... ... ... ... ... ..T
0592L     ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00403L    ... ... ..G ... ... G.G ... ... ... ... ... ... ... ... ... ...
01195L    ... ... ..G C.. ... ..G ... ... ... ... ... ... ... ... ... ...
01909L    ... ... ..G ... ... .CG ... ... ... ... ... ... ... ... ... ...
02776L    ... ... ..G ... ... ..G ... ... ... ... ... .G. ... ... G.. ...
0808M     ... ..G G.G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00409M    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
01752M    ... .GG ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
00448N    ... ... ... .G. ... ... ..A ... ... ... ... ... ... ... ... ...
01266N    ... .G. ... G.. ... ... ..A ... ... ... ... ... ... ... ... ...
02009N    ... .G? ... G.. ... ... ..A ... ... ... ... ... ... ... ... ...
02794N    ... .G. ... GG. ... ... ..A ... ... ... ... ... ... ... ... ...
00499X    ... ..C ... G.. ... ... ... ... ... ... ... ... ... ... ... ...
01296X*   ... ..C ... G.. ... ... ... ... ... ... ... ??? ??? ??? ??? ???
01893X    ... ... ... G.. ... ... ... ... ... ... ... ... ... ... ... ...
01094O    ... ..C ..G .G. ... ..G ... ... ... ... ... ... ... ... ... ...
0764P     ... ..G C.G ... ... ..G ... ... ... ... ... ... ... ... ... ...
02050P    ... ..C C.G ... ... ..G ... ... ... ... ... ... ... ... ... ...
7202Q     ... ... C.G .G. ..C ..G ... ... ... ... ... ... ... ... ... ...
7194R     ... ... G.. G.. ..C ... ... ... ... ... ... ... ... ... ... ...
01732R    ... ... ... G.. ..C ... ... ... ... ... ... ... ... ... ... ...
7050S     ... ... ..G ... ... G.G ..A ... ... ... ... ... ... ... ... ...
01100S*   ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
01746S    ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ...
7233T*    ... ..T ..G G.. ... G.. ... T.. CCC ... ... ... ... ... ... A..
7043U     ... ..C G.G ?.. ... G.G ..A ... ... ... ... ... ... ... ... ...
01842U    ... ... ..G ... ... G.G ..A ... ... ... ... ... ... ... ... ...
02191V    ... ... ... G.. ... ... ... ... ... ... ... ... ... ... ... ...
```

```
CONSENSC  TAT GCA ACA GGA GAC ATA ATA GGA GAC ATA AGA CAA GCA CAT TGT AAC

00099A    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
00101B    ... ... ... .-- -.. ... ... ... ..T ... ... ... ... ... ... ...
01500B    ... ... ... .-- -.. ... ... ... ..T ... ... ... ... ... ... ...
00143C    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
00153D    ... ... ... ... .C. ... ... ... ..T ..C ... A.. ... ... ... ...
00707D    ... ... ... ... .C. ... ... ... ..T ... ... ... ... ... ... ...
01706D    ... ... ... ... .C. ... ... ... ..T ... ... A.. ... .T. ... ...
0724E     ... ... ... ..G ... ... ... ... ..T ... ... ... ... ... ... ...
00156E    ... ... ... ..G ... ... ... ... ..T ... ... ... ... ... ... ...
00713E    ... ... ... ..G ... ... ... ... ..T ... ... ... ... ... ... ...
01625E    ... ... ... ..G ... ... ... ... ..T ... ... ... ... ... ... ...
01105F    ... ... ... ... ..T ... ... ... A.T ... ... ... ... ... ... ...
00258G    ... ... ... ... ..A ... ... ... ..T ... .A. ... ... T.. ... ...
00849G    ... ... ... ... ..A ... ... ... ..T ... .?. ..? ... T.. ... ...
01687G    ... ... ... ... ..A ... ... ... ..T ... .A. ..T ... T.. ... ...
0688H*    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
00965H    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
02357H    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
00647I    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
00330J    ... A.. ... ... A.A ... ... ... ..T ... ..G A.. ... ... ... ...
00979J    ... A.. ... ... C.A ... ... ... ..T ... ..G A.. ... ... ... ...
01748J    ... A.. ... ... C.A ... ... ... ..T ... ..G ... ... ... ... ...
00363K    ... ... ... ... ... ... ... ... A.T ... ... .T. ... ... ... ...
01169K    ... ... ... ... ... ... ... ... A.T ... ... .T. ... ... ... ...
01763K    ... ... ... ... ... ... ... ... ..T ... ... .G. ... ... ... ...
0592L     ... ... ... ... .C. ... ... ... ..T ... ... ... ... T.. ... ...
00403L    ... ... ... ..G .C. ... ... ... ..T ... ... ... ... T.. ... ...
01195L    ... ... ... ... .C. ... ... ... ..T ... ... ... ... T.. ... ...
01909L    ... ... ... ... .C. ... ... ... ..T ... ... ... ... ... ... ...
02776L    ... .G. ..G ..G .C. ... ... ... ..T ... ... ... ... T.. ... ...
0808M     ... ... ... ... ..A ... ... ... ..T ... ... .G. ... ... ... ...
00409M    ... ... ... ... ..A ... ... ... A.T ... ... ... ... ... ... ...
01752M    ... ... ... ... ..A ... ... ... A.T ... ... ... ... ... ... ...
00448N    ... ... ... ... ..T G.. ... ... ..T ... ... ... ... ... ... .C.
01266N    .T. ... ... ... ..T G.. T.. ... ..T ... ... .T. ... ... ... .C.
02009N    .T. ... ... ... ..T G.. ... ... ..T ... ... .T. ... ... ... .C.
02794N    ... ... ... ... ... ... ... ... ..T ... ... A.. ... ... ... .C.
00499X    ... ... ... ... ... ... ... ... ..T ... ..G ... ... T.. ... ...
01296X*   ??? ??? ??? ??? ??? ??? ??? ??? ??? ?GC T.T TG. ... T.. ... ...
01893X    ... ... ... ... ... ... ... ... ..T ... ..G ... ... ... ... C..
01094O    ... ... ... ... ... ... .C. ... ..T ... ... ... ... ... ... ...
0764P     ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
02050P    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
7202Q     ... ... ... ... .G. ... ... ... ..T ... ... ... ... ... ... ...
7194R     ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
01732R    ... ... ... ... .G. ... ... ... ..T ... ... ... ... ... ... ...
7050S     ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
01100S*   ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
01746S    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
7233T*    ... ... ... ..T ..T ... G.. ... ..T ... C.. ..G ... ... ... ...
7043U     ... ... ... ... ... ... ... ... ..T ... ..C ... ..C ... ... ...
01842U    ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ...
02191V    ... ... ... ... ... ... ... ... ..T ... ... A.. ... ... ... ...
```

```
CONSENS   ATT AGT AAA GA? AAA TGG AAT ?AA ACT TTA CAA AAG GTA AGT AAA AAA

00099A    ... .A. GG. C.A ..T ... ... AG. ... ... A.C .T. ... .AA GG. ..?
00101B    ... ... G.. ..G ... ... ... G.. ... ... ..T ... ... G.. G.. G.T
01500B    ... ... G.. ..G ... ... .T. A.. ... ... ... G.. ... G.. G.. ..T
00143C    ... ... GG. A.G ..C ... .G. C.C ... ... ... G.. ... GT. ... ...
00153D    ... ..C G.. A.G G.. ... ... A.. ... ... ... ... ... G.. G.. ...
00707D    ... ..C G.. A.G GC. ... ... A.. ... ... ... ... ..G GT. G.. ...
01706D    ... ..C G.. A.G ... ... ... G.. ... ... ... .G. TAG GTG ... TTT
0724E     ... ... G.. A.G .?. ... ... G.C ... ... ... G.. ... G.. .G. ...
00156E    ... ... G.. ..G ... ... ... G.C ... ... ... G.. ... G.. ... G..
00713E    ... ... G.. A.G ... ... ... G.C ... ... ... G.. ... G.. ... ...
01625E    ... ... G.. A.G ... ... ... G.C ... ... ... G.. ... G.. ... ...
01105F    ... ... G.. ?GG ... ... ... G.. ... ... ... G.. ... G.. GG. ?G.
00258G    ... .A. G.. A.G GC. ... ... G.C ... ... ... GC. ... G.. ... ...
00849G    ... .A. G.. A.G GC. ... ?.. G.C ... ... ... GC. ... G.. ... ...
01687G    ... .A. G.. A.G GC. ... G.. G.C ... ... ... GC. ... G.. ... ...
0688H*    ... ... GG. .GA ... ... ... G.C ... ... ..T ... ... ..C G.. ...
00965H    ... ... GG. .GA ... ... ... G.C ... ... ..T G.. ... ..C G.. ...
02357H    ... ... GG. .GA ..T ... ... G.C ... ... .TT ... ... ..C C.. ...
00647I    ... ..C GG. ..A ..T ... ... AC. ... ... G.. .G. ... .AG G.. ..T
00330J    ... ... GG. A.G GC. ... ... G.C ... ... ... ... ... G.. ... ...
00979J    ... ... GG. A.G GC. ... ... G.C ..C ... ... ... ... G.. ... ...
01748J    ... ... .G. A.G GC. ... ... G.C ..A C.. ... G.. ... G.. ... C..
00363K    G.. ... G?. ..G ... ... ... A.. ... ... A.. G.. ... GC. ... ...
01169K    G.. ... GG. ..G ... ... ... G.. ... ... A.. G.. ... GC. ... ...
01763K    G.. ... GG. .GG ... ... ... G.. ... ... G.. G.. ... G?G G.. ...
0592L     ... ... G.. A.G .C. ... ... G.. ... ... ... G.. ... G.G G.. ...
00403L    ... ... GG. A.G .C. ... ... G.. ... ... ... G.. ... G.G G.. ...
01195L    ... ... G.. A.G .C. ... ... G.. ... ... ... G.. ... G.G ... ...
01909L    ... ... G.. ..G ... ... ... A.C ... ... ... G.. ... G.. ... ...
02776L    ... ... G.. A.G .C. ... ... A.. ... ... ... G.. ... G.G ... ...
0808M     ... ... G.G C.G GC. ... ... A.. ... ... ... G.. ... G.. G.. G..
00409M    ... ... G.G C.G GC. ... ... A.. ... ... ... G.. ... G.. G.. G.G
01752M    ... ... G.G C.G GC. ... ... A.. ... ... ... G.. ... G.. G.. GCG
00448N    ... GAG GG. ..A ... ... ... AC. G.. ... ... .G. ... .AA GG. ...
01266N    ... GAG GG. ..A ... ... ... AC. G.. ... ... GG. ... .AA GG. ...
02009N    ... GAG GG. ..A ... ... ... AC. G.. ... ... GG. ..? .AA GG? ...
02794N    ... .AG GG. ..A ... ... ... AC. G.. ... ... GG. ... .AA GG. ...
00499X    ... ... .G. ..A ... ... ... A.T ... C.. ... G.. ... ..G GG. ...
01296X*   ... ... .G. ..A ... ... ... A.T ... C.. ... G.. ... .AG ..G ...
01893X    ... ... .G. ..A ... ... ... T.T ... C.. ... G.. ... GAG ... ...
01094O    ... ... G.. ..G .C. ... ... G.. ... ... ... .G. T.. G.. T.. ...
0764P     ... ... G.. A.G GC. ... ... G.C ... ... ... ... ... G.. .GG ...
02050P    ... ... G.. C.G GC. ... ... ..C ... ... ... ... ... G.. .GG ..T
7202Q     ..A ... .G. AGT .T. ... ... G.. ... ... ... ... ... .AG GG. G..
7194R     ... ... GG. ..A ..G ... ... A.. ... ... ... GC. ... .T. ... ...
01732R    ... ... GG. ?.A ..G ... ... A.. ... ... A.. GC. ... .T. ... ...
7050S     ... ... G.. ..A ..G ... ... G.. ... ... ... ... ... G.G GG. ...
01100S*   ... .C. GG. .GG .?. ?.. ... AC. C.C ... .C. G.A ... CCC ??? ???
01746S    ... .C. G.. .GG ... ... ... A.. ... ... ... G.. ... G.. ... ...
7233T*    ... ... GG. ..A .GT ... ... GC. ... ... ... G.. ..G .AG G.. ..T
7043U     ?.. ... G.. .GG ... GT. G.A T.. ... ... ..? G.. ... G.G GG. ..T
01842U    ... ... G.. ..A ..G ... ... G.. ... ... ... ... ... G.G G.. ...
02191V    ... ... GG. .CA G.G ... ... A.. ... ... ... GC. ... .AG G.. ...
```

```
CONSENSC      TTA GCA GAA CAC TTC CCT AAT --- AAA ACA ATA AAA TTT

00099A        ... AA. ..G ... ... AA. ..A ... --- ... ... .C. ...
00101B        ... AA. A.G .C. ... ... ... ... .G. ... ..C .T. ...
01500B        ... AA. A.G .CT ... ... ... ... .GT ... .C. CT. ...
00143C        ... CA. A.? G.? ... ... .?. ... .G. ... ?.. .C? ...
00153D        ... CA. C.. ... ... ... ... ... T.. G.C ... ... ...
00707D        ... CA. C.. A.. ... ... ... ... ..G ... ... .G. ...
01706D        AC. C.. C.T TT. CCT ?G. ..G ... --- G?C CAT ?.. ...
0724E         ... CA. A.. ... ... ... ... ... ..G ..C ... .GC ...
00156E        ... CA. A.? ... ... ... ... ... ... ?.. ... ..? ...
00713E        ... CA. ... ... ... ... ... ... ... ... ... .GC ...
01625E        ... CA. ... ... ... ... G.. ... ... ..C ..T .CC ...
01105F        ... CA. A.? .C. ... ... G.. ... ... C.. T.. ?GG ...
00258G        ... CA. A.. ... ... .TG ... ... ... C.. ... .GC ...
00849G        ... CA. A.. ... ... .TG ... ... ... ... ... .GC ...
01687G        ... CA. A.. ... ... .TG ... ... ... G.. ... ?GC ...
0688H*        ... AA. ..G ... ... ... ... ... .G. ... ... .G. ...
00965H        ... AA. ..G ... ... ... ... ... .G. ... ... ... ...
02357H        ... AA. ..G ... ... ... ... ... .G. ... ... ... ...
00647I        ... AA? ..G .T. ... ... ... ... .C. C.. CCC ?T. ...
00330J        ... AA. ... .T. ... ... ... ... ... ... ... .GT ...
00979J        ... AA. ..C .T. ... ... ?T. ... .C. ... T.. CGC ...
01748J        ... AA. ... .T. ... ... ... ... ... ... ... ... ...
00363K        ..G CA. ?.. ... ... ... ... ... ... T.. ... .GC ...
01169K        ..G CA. ... ... ... ... ... ... ... T.C ... .GC ...
01763K        ..G CA. A.. ... ... ... ... ... ... ... ... .GC ...
0592L         ... CG. ... ... ... ... ... ... ... ... ... .GC ...
00403L        ... CA. ... ... ... ... ... ... ... ... ... .GC ...
01195L        ... CA. ... ... ... ... ... ... ... ... ... .GC ...
01909L        ... CA. .GT ... ... ... ... AAA .C. ... ... .TC ...
02776L        ... CA. ... ... C.. ... T.. ... T.. .A. ... GTC ...
0808M         GCT A.. A.. .G. ... ... ... ... ... C.. ... .GC ...
00409M        C.. CA. A.. .G. ... ... ... ... ... ... ... .GC ...
01752M        CCC CAC CC. .G. ... ... ... ... ... C.. T.. GGT ...
00448N        ... AA. ..G ... ... ... GGG ... .G. ... ... ... ...
01266N        ... AA. ..G ... ... ... GGG ... .G. ... ... ... ...
02009N        ... AA. A.G ... ... ... ..G ... --- .G. ... ... ...
02794N        ... AA. ..G ... ... ... GGG ... .G. .A. ... ... ...
00499X        ... AA. ..G ... ... .T. ... ... ..C ... ... C.C ...
01296X*       ... AA. ..G ... ... ... ... ... .G? ??? ??? ??? ???
01893X        ... AA. ..G ... ... .T. .C. AAC .C. GT. CAC TT? ???
01094O        ... AA. .G. ... ... ... ... ... ... C.. ... GGC ...
0764P         ... CA. A.. ... ... ... ... ... ... ..G ... GCC ...
02050P        .AC AG. ACC .TT CC. TA. T.A ... .CG GTG ?.. .GC ...
7202Q         ... C.. A.. AGA ... ... GG. ... ... .AT ... .G? ...
7194R         ... .A. .GG T.T ... ..C .T. ... .G. ?GC .G. .T. ?..
01732R        ... .A. ..G ..T .C. ..C ... ... ... .GC ... .C. ...
7050S         ... AA. ..G .G. ... AA. ..A ... --- ?.. ... .C. ...
01100S*       ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
01746S        ... AA. ... .T. .AT AA. GCA ... TC. T.. ... ..C ...
7233T*        ..? C.T CCC A?? ??. ... .GC ... ... CA. TAG GTT ..?
7043U         .A. AG. CCC TT. A.T AAA CC. ... T.. GTT TGC CCC ...
01842U        ... AA. ..G .G. ... AA. ..A ... --- ... ... .C. ...
02191V        ... AG. ... ... ... ... ... ... ... ..C ... .CC ...
```

From the DNA sequences the amino acid sequences of the HIV-1 gp120 V3 regions were predicted (see Figure 5). These amino acid sequences occasionally had stop codons (indicated by *). The V3 loop (a 35 amino acid region between 2 conserved cystein (C) residues) is indicated in bold. As can be seen, there are more amino acid replacements in the flanking regions than in the loop itself. Most mutations are observed after the second cystein.

Figure 5. Aligned predicted HIV-1 gp120 V3 amino acid sequences of Akaki and Wonji HIV+ cohort participants. Legend: Sequences are aligned relative to a consensus subtype C sequence (CONSENSC), which was obtained from the Los Alamos 1998 databases. A dot (.) represents the same amino acid as the consensus, a dash (-) represents a deletion relative to the consensus, a question mark (?) represents a sequencing ambiguity. Sample codes: see Figure 4. In bold are the V3 loop encoding sequences. Sequences marked by a * have a high number of ambiguous positions. * inside the sequence equals a stop codon.

```
CONSENSC    IIIRSENLTN NVKTIIVHLN ESVEIVCTRP NNNTRKSIRI GPGQ--TFYA TGDIIGDIRQ

00099A      .V........ .A.I...Q.. .T........ .....E.M.. .......... ..........
00101B      .VV.A..... .A.IL..Q.. .......... ..?PG..V.. .......... .--.......
01500B      .V..A...E. .A.IL..Q.. .......... .......M.. .......... .--.......
00143C      .V.....I.. .........K KP........ .......M.. .......... ..........
00153D      .V........ .I.....Q.Q DP........ S..A...M.. .......... ..A......K
00707D      ?V........ .I.....Q.Q DP.K...... S..?...M.. .......... ..A.......
01706D      .V......AS D......Q.Q DP........ G..?.E.M.. .......... ..A......K
0724E       .V.......D ...N.....Q NP...E.... .......M.. .......... ..........
00156E      .V.......D ...N...Q.K .P........ .......M.. .......... ..........
00713E      .V.......D ...N...Q.K .P.Q...... .......M.. .......... ..........
01625E      ?V.......D ...N...Q.K .P.Q...... .......M.. .......... ..........
01105F      .V...Q.... .A.I...Q.K NP.....A.. ....GGGM.. .......... ......N...
00258G      .V........ ...I.....Q ?P........ .......M.. .......... ..E.....K.
00849G      .V........ ...I.....Q KP........ ....?..M.. .......... ..E.....??.
01687G      .V......A. ...I...Q.Q .P........ .......M.. .......... ..E.....KH
0688H*      ????.K.?.. .?.I...Q.K .P........ T......M.. .......... ..........
00965H      VV........ .A.I...Q.K .P........ .......M.. ...-...... ..........
02357H      .V........ .A.I...Q.K .P........ .......M.. .......... ..........
00647I      .......... .A.I...Q.. .T.G.N.... .......... ......A... ..........
00330J      .......... .A.V...Q.R QP...N.... ....IRNM.. .........T ..K......K
00979J      .......... .A.V...Q.E QP...N.... ....IR.M.. ......A..T ..Q......K
01748J      ?.......?. .A.V...Q.K QP...N.... ....IRNM.. .........T ..Q.......
00363K      ?..?...... KA.I...Q.K .P........ S...S..M.. .......... ......N..L
01169K      T......... .A.I...Q.K .P........ G....E.M.. .......... ......N..L
01763K      ?......... .A.I...Q.K .P........ G....E.M.. .......... .........R
0592L       ?V.K...... .A.I.....K KP........ S......M.. .......... ..A.......
00403L      .V.K...... .A.I.....K NP........ S......V.. .......... ..A.......
01195L      .V.K...... .A.I.....K NP........ S....Q.M.. .......... ..A.......
01909L      ?V.K.K..SD .A.I.....K KP........ .......T.. .......... ..A.......
02776L      .V.K...... .A.I.....K NP........ S......M.. ...R..A..G ..A.......
0808M       .V........ .......Q.K TP........ ....G..M.. .......... ..E......R
00409M      .V.......? .......Q.K TP........ .......M.. .......... ..E...N...
01752M      .V.....?.. .......Q.K TP........ ...R...M.. .......... ..E...N...
00448N      .......... .A.I...Q.K DP.K.D.I.. .....R.... .......... ...V......
01266N      .......... .A.I...Q.K DP.Q.N.... ...R.E.... .........F. ...VL....L
02009N      .......... .A.I...Q.K DP.Q.N.... ...?.E.... .........F. ...V.....L
02794N      .......... .A.I...Q.K NP.Q.N.... ...R.G.... .......... .........K
00499X      ..?....... .A.....Q.. AP...N.... .....E.... .......... ..........
01296X*     ..N..?.... .AQ....Q.K A?...N.... .....E.... ...??????? ????????C*
01893X      .......... .A.....Q.K AP...N.... .....E.... .......... ..........
01094O      .V....D..D .A.I...Q.K .P........ S....R.M.. .......... ....T.....
0764P       .V...K.... .A.I.....K NP........ .......M.. .......... ..........
02050P      .V........ ...I.....K DP........ T......M.. .......... ..........
7202Q       .V........ DA.I...R.. K..D...... G....R.M.. .......... ..G.......
7194R       ..V....I.. .A.....Q.. .T...N...A G...GE.... .......... ..........
01732R      ..V....I.. .A.....Q.. .T...N.... G....E.... .......... ..G.......
7050S       .V.......D K..I.....K D?.K.E.... G......V.. .......... ..........
01100S*     .....K...D .T.L...Q.K DP.Q...... .......M.. .......... ..........
01746S      .....K...D .T.L...Q.K DP........ .......M.. .......... ..........
7233T*      T..S..T.PH .A.I...Q.K TP...T.LP. S....E.V.L P......I.. ....V.....
7043U       ?V.......D ...I.....K DP...E.I.. G...G?.V.. .......... ........S.
```

```
01842U      ?V.......D ...I.....K DP.K.E.... G......V.. .......... ..........
02191V      .......I.. .A.....Q.. KT.D.N.... G....E.... .......... .........K

CONSENSC    AHCNISK?KW N?TLQKVSKK LAEHFPN-KT IKF

00099A      .....NGQN. .R..NM.KG? .K...NK.-. .T.
00101B      ......EE.. .E..H..GED .KKP....R. .I.
01500B      ......EE.. IK...E.GEN .KKP....S. TL.
00143C      ......GKN. SH...E.V.. .Q??..?.R. ??.
00153D      ......EKE. .K.....GE. .QQ.....*A ...
00707D      ......EKA. .K.....VE. .QQN...... .R.
01706D      .L....EK.. .E...R*V.F TPHFP?K.-? H?.
0724E       ......EK?. .D...E.GR. .QK....... .S.
00156E      ......EE.. .D...E.G.E .Q?......? .?.
00713E      ......EK.. .D...E.G.. .Q........ .S.
01625E      ......EK.. .D...E.G.. .Q....D... .T.
01105F      ......E.... .E...E.GG? .Q?P..D..P L?.
00258G      .Y...NEKA. .D...A.G.. .QK..L...P .S.
00849G      .Y...NEKA. ?D...A.G.. .QK..L.... .S.
01687G      .Y...NEKA. DD...A.G.. .QK..L...A .?.
0688H*      ......GG.. .D..H...E. .K......R. .R.
00965H      ......GG.. .D..HE..E. .K......R. ...
02357H      ......GGN. .D..L...Q. .K......R. ...
00647I      ......GEN. .T..ER.KEN .?.L....TP P?.
00330J      ......GKA. .D.....G.. .K.L...... .S.
00979J      ......GKA. .D.....G.. .KDL..?.T. LR.
01748J      ......RKA. .D...E.G.Q .K.L...... ...
00363K      ....V.?E.. .K..KE.A.. .Q?......S .S.
01169K      ....V.GE.. .E..KE.A.. .Q.......S .S.
01763K      ....V.GG.. .E..EE.?E. .QK....... .S.
0592L       .Y....EKT. .E...E.GE. .R........ .S.
00403L      .Y....GKT. .E...E.GE. .Q........ .S.
01195L      .Y....EKT. .E...E.G.. .Q........ .S.
01909L      ......EE.. .N...E.G.. .QG....KT. .I.
02776L      .Y....EKT. .K...E.G.. .Q..L.Y.*K .V.
0808M       ......EQA. .K...E.GEE ATKR.....P .S.
00409M      ......EQA. .K...E.GEE .QKR...... .S.
01752M      ......EQA. .K...E.GEA PHPR.....P LG.
00448N      ...T.EGE.. .TA..R.KG. .K....G.R. ...
01266N      ...T.EGE.. .TA..G.KG. .K....G.R. ...
02009N      ...T.EGE.. .TA..G?K?. .KK...K.-R ...
02794N      ...T.KGE.. .TA..G.KG. .K....G.RK ...
00499X      .Y....RE.. .N...E.RG. .K...L..N. .H.
01296X*     .Y....RE.. .N...E.K.. .K......?? ???
01893X      ...H..RE.. .Y...E.E.. .K...LTNTV H??
01094O      ......EET. .E...RLG*. .KG......P .G.
0764P       ......EKA. .D.....GR. .QK....... .A.
02050P      ......EQA. .......GRN YRTLPY*.TV ?S.
7202Q       ......RSI. .E.....KGE .PKR..G..N .?.
7194R       ......GE.. .K...A.I.. .EGY..I.R? RI?
01732R      ......G.... .K..KA.I.. .E..S....S .T.
7050S       ......EE.. .E.....GG. .K.R.NK.-? .T.
01100S*     .....TGG?? .TP.PE.P?? ?????????? ???
01746S      .....TEG.. .K...E.G.. .K.LYNA.SS .N.
7233T*      ......GES. .A...E.KEN ?PP??.S..Q *V?
7043U       ....?.EG.V E*..?E.GGN *RPFIKP.*V CP.
01842U      ......EE.. .E.....GE. .K.R.NK.-. .T.
02191V      ......GAE. .K...A.KE. .R........ .T.
```

As is shown in the phylogenetic tree of Figure 6, all the samples are of the C subtype. This can be concluded from the fact that all non-subtype C sequences clustered with significant bootstrap values away from the cohort participant-derived sequences and that the Los Alamos consensus C sequence clustered together with the cohort sequences.

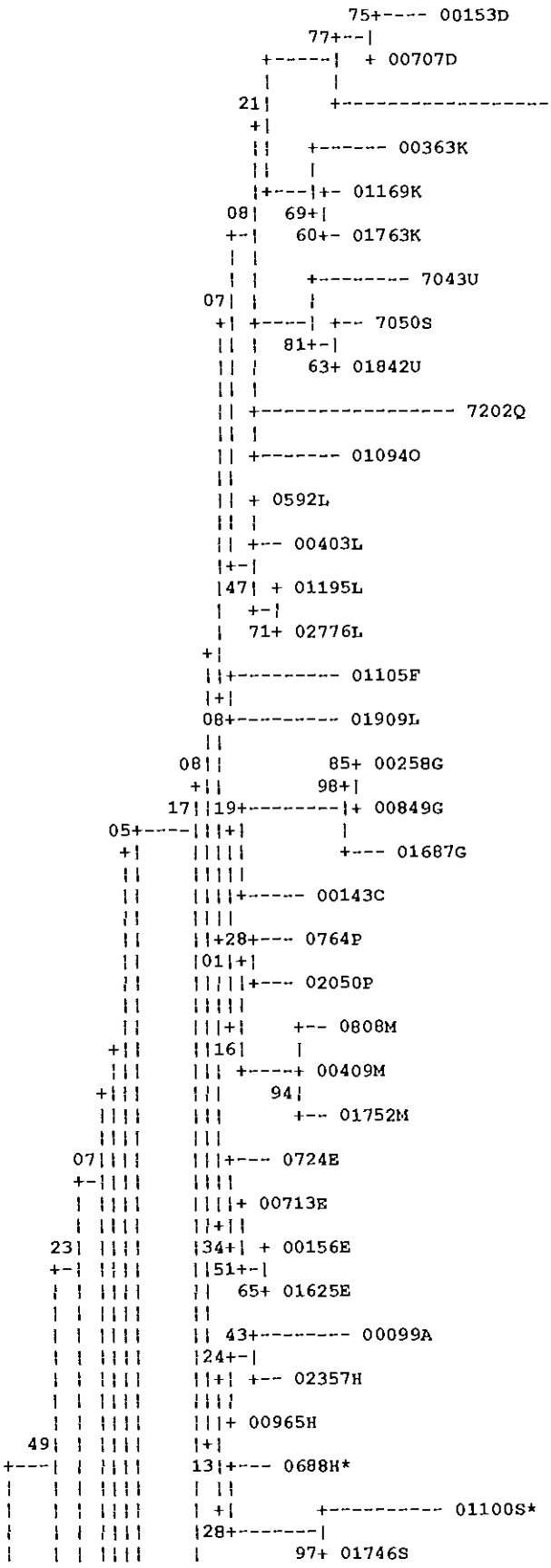
A second conclusion that can be drawn is, that sequences of the same individual usually clustered together in this tree, often with significant bootstrap values (>80). This implies that the HIV-1 DNA sequence differences between samples from the same individual (intra-person) were smaller than those between different individuals (inter-person). One sequence from subject S, 7050 was split from its group and aligned near a group of sequences from another subject from Wonji, U.

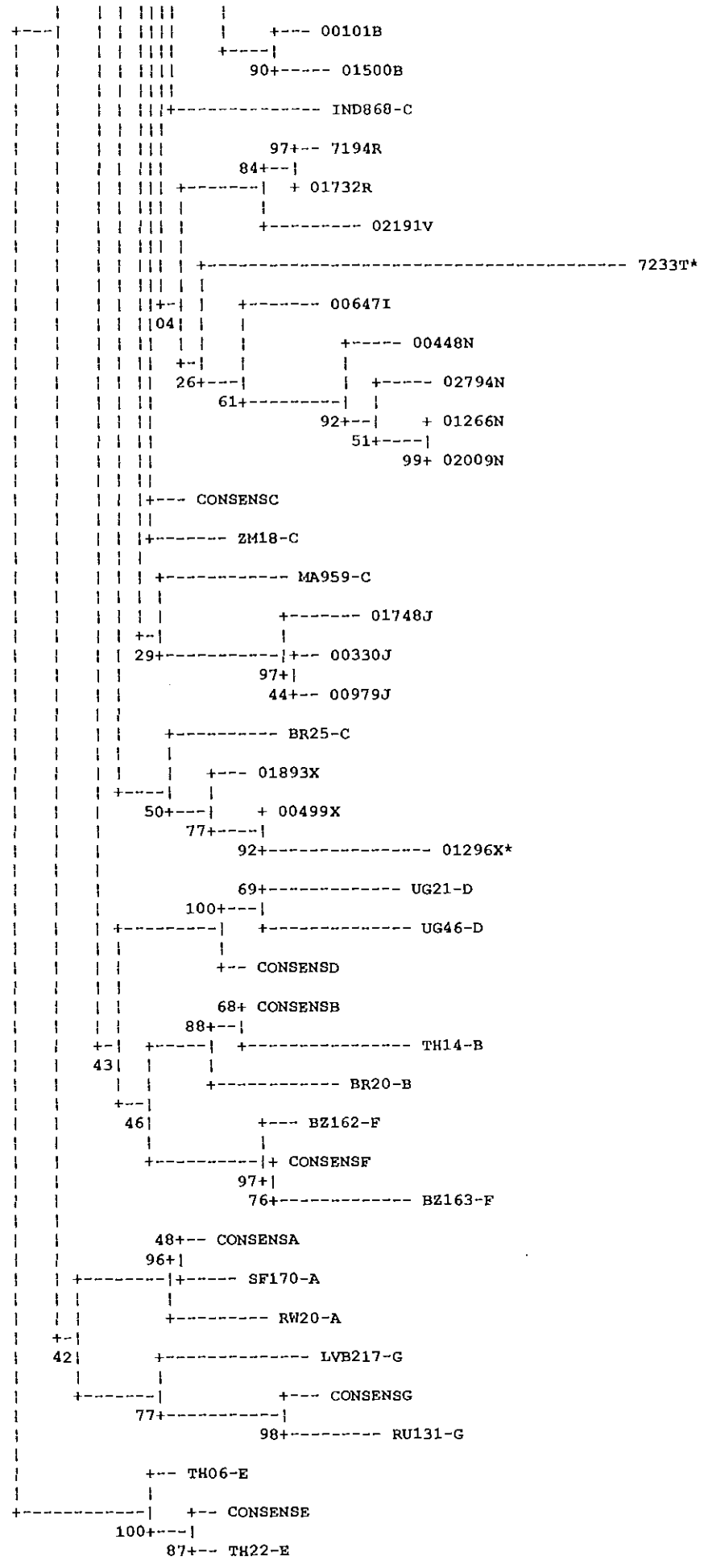
Third, the phylogenetic analysis demonstrated no separate clustering of samples from Akaki versus Wonji. Rather, sequences from both locales were intermixed. This indicates the homogeneity of the HIV-1 epidemic in these area's.

Fourth, most cohort HIV-1 gp120 V3 sequences clustered together, showing the higher relatedness of Ethiopian samples amongst each other than compared to C-subtype sequences from other locales. The only two exceptions were subjects J and X. The HIV-1 gp120 V3 DNA sequences from subject J clustered with a bootstrap value of 97 away from the other Ethiopian sequences and showed some (although not with significant bootstrap value) homology with sequences of HIV-1 subtype C from Malawi (MA959-C) and Zambia (ZM18-C). Subject X sequences also clustered together (and away from the other cohort sequences) with a bootstrap value of 77 and showed highest relatedness to a subtype C sequence from Brazil (BR25-C). The Indian C-subtype sequence ((IND868-C) clustered together with Ethiopian samples.

Fifth, there were no sequences found to be exactly identical to each other (which would have been expressed by the lack of any branch length in the phylogenetic tree). This implies that there were no contaminations in the PCR products sequenced.

Figure 6: Phylogenetic tree analysis of HIV-1 gp120 V3 sequences derived from cohort participants, as compared to HIV-1 consensus sequences. Legend: numbers at the branches indicate bootstrap values out of a total of 100 comparisons. Scale: each - is approximately equal to the distance of 0.003932




00101B
90
01500B
IND868-C
97
7194R
84
01732R
02191V
7233T*
00647I
04
00448N
26
02794N
61
92
01266N
51
99
02009N
CONSENSC
ZM18-C
MA959-C
01748J
29
00330J
97
44
00979J
BR25-C
01893X
50
00499X
77
92
01296X*
69
UG21-D
100
UG46-D
CONSENSD
68
CONSENSB
88
TH14-B
43
BR20-B
46
BZ162-F
CONSENSF
97
76
BZ163-F
48
CONSENSA
96
SF170-A
RW20-A
42
LVB217-G
CONSENSG
77
98
RU131-G
TH06-E
CONSENSE
100
87
TH22-E

In order to assess mutation speed in the various sequences, the proportions of silent and non-silent mutations (ds and dn) were calculated for several combinations of sequences. Figures 7A (for Akaki) and 7B (for Wonji) give examples of these analyses. Table 3 summarises some data of interest, including ds, dn and the frequently used dn/ds ratio.

First, it can be concluded that inter-person comparisons yielded higher ds and dn values (12-19%) than intra-person comparisons (2-5%). Second, the ds and dn values for Wonji cohort participants (both when the complete 279 bp of the V3 region and when the 105 bp of the V3 loop were used for the dn/ds calculations) were higher than for Akaki cohort participants, indicating a slightly higher accumulation of mutations in HIV-1 of individuals from Wonji. Third, it can be concluded that dn/ds calculations based on V3 loops give similar results as based on V3 loops plus flanking sequences.

Table 3. Proportions of synonymous (ds) and non-synonymous mutations (dn) for various selected groups of HIV-1 gp120 V3 sequences.

groups of sequences	ds	dn	dn/ds	number of bp analysed
inter-person comparisons:				
Akaki (n=16)	15.3 ± 2.8	11.8 ± 1.3	0.80	Complete (279)
Akaki (n=16)	14.3 ± 2.8	12.0 ± 1.2	0.84	V3 loop (105)
Wonji (n=5)	18.6 ± 4.1	18.4 ± 2.2	1.00	Complete (279)
Wonji (n=5)	18.3 ± 4.0	17.0 ± 2.1	1.00	V3 loop (105)
intra-person comparisons:				
L (n=5)	5 .0 ± 2.1	5.1 ± 1.1	1.0	V3 loop (105)
E (n=4)	4.3 ± 2.2	2.6 ± 0.8	0.6	V3 loop (105)
N (n=4)	1.8 ± 1.5	4.5 ± 1.2	2.5	V3 loop (105)

Figure 7A. Example analysis of frequency of synonymous mutations per synonymous site (ds) analysed for 16 different subjects from Akaki, over full length of sequences.

general information:

Distance: Jukes-Cantor corrected proportion of synonymous differences
No. of Codons in Subset: 93 of 93
Genetic Code: "Universal"
Input File Name: C:\MEGA\ASEFAW.MEG on Tue Jun 01 11:17:25 1999
Input Data: DNA
No. of Otus Used: 16 of 79
Otus Used:27 28 30 32 35 38 39 43 45 46 49 53 58 61 66 68
Gap Sites and Missing Information Data: All such sites were removed from the subset data OTU Labels

Sequences used for comparison:

1.. 00099A
2.. 00101B
3.. 00143C
4.. 00707D
5.. 00156E
6.. 01105F
7.. 00258G
8.. 00965H
9.. 00647I
10.. 00330J
11.. 00363K
12.. 00403L
13.. 00409M
14.. 01266N
15.. 01893X
16.. 0764P

Distances in the upper-right matrix
Standard Errors in lower-left matrix
'*' indicates an invalid distance value

OTUs	1	2	3	4	5	6	7	8
1		0.1530	0.1342	0.2311	0.1198	0.1005	0.1527	0.1006
2	0.0626		0.0335	0.1165	0.0561	0.0436	0.0668	0.0693
3	0.0587	0.0275		0.1193	0.0694	0.0332	0.0334	0.0346
4	0.0808	0.0532	0.0545		0.1575	0.1156	0.1419	0.1474
5	0.0548	0.0358	0.0406	0.0637		0.0557	0.1040	0.1080
6	0.0492	0.0311	0.0273	0.0528	0.0356		0.0663	0.0688
7	0.0625	0.0390	0.0274	0.0595	0.0500	0.0388		0.0937
8	0.0502	0.0405	0.0284	0.0619	0.0519	0.0402	0.0476	
9	0.0962	0.0666	0.0780	0.0781	0.0708	0.0692	0.0821	0.0736
10	0.0708	0.0344	0.0533	0.0750	0.0472	0.0527	0.0615	0.0607
11	0.0710	0.0565	0.0644	0.0731	0.0764	0.0684	0.0720	0.0750
12	0.0568	0.0311	0.0318	0.0528	0.0429	0.0309	0.0388	0.0276
13	0.0594	0.0313	0.0320	0.0532	0.0433	0.0311	0.0390	0.0405
14	0.0806	0.0540	0.0720	0.0835	0.0702	0.0686	0.0771	0.0729
15	0.0965	0.0616	0.0826	0.0983	0.0817	0.0765	0.0899	0.0849
16	0.0540	0.0313	0.0400	0.0596	0.0358	0.0311	0.0460	0.0477

OTUs	9	10	11	12	13	14	15	16
1	0.3034	0.1869	0.1873	0.1300	0.1398	0.2346	0.3041	0.1181
2	0.1725	0.0520	0.1292	0.0436	0.0440	0.1219	0.1506	0.0440
3	0.2197	0.1148	0.1591	0.0446	0.0450	0.1958	0.2402	0.0685
4	0.2218	0.2078	0.1987	0.1156	0.1165	0.2508	0.3158	0.1422
5	0.1889	0.0928	0.2135	0.0792	0.0798	0.1892	0.2376	0.0561
6	0.1847	0.1154	0.1809	0.0433	0.0436	0.1851	0.2177	0.0436
7	0.2429	0.1503	0.1958	0.0663	0.0668	0.2231	0.2790	0.0905
8	0.1981	0.1426	0.2038	0.0336	0.0693	0.1983	0.2485	0.0939
9		0.1518	0.2827	0.1756	0.2001	0.1186	0.1630	0.2434
10	0.0621		0.1936	0.1112	0.1121	0.1147	0.1048	0.1290
11	0.0912	0.0719		0.1674	0.1687	0.2508	0.2833	0.1962
12	0.0671	0.0516	0.0653		0.0436	0.1896	0.1896	0.0665
13	0.0729	0.0521	0.0659	0.0311		0.1866	0.2195	0.0670
14	0.0534	0.0524	0.0835	0.0696	0.0692		0.1638	0.2283
15	0.0647	0.0504	0.0914	0.0703	0.0772	0.0643		0.2796
16	0.0823	0.0563	0.0721	0.0388	0.0391	0.0783	0.0901	

Final Result: ds = 14.33% +/- 2.80%. Thus, when comparing the above 16 complete gp120 V3 sequences amongst each other in a pair-wise manner, the average percentage of actual silent mutations compared to the possible silent mutations was 14.33 % with a standard deviation of 2.80%.

Figure 7B. Example analysis of frequency of synonymous mutations per synonymous site (ds) analysed for 5 different subjects from Wonji, over V3 loop only.

general information:

Distance: Jukes-Cantor corrected proportion of synonymous differences.
No. of Codons in Subset: 37 of 93 (V3-loop only)
Genetic Code: "Universal"
Gap Sites and Missing Information Data: All such sites were removed from the subset data
Input File Name: C:\MEGA\ASEFAW.MEG on Tue Jun 01 11:27:56 1999
Input Data: DNA
No. of Otus Used: 5 of 79
Otus Used:70 72 75 78 79
Analysed sequences:

1.. 7202Q
2.. 01732R
3.. 01746S
4.. 01842U
5.. 02191V

Distances in the upper-right matrix, Standard Errors in lower-left matrix, '*' indicates an invalid distance value

OTUs	1	2	3	4	5
1		0.1793	0.1373	0.1813	0.2928
2	0.0928		0.1378	0.2082	0.0858
3	0.0813	0.0816		0.0436	0.1399
4	0.0939	0.1023	0.0440		0.2115
5	0.1273	0.0616	0.0830	0.1040	

Final Result: ds = 16.18% +/- 6.09%. Thus, when comparing the above 5 partial gp120 V3 sequences amongst each other in a pair-wise manner, the average percentage of actual silent mutations compared to the possible silent mutations was 16.18 % with a standard deviation of 6.09%.

7. DISCUSSION AND CONCLUSION

As can be concluded from the results presented in Table 2 and in Figures 8A and 8B, the average follow up time of 2 and a half years for both Akaki and Wonji cohort participants did not allow for the detection of significant declines of CD4 T-cells or increases of HIV-1 viral loads. Rather, a stable situation was encountered for both HIV-1 disease progression markers studied. Apparently the follow up time of the cohort participants is not yet long enough to allow for a clear detection of decrease of CD4 T-cells or increase of HIV-1 viral loads, being important surrogate markers for HIV-1 disease progression, as reported for other cohorts.

From the above 23 cohort participants, a total of 75 plasma samples (17 from Akaki, 6 from Wonji) was subjected to cDNA preparation and nested PCR. As a result, 61 PCR positive products were generated. This equals a success rate of 81%. The reasons why 19% of the samples were not amplified can be various:

1. The specific primer combination ED31/33 followed by M13 5'F/M13 3'R did not anneal to the target sequence. In our setting this is a rather unlikely situation, since for all PCR negative experiments on a certain subject, there were also PCR positive results on that same subject (see Table 2). It is unlikely that PCR primers fitting to a HIV-1 target sequence of a given individual at a given time point, will not anneal to HIV-1 target sequences of that same individual at a different time point. This may be best illustrated by the detected % of intra-person mutations (both ds and dn) in this study, which is in the order of 2-5% only. On top of this, the primers for HIV-1 PCR are chosen in a very conserved region of HIV-1, which will mutate at considerable lower frequency than 2-5%.
2. There was not enough target RNA. It could be that the input RNA for the cDNA reaction was too low and therefore created some PCR negative results. However, the mean ^{10}log viral load for PCR negative samples was 3.24 and for PCR positive samples 3.35, which is not significantly different. Therefore, it is unlikely that a too low input of viral RNA is the most important cause for lack of success of certain PCR reactions.
3. Technical errors. This always remains a possibility and its contribution to PCR negativity, although unpredictable, will be a reality.

Out of the 61 PCR products 53 resulted in an interpretable DNA sequence. This equals a sequencing success rate of 87%. So, in summary, of the 75 plasma samples, 53 (71%) resulted in an interpretable HIV-1 gp120 V3 sequence. This result is in accordance with other sequencing studies performed by EHNRI-ENARP, where on the average 1 out of 4 plasma-derived sequences is unsuccessful (personal communication Rinke de Wit). Reasons for failure of the sequencing reaction could be:

1. Too low quantities of input DNA (normally, 100 ng should be added to a cycle sequence reaction. In fact, reinspection of the agarose gels revealed that 5/8 of the samples that failed to be sequenced had either no, very faint, or double bands after nested PCR.
2. The presence of several abundant viral clones in the original plasma sample. Normally, one HIV-1 viral clone is predominant in a plasma sample of an infected subject at a given time point. However, sometimes there are >1 HIV-1 virus clones present in the plasma at sufficient quantities to allow for PCR amplification and thus the generation of a mixed PCR product, which will give an uninterpretable sequencing result.
3. Technical errors during the sequencing procedure.

The relative contribution of each of the above explanations is not clear, but certainly any of them could have played a role.

Based on the DNA sequences of the 53 HIV-1 gp120 V3 regions, amino acid sequences were derived (as depicted in Figure 5). When these amino acid sequences are studied in more detail, the far majority of them would be predicted to belong to HIV-1 viruses of the non syncytium inducing (NSI) phenotype. As reported by Abebe et al. (47), amino acid positions 8 and 29 of the V3 loop may play a role in determining viral SI phenotype. According to literature, the SI phenotype of HIV-1 is associated to poor prognosis, accelerated decline of CD4 T-cells and usually to AIDS in many HIV-1 subtypes, like B, D, E (48). In our data set, there were 3 sequences from two cohort participants (M and N) that had a positively charged amino acid residue at position 8 (01752M, 01266N and 02794N: see Figure 5). No subjects had a positively charged mutation at position 29. When studied in some more detail, the visit 01752 of subject M was characterised by a very low CD4 T-cell count (101/ul) and relatively low ^{10}log viral load (2.63). Subject M in fact is an AIDS patient (according to the clinical WHO definition) for a period of almost 3

years. So, it could well be that in this patient HIV-1 SI phenotype viruses have emerged. The second subject (N), had reasonable CD4 T-cell counts at both visits 01266 and 02794: 352 and 302/ul, respectively. However, the ^{10}log viral loads at those visits were very high: 4.94 and 5.64 respectively. It could well be, that amongst these very high viral loads of subject N, viruses of the SI phenotype are present. According to literature (49) there are 2 other positions involved in determining "SI-ness": positions 11 and 27 (in our Figure 5). However, in our case, there were no positively charged residues detected in either of those positions for all 53 predicted amino acid sequences studied. This again strengthens the prediction that the far majority of HIV-1 positive subjects of the EHNRI-ENARP cohort studies in Akaki and Wonji harbour viruses of the NSI phenotype.

The phylogenetic analyses of the DNA sequences from the 23 cohort participants have led to some important conclusions:

1. All HIV-1 sequences derived from the cohort participants are of the C-subtype. This confirms earlier observations, by us (39,47) and others (13) that the HIV-1 epidemic in Ethiopia is of a rather homogeneous nature. The explanation would be, that HIV-1 got introduced into Ethiopia in the early 1980's and managed to saturate its main transmission network (commercial sex workers) very rapidly. Since then, introductions of possible other HIV-1 subtypes have been counteracted, because of the phenomenon of viral exclusion: infection with HIV-1 C subtype protects against a second infection with another HIV-1 subtype.

2. The intra-person variation of HIV-1 is much lower than the inter-person variation, as reflected by tight clustering of DNA sequences on a per-individual basis. This is a well established fact from literature. In general, HIV-1 tends to be clonal in the very early phases of infection and then, dependent on the selective force of the immune system of the host, mutates to a swarm or quasispecies of related viruses over time. However, within individuals the mutation percentage is reportedly between 2 and 5% (50), which is in exact concordance with our results presented in Table 3. In fact, there was only one sample (from subject S): 7050, that clustered with DNA sequences from another subject (subject U). This could either be due to a contamination or mix-up of samples, or to the fact that both individuals were infected by a common source. This could well be possible, given the fact that these subjects are both residents from Wonji, which is a rather closed community.

3. There is no evidence for a founder effect of HIV-1 in Akaki versus Wonji. If the HIV-1epidemics
in both locales would have been introduced separately and would have had time to evolve
independently, this would have been reflected by a separate clustering of the DNA sequences
derived from these locales. However, this is not the case: the sequences branch in an
intermixed way, demonstrating their strong phylogenetic relatedness. This observation is again
consistent with the above described fact introduction and spread of the HIV-1 subtype C epidemic
in Ethiopia: both Akaki and Wonji got into contact with this virus subtype at an early stage and
were not invaded by other subtype epidemics later on. The silent and non-silent mutation analysis
of Wonji versus Akaki (Table 3) revealed a possible higher mutation frequency in Wonji. This
could imply that the epidemic has reached this locale earlier in time than Akaki.

4. Ethiopian C-subtype HIV-1 sequences show high inter-relatedness and cluster most of the time
with subtype C sequences from India. This has been reported in literature before (e.g. 39). It could
be that the Ethiopian C-subtype viruses have contributed to the raise of the HIV-1 epidemic in
India, which took place several years after the introduction of HIV-1 into Ethiopia.

5. The fact that amongst our 53 sequences there were no identical ones, supports the idea that there
were no PCR contaminations and that all presented DNA sequences are genuine.

Finally, it was attempted to correlate mutation patterns, as expressed by ds and dn, to immune status

or viral loads of cohort participants. According to Holmes & Zanotto, (1998) a ratio of dn/ds < 1.0
is indicative of purifying (negative) selection, dn/ds >1.0 for positive selection and dn/ds =1 for
neutral mutation, without selection. When comparing dn/ds ratios of cohort participants as
mentioned in Table 2 to CD4+ T-cell counts or viral loads, no significant correlation was found (not
shown). In other words, subjects with relatively high CD4+ T-cell counts, who would be suspected
to have a still rather strong immune system, did not show signs of strong positive selection (dn/ds
>1.0). Vice versa, subjects with high viral loads and low CD4 T-cell counts did not show sequence
homogeneity in their consecutive visits.

In conclusion, all 23 HIV-1 positive cohort participants tested harboured viruse of C subtype, which
indeed is the subtype that is predominantly present in Ethiopia. There is no difference between the
molecular structures of the Akaki and Wonji epidemics, indicating a common origin, rather than a
founder effect. The intra-person differences of HIV-1 gp120 V3 sequences were smaller than the
inter-person differences. More than 95% of the samples were predicted to belong to HIV-1 viruses
of the NS1 phenotype. The HIV-1 epidemic in Wonji might be slightly older than the one in Akaki.
Finally, the proportions of non-synonymous (dn) versus synonymous (ds) mutations are variable per subject and do not dependent on CD4 counts or viral load.

8. REFERENCES

1. Kuiken, Carla L., Gabriel Zwart, Elly Baan et al, (1993). **Increasing antigenic and genetic diversity of the V3 variable domain of the human immunodefiecency virus envelope protein in the course of the AIDS epidemic.** *Proc. Natl. Acad. Sci. USA* Vol. 90 pp. 9061-9065.

2. Shioda, T., Levy J. A., and Cheng-Meyer C. (1992). Proc. Natl. Acad. Sci. USA Vol. 89 pp. 9534-9438.

3. Freed, E. D., Delwart E. L., Buchschacher G. L., and Panganiban A. T. (1992). *Proc. Natl. Acad. Sci. USA* Vol. 89 pp.70-74.

4. de Jong, J. J., de Ronde A, Kuelen W., Termette M., and Goudsmit J. (1992). *J. Virol.* Vol. 66 pp. 6777-6780

5. Korber, Bette T. M., Kersti MacInnes, Randall F. Smith and Grerald Meyers (1994). **Mutational Trends in V3 loop protein sequences observed in different genetic lineages of HIV-1.** *J. Virol.* Vol. 68 pp. 6730-6744.

6. Meyers, Gerard , Kristi MacInnes, and Lynda Meyers (1993). **Phylogenetic moments in the AIDS epidemic.** In Morse, Stephen S. (ed) Emerging Viruses. Oxford University Press. Oxford. Pp. 120-137.

7. Baumforth, K.R.N., P. N. Nelson, J. E. Dinghy, D. O'Neil and P. G. Murray (1999). **The Polymerase Chain Reaction.** *J. Clin. Pathol.* Vol. 52 pp. 1-10

8. Leonard, C. K., Spellman et al (1990). *J. Biol. Chem.* Vol. 265 pp. 10373-10382

9. Ohagen, Asa, Roald B. Luftig, Alise S. Reicin et al (1997). **The morphology of immature HIV-1 viron.** *Virol.* Vol. 228 pp. 112-114

10. Goudsmit, Jaap, (1995). **Effects of variation within HIV-1 envelope region containing neutralising epitopes and virulence determinants.** In Gibbas, Adrian, Charles H. calither and Fernando Garcia Arenal (ed) Molecular Basis of virus evolution. Cambridge University Press Cambridge. Pp. 165-174

11. Wyatt, Richard, Peter Kwong, Elizabeth Desjardins, et al (1998). **The antigenic structure of the HIV gp 120 envelope glycoprotein.** *Nature* Vol. 393. pp. 705-711.

12. Kwong, Peter D., Richard Wyatt, James Robinson et al (1998). **Structure of an HIV gp 120 envelope glycoprotein in complex with CD4 receptor and a neutralising human antibodies.** *Nature.* Vol. 393. pp . 648-659.

13. Sherefa, Kimder, Bo Johansson, Mika Salminen and Andens Sonnerborg (1998). **Full length sequence of human immunodeficiency virus type 1 subtype A, recombined with subtype C in the *env* V3 domain.** *AIDS. Res. Human. Retrovirus.* Vol. 14 pp. 289-291

14. Boom, R. Sol C.J.A., Salimans MMM. Et al (1990). **A rapid and simple method for purification of nucleic acids.** *J. Clin. Microbiol.* Vol.28 pp 495-503.

15. Charneau, P. Borman A. M. Quillent C. et al (1994). **Isolation and envelope sequence of a highly divergent HIV-1 isolate: Definition of a new HIV-1 subgroup.** *Virology.* Vol. 205 pp.247-257.

16. Delwart, E. L. , Shaper E.G. Louwagie j. et al (1993). **Genetic relationships determined by a DNA hetroduplex mobility assay : analysis of HIV-1 genes.** *Science* Vol.262 pp.1257-1261.

17. Eshete, H. and Sahlu T. (1996). Review article: **The progression of HIV/AIDS in Ethiopia.** *Eth. J. Heal. Dev.* Vol.10 pp. 179-190.

18. Esparaza, J. and Osmannov D. (1993). **The development and evaluation of HIV-1 vaccines.** *Current Opinion of Diseases. Vol. 6 pp. 218-229.*

19. Kieviet, T , Van Gemen B, Van Strijp D, et al (1991). **NASAB TM isothermal enzymatic in vitro nucleic acid amplification optimised for the diagnosis of HIV-1 infection.** *J. Virol Methods. Vol. 35 pp.273-286.*

20. Koot, M. Vos AHV., Keep RPM. et al (1992). **HIV-1 Biological phenotype in long term infected individuals, evaluated with an MT-2 cocultivation assay.** AIDS Vol. 6 pp.49-54.

21. Kostikis, L. G. , Cao Y. , Nagi H. , et al (1995). **Genetic analysis of HIV-1 strains from patients in Cyprus : identification of a new subtype designated Subtype I.** J. Virol. Vol.69 pp 6122-6130.

22. Kumar, S. , Tamura K. Nei M. (1993). **Molecular evolutionary genetics analysis (MEGA). Version 1.01.** Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park.

23. Kunanusont, C. FoyH.M. , Kreiss J.K. st al (1995). **HIV-1 subtypes and male-to-female transmission in Thailand.** *Lancet* Vol. 345 pp. 1078-1083.

24. Leitner, T. Escanilla D. Marquina S., et al (1995. **Biological and molecular characterisation of subtype D,G. and A/D recombinant HIV-1 transmission in Sweden.** Virology. Vol. 209 pp.136-146.

25. Louwagie, J. Janssen W. Mascola J. et al (1995). **Genetic diversity of the envelop glycoprotein from human immunodeficiency virus type 1 isolates of African origin.** *J. Virol.*. Vol.669 pp 263-271.

26. Sharp, P. M. , Robertson D.L. , Gao F. and Hahn B. H. (1994). **Origins and Diversity of human immunodeficiency viruses.** AIDS Vol.8 pp. S27-S42.

27.UNAIDS and WHO Report on the Global HIV/AIDS Epidemic. (1998) **Global estimates of the HIV/AIDS epidemic as of end 1997.** Geneva , Switzerland.

28. World Health Organisation (1994). **WHO Guidelines for standard HIV isolation procedures.** Geneva Switzerland.

29. Zwart, G. , Wolfs T. F. W. , Brookelman R. , et al (1993).**Greater diversity of the HIV-1 V3 neutralisation domain in Tanzania compared with Netherlands : Serological and analysis.** *AIDS* Vol.7 pp 467-474.

30. Andeweg, A. C., P. Leeflang A. D., M. E. Ostterhaus, and M. L. Bosch (1993). **Both the V2 and V3 regions of the human Immunodeficiency virus type 1 surface glycoprotein functionally interact with other envelope regions in syncytium formation.** *J. Virol.* Vol. 67 pp.3232-3239.

31. de Jong, J. J., J. Goudsmit, W. Kuelen B., Klaver W. J. A. Krone, M. Tresmette, and A. de Ronde (1992(. **Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity.** *J. Virol.* Vol. 66 pp. 757765.

32. McKeating, J. A., J.Cordell, C. J. Dean and P. Balfe (1992). **Synergistic interactions between ligands binding to the Cd4 binding cite and V3 domain of human immunodeficiency virus type 1 gp 120.** *Virology* Vol. 191 pp. 732-742

33. Cheng-Mayer, C., D. Seto, M. Tateno, and J. A. Levy (1988). **Biologic features of HIV-1 that correlate with virulence in the host.** *Science* Vol. 240. pp. 80-82

34. Zhu, T., H. Mo, N. Wang, D. S. Nam, Y. Cao, R. A. Koup, and D. D. HO (1993). **Genotypic and phenotypic characterisation of HIV-1 in patients with primary infection.** *Science.* 261, 1179-1181

35. Kassa, E., Rinke de Wit, T.F., Hailu, E., Girma, M. et al (1999). **Evaluation of the World Health Organisation staging system for HIV infection and disease in Ethiopia: Association between clinical stages and laboratory markers.** *AIDS* Vol. 13 pp. 381-389.

36. Tsega, E., Mengesha B., Nordenfelt E., et al (1988). **Serological survey of HIV infection in Ethiopia.** *J. Med. Virol.* Vol. 28 pp. 21-24.

37. Ayehunie, S., Johansson B., Sonnerborg A., et al (1990) **New subtype of HIV-1 in Ethiopia.** *Lancet* Vol. 336. pp. 942

38. Salminen, M., Johansson B.,Sonnerberg A., et al (1996). **Full-length sequence of an Ethiopian HIV-1 isolate of genetic subtype C.** *AIDS. Res. Hum. Retrovir.* Vol. 12 pp. 1329-1339

39. Abebe, A., Kuiken C. L., Goudsmit J., et al (1997). **HIV-1 subtype C in Addis Abeba, Ethiopia.** *AIDS. Res. Hum. Retrovir.* Vol. 13 pp. 1071-1075

40. Reddy, Bindu, and John Yin (1999). **Quantitative intracellular kinetics of HIV-1.** *AIDS. Res. Hum. Retrovir.* Vol.15 pp. 273-283

41. Lole, Kavita S., Robert C. Bollinger, Ramesh S. P. et al (1999). **Full length sequence of HIV-1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination.** *J. Virol.* Vol. 68 pp. 152-160

42. Kuiken, Carla L., and Bette T. M. Korber (1994). **Epidemiological significance of intra- and inter-person variation of human immunodeficiency virus type 1.** *AIDS. Vol. 8 Suppl. 1.* pp. S73-S83

43. Bjorndal, Asa, Andreas Sonnerborg, Charlotte Tscherning et al (1999). **Phenotypic characteristics of human immunodeficiency virus type 1subtype C isolates of Ethiopian AIDS patients.** *AIDS. Res. Hum. Retroviru.* Vol. 15 pp.647-653

44. Simmonds, P. Zhang L. Q., McOmish F., et al (1991) *J. Virol.* Vol. 65 pp. 6266-6277.

45. Wolfs, T. F. W., Zwart G. Bakker M., and Goudsmit J. (1992). Virology Vol. 189. Pp103-110.

46. Holmes, C. Edwards and Paolo M. de A. Zanotto (1998). **Genetic Drift of Human immunnodeficiency virus type 1 ?** *J.Virol.* Vol. 72 pp.886-887.

47. Abebe, A., Demissie, D., Goudsmit, J., Brouwer, M., Kuiken, C., Pollakis G, Schuitemaker H., Fontanet, A., and Rinke de Wit, T.F. (1999). **HIV-1 subtype C syncytium- and non-syncytium-inducing phenotypes and coreceptor usage among Ethiopian patients with AIDS.** *AIDS* 13, - , in press.

48. Schuitemaker, H., Koot, M., Kootstra, N..A., *et al.* (1992): **Biological phenotype of human immunodeficiency virus type 1 clones at different stage of infection: progression of disease is associated with a shift from monocytotropic to T-cell tropic virus populations.** *J.Virol.* 66:1354-1360.

49. Fouchier, R.A.M., Groenink, M., Kootstra, N.A., Tersmette, M., Huisma, H.G., Schuitemaker, H. (1992). **Phenotype associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule.** J.Virol.66, 3183-3187.

50. Delwart, E.L., Busch, M.P., Kalish, M.L., Mosley, J.W., Mullins, J.I. (1995). **Rapid molecular epidemiology of of human immunodeficiency virus transmission.** AIDS Res.Hum.Retrovir. 11, 1081-1093.

Table 1. An overview of clinical manifestations in HIV positive cohort participants from Akaki, 1997-1999.

Adapted from Kassa et al., 1999; n = number of cohort participants with the indicated clinical symptoms.

Clinical condition	n	median CD4 T-cells ($x10^6$/l)
Asymptomatic	48	324 (89-966)
Minor weight loss	15	270 (79-504)
Pulmonary tuberculosis	9	140 (45-436)
Persistant generalised lymphadenopathy	8	375 (286- 798)
Oral candidiasis	6	268 (94-436)
Major weight loss	5	317 (151-490)
Herpes Zoster	4	144 (92-270)
Minor mucocutanious lesions	2	140 & 317
Herpes simplex infection >1 month	1	198

Table 2: Viro-immunological data on cohort participants selected for gp120 V3 sequencing.

Legend: Subject ID = code number of the pertinent cohort participant, Dob = date of birth; Lab ID = visit code number, CD4 = number of CD4 T-cells/ul of whole blood; CD8 = idem for CD8 T-cells; 4/8 = CD4/CD8 ratio; load = viral load in ^{10}log numbers of HIV-1 RNA molecules per ml plasma (NB the lower detection limit of Nuclisens is 80 RNA molecules per ml of plasma, which equals a ^{10}log value of 1.90; PCR = result of nested PCR reaction; sequence = result of sequencing reaction; ds = percentage of silent mutations per possible silent mutation site; dn = percentage of non-silent mutations per possible non-silent mutation site. The values of ds and dn refer to a comparison between the V3 loops of 2 sequences: one on the same line and one above. NT = not tested. NA = not applicable (when there is only 1 sequence, ds or dn cannot be calculated).

Sub. ID	Dob	Sex	Lab ID	Visit date	Days	CD4	CD8	4/8	load	PCR	Sequence	ds	dn
A	1960	M	00099	04-03-1997	1	406	624	0.7	3.60	+	+	NA	NA
			00670	27-08-1997	176	572	744	0.8	2.96	NT	NT		
			01564	23-03-1998	385	425	480	0.9	<1.90	-	-		
			02205	13-10-1998	587	533	730	0.7	3.20	NT	NT		
			02979	26-04-1999	837	521	605	0.9	2.30	NT	NT		
B	1955	M	00101	04-03-1997	1	149	1014	0.1	1.90	+	+		
			00671	27-08-1997	152	120	506	0.2	2.92	NT	NT		
			01500	09-03-1998	347	NT	NT	NT	2.56	+	+	0.00	0.42
			02099	17-09-1998	526	119	393	0.3	2.96	NT	NT		
			02867	15-03-1999	742	99	563	0.2	<1.90	NT	NT		
C	1965	M	00143	18-03-1997	1	NT	NT	NT	5.64	+	+	NA	NA
D	1957	M	00153	19-03-1997	1	198	991	0.2	2.45	+	+		
			00707	17-09-1997	186	233	1183	0.2	3.36	+	+	4.51	1.27
			01706	07-05-1998	416	288	1170	0.2	3.70	+	+	0.00	5.28
			02443	01-12-1998	635	312	1455	0.2	3.08	+	-		
			03160	03-06-1999	849	367	1748	0.2	NT	NT	NT		

E	1969	M	0724	26-01-1996	1	NT	NT	NT	2.63	+	+		
			00156	20-03-1997	417	179	1476	0.1	2.13	+	+	9.00	0.00
			00713	18-09-1997	587	144	1350	0.1	2.46	+	+	9.00	0.00
			01625	08-04-1998	782	185	1313	0.1	<1.90	+	+	4.43	0.00
			02189	07-10-1998	966	150	1215	0.1	<1.90	+	-		
			02972	22-04-1999	1235	128	1149	0.1	2.34	NT	NT		
F	1970	M	00256	14-04-1997	1	133	1065	0.1	5.04	NT	NT	NA	NA
			01105	24-11-1997	222	110	922	0.1	4.70	+	+		
			02944	13-04-1999	696	114	930	0.1	4.67	NT	NT		
G	1952	M	1087	16-02-1996	1	NT	NT	NT	3.65	-	NT		
			00258	14-04-1997	425	943	1185	0.8	3.80	+	+	0.00	0.00
			00849	22-10-1997	615	319	1229	0.3	3.66	+	+	0.00	0.00
			01687	29-04-1998	795	318	1290	0.2	3.77	+	+		
H	1975	F	0688	25-01-1996	1	NT	NT	NT	2.80	+	+		
			00324	14-05-1997	474	321	947	0.3	2.45	+	-		
			00965	12-11-1997	655	232	645	0.4	3.82	+	+	0.00	1.28
			01740	20-05-1998	831	351	837	0.4	2.86	+	-		
			02357	16-11-1998	1016	338	852	0.4	3.79	+	+	0.00	0.00
I	1971	F	0647	23-01-1996	1	NT	NT	NT	3.90	+	+	NA	NA
			00326	14-05-1997	476	530	1102	0.5	3.79	+	-		
			00966	12-11-1997	657	380	686	0.6	3.36	-	NT		
			01741	20-05-1998	832	459	792	0.6	2.90	+	-		
			02358	16-11-1998	1018	491	1058	0.5	4.20	-	-		
			03105	19-05-1999	1202	410	760	0.5	4.04	NT	NT		
J	1964	F	0578	18-01-1996	1	NT	NT	NT	3.04	-	NT		
			00330	15-05-1997	481	141	437	0.3	4.04	+	+		
			00979	13-11-1997	670	125	437	0.3	4.00	+	+	0.00	3.78
			01748	01-06-1998	870	36	510	0.1	4.38	+	+	0.00	3.78

K	1971	F	0976	12-02-1996	1	NT	NT	NT	2.89	-	NT		
			00363	21-05-1997	464	425	867	0.5	3.50	+	+		
			01169	03-12-1997	660	371	800	0.5	3.46	+	+	4.23	3.81
			01763	04-06-1998	843	420	958	0.4	2.98	+	+	0.00	2.54
			02470	03-12-1998	1025	307	861	0.4	3.56	-	NT		
L	1961	F	0592	19-01-1996	1	NT	NT	NT	4.15	+	+		
			00403	02-06-1997	468	250	451	0.6	3.64	+	+	4.11	1.26
			01195	10-12-1997	659	244	568	0.4	3.26	+	+	8.46	2.54
			01909	28-07-1998	889	263	504	0.5	3.85	+	+	0.00	5.15
			02776	16-02-1999	1092	398	890	0.4	3.77	+	+	8.37	7.93
M	1957	M	0808	01-02-1996	1	NT	NT	NT	< 1.90	+	+		
			00409	03-06-1997	495	89	850	0.1	2.30	+	+	4.30	3.79
			01752	02-06-1998	860	101	1431	0.1	2.63	+	+	4.48	1.23
			02512	09-12-1998	1065	103	1982	0.1	2.34	+	-		
N	1962	M	1043	14-02-1996	1	NT	NT	NT	4.68	-	NT		
			00448	11-06-1997	483	419	1511	0.3	4.01	+	+		
			01266	24-12-1997	681	352	1071	0.3	4.94	+	+	2.12	8.82
			02009	24-08-1998	922	446	2677	0.2	5.18	+	+	0.00	1.29
			02794	22-02-1999	1103	302	1229	0.2	5.64	+	+	4.29	6.70
X	1954	F	0779	31-01-1996	1	NT	NT	NT	< 1.90	-	NT		
			00499	26-06-1997	509	239	796	0.3	< 1.90	+	+		
			01296	14-01-1998	712	205	614	0.3	< 1.90	+	+	NA	NA
			01893	22-07-1998	901	280	786	0.4	< 1.90	+	+	4.23	1.25
			02682	20-01-1999	1083	213	451	0.5	< 1.90	+	-		
O	1965	F	1094	16-02-1996	1	NT	NT	NT	4.93	+	+	NA	NA
			00666	26-08-1997	194	94	508	0.2	4.85	NT	NT		
			01539	17-03-1998	391	122	682	0.2	4.94	-	NT		
			03147	01-06-1999	834	98	267	0.4	NT	NT	NT		
P	1957	M	0764	30-01-1996	1	NT	NT	NT	3.04	+	+		
			00701	09-09-1997	588	581	1769	0.3	2.70	+	-		
			02050	02-09-1998	947	524	2125	0.2	4.26	+	+	4.26	1.25
			02975	22-04-1999	1161	507	1814	0.3	4.26	NT	NT		

Q	1960	M	7202	26-07-1996	1	NT	NT	NT	4.54	+	+	NA	NA
			C0363	11-05-1998	691	NT	NT	NT	<1.90	-	NT		
			00917	23-12-1998	907	NT	NT	NT	NT	-	NT		
			C1470	02-06-1999	1016	559	641	0.9	NT	NT	NT		
R	1958	M	7194	25-07-1996	1	NT	NT	NT	3.34	+	+		
			00948	11-11-1997	493	163	567	0.3	NT	NT	NT		
			01732	18-05-1998	671	218	886	0.2	3.30	+	+	0.00	3.86
S	1967	M	7050	28-06-1996	1	NT	NT	NT	3.30	+	+		
			01100	20-11-1997	510	441	1373	0.3	4.18	+	+	4.27	3.80
			01746	26-05-1998	687	489	1467	0.3	3.99	+	+	0.00	0.00
			02546	14-12-1998	759	455	1317	0.3	4.11	NT	NT		
T	1969	M	7233	01-08-1996	1	NT	NT	NT	3.11	+	+	NA	NA
			01225	16-12-1997	501	259	678	0.4	3.85	NT	NT		
			01748	17-06-1998	684	238	475	0.5	3.84	-	-		
			02547	14-12-1998	864	NT	NT	NT	3.86	NT	NT		
U	1954	M	7043	27-06-1996	1	NT	NT	NT	NT	+	+		
			01338	26-01-1998	731	368	536	0.7	3.20	NT	NT		
			01842	08-07-1998	834	228	456	0.5	3.93	+	+	8.67	3.97
V	1959	M	7043	25-06-1996	1	NT	NT	NT	2.20	-	NT	NA	NA
			01362	02-02-1998	587	373	594	0.6	3.46	NT	NT		
			02191	07-10-1998	834	349	497	0.7	2.90	+	+		

Figure 8a

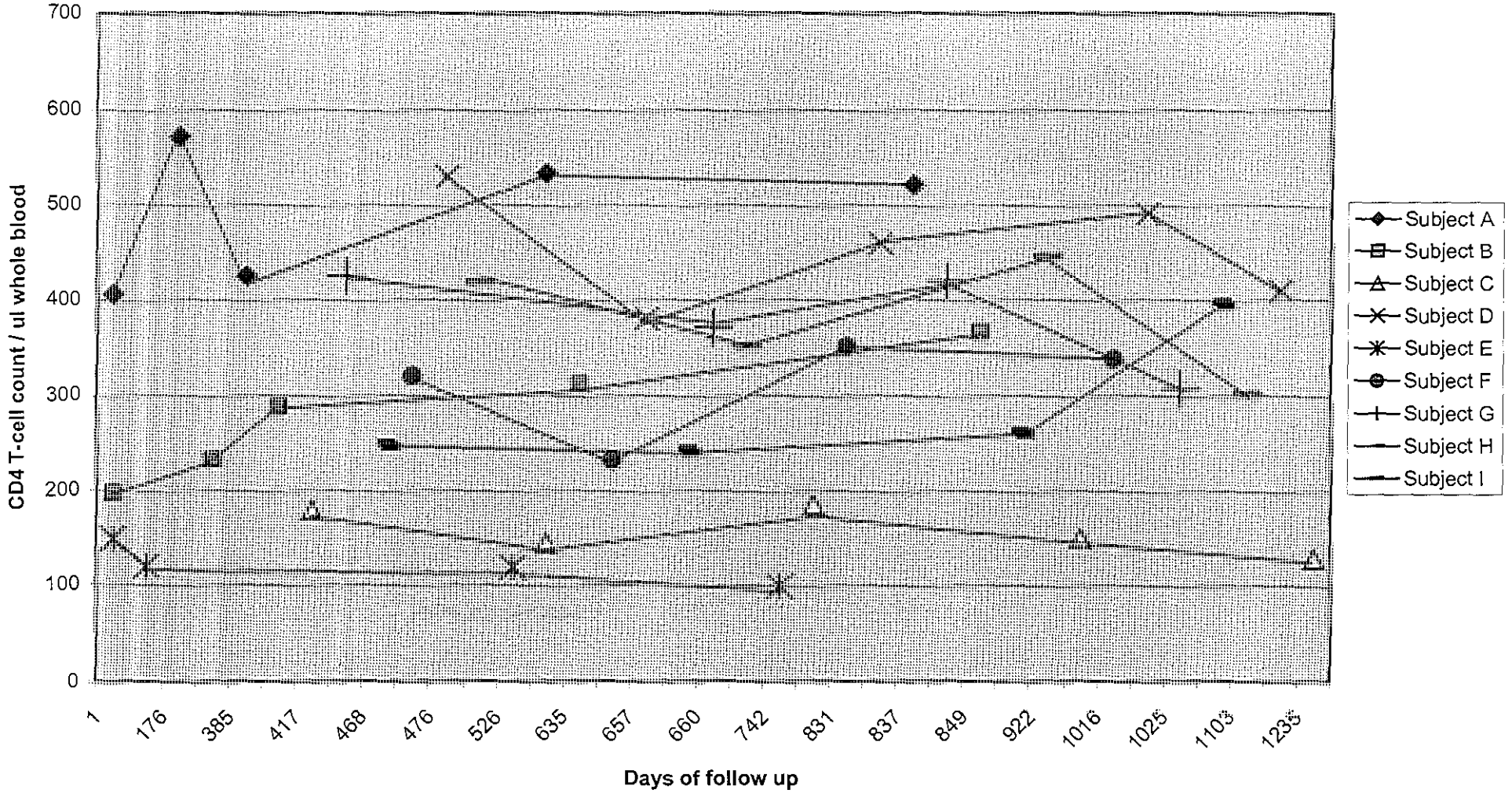
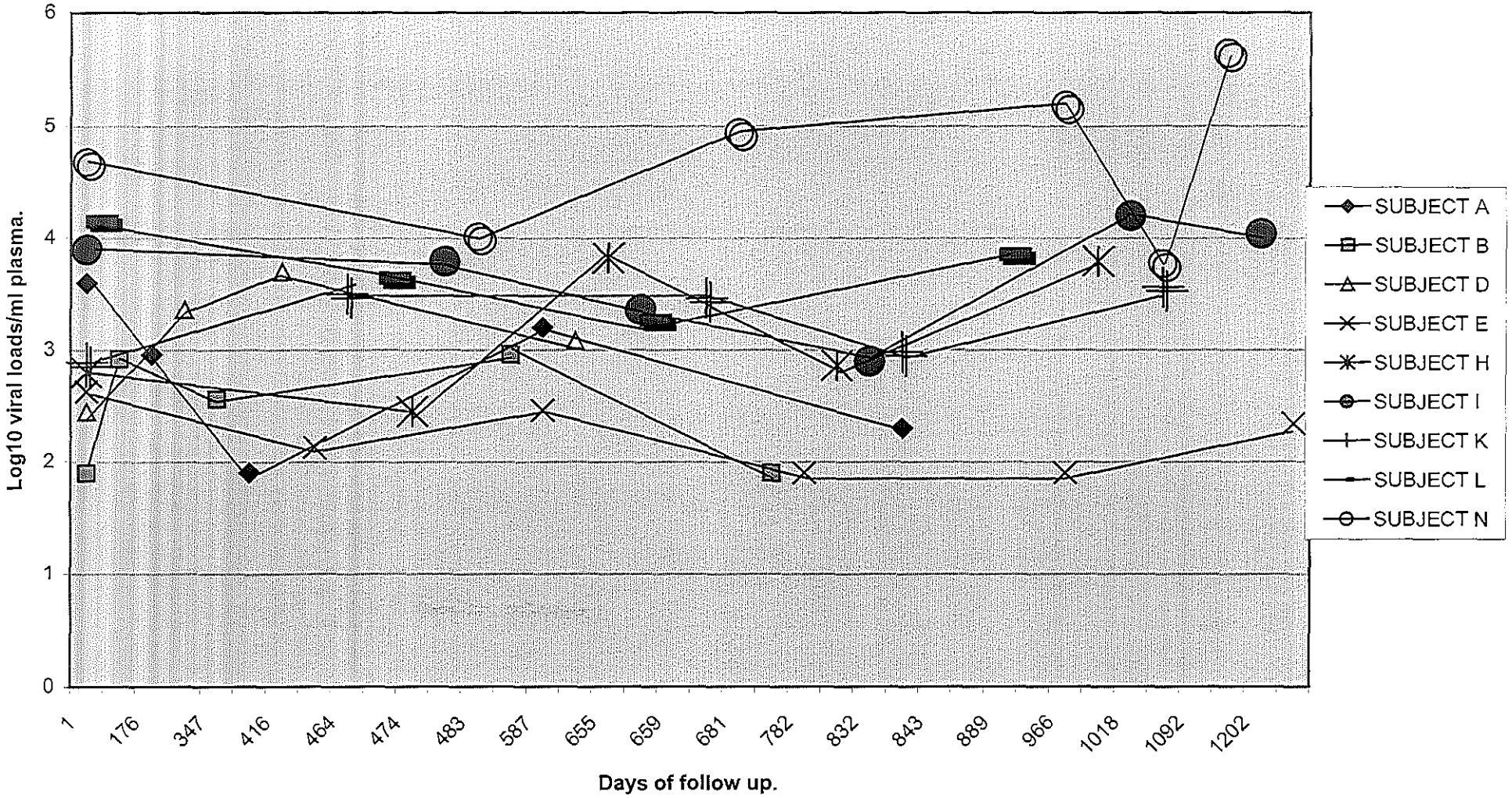

700
600
500
400
300
200
100
0
CD4 T-cell count / ul whole blood
1
176
385
417
468
476
526
635
657
660
742
831
837
849
922
1016
1025
1103
1238
Days of follow up
Subject A
Subject B
Subject C
Subject D
Subject E
Subject F
Subject G
Subject H
Subject I

Figure 8b


Log10 viral loads/ml plasma.
0
1
2
3
4
5
6
1
176
347
416
464
474
483
587
655
659
681
782
832
843
889
966
1018
1092
1202
Days of follow up.
SUBJECT A
SUBJECT B
SUBJECT D
SUBJECT E
SUBJECT H
SUBJECT I
SUBJECT K
SUBJECT L
SUBJECT N

ANNEXES: ENARP LAB PROTOCOLS

PROTOCOL 044: BLOOD COLLECTION BY VACUTAINER SYSTEM, ENARP.

update: 17-2-1997

Introduction:

The vacutainer system is used to collect blood specimens for diagnostic analysis in the laboratory. The system is composed of three basic elements:

. An evacuated glass tube containing a pre-measured vacuum to provide a controlled specified draw.
. A holder which is used to secure the needle during insertion into the tube stopper and subsequent venapuncture.
. A sterile single use blood collection needle suitable for drawing either single or multiple samples.

Materials needed:

- 2x purple EDTA 10 ml vacutainer tubes, holder and needle
- marker
- 70% alcohol swab
- dry sterile swab
- tourniquet
- protective dressing

1. Open needle package, but don't remove needle shield.
2. Insert needle into holder. Cleanse the stopper of the tube with 70% alcohol and allow to dry.
3. Insert tube into holder, push tube stopper onto needle until leading edge of stopper meets guide line of holder.
4. Select site for venapuncture.
5. Apply tourniquet. Prepare venapuncture site with an appropriate antiseptic (70% alcohol). Do not palpate venapuncture area after cleausing.
6. Push tube to end of holder puncture diaphragm of stopper.
7. Remove tourniquet as soon as blood appears in tube. Do not allow contents of tube to contact the stopper or the end of the needle during the procedure.
8. If no blood flows into the tube or ceases to flow before an adequate sample is collected, the following steps are suggested:
   a. confirm correct position of the needle in the vein.
   b. if a multiple sample needle is being used, remove the tube and place a new tube into the holder.
   c. remove needle and discard, repeat the procedure.
9. Invert the tube 5-6 times, gently. Do not shake vigorously: mixing can cause haemolysis.
10. As soon as blood stops flowing into the tube, remove the needle from the vein. Apply pressure to the puncture site with dry sterile swab. Keep swab in place by elevating arm.

11. Label the tube with the Subject ID, cross check Subject ID with cohort forms.
12. StorE tube in rack at RT, but not in direct sunlight, until transportation.

PROTOCOL 017: PLASMA ISOLATION AND FREEZING, ENARP

update: 2-12-1996

1. Label plasma freezing vials: LAB ID + "A","B","C","D","E".
2. Label per subject one sterile 15 ml tube with the corresponding LAB ID.
3. Centrifuge EDTA vacutainer tubes at 300 g (1640 rpm) for 10 min, with brake. Mark upper level with lab marker.
4. Transfer with a sterile Pasteur pipet (pastette) plasma to the labeled 15 ml tube.
5. Add EHN medium up to mark, close vacutainer tubes and mix carefully.
6. Centrifuge the 15 ml plasma tube 10 min 3000g at RT.
7. Transfer with pastette CELL-FREE plasma to the freezing vials. Avoid contact with the cells/debris.
8. Store freezing vials immediately at -80°C according to instructions. Keep vial "B" in -20°C for serological assays. Write stored plasma volumes and positions on worksheet.

PROTOCOL 011: RNA ISOLATION FROM PLASMA, ENARP.

update: 17-02-1999

N.B.1 Work RNAse free: changes gloves after every step in the protocol, prepare buffers from newly opened reagents, sterilize all bottles, tubes etc. by autoclaving, etc.
N.B.2 All reagents/stocks/buffers below should be sterilized by autoclaving in glass bottles.
N.B.3 All aliquots are made assuming an experiment on 10 samples. Aliquots should be used ONCE only. Any left-overs should be discarded.

stocks/reagents needed:

10 N NaOH
. add 400 g NaOH pellets to a 1 liter measure cylinder
. add dH2O up to 1 liter. Be carefull, this solution becomes very hot, so use proper plastic containers. Never use glass measure cylinders.
. 100 ml in waste flask of aspiration device.

1 M Tris-Cl pH=8.0
. add 121.1 g Tris to 800 ml dH2O in a 1 liter measure cylinder
. add 42 ml 37% HCl until pH = 8.0
. make 1 liter with dH2O

L1 buffer
. add 12.1 g Tris to 900 ml dH2O in a 1 liter measure cylinder
. add 8.1 ml of 37% HCl
. make 1 l with dH2O
. check pH with pH meter (should be 6.4)

L2 buffer (wash)
. add 120 g guanidium isothiocyanate (GuSCN) to 100 ml L1 buffer at 60°C in a 250 ml measure cylinder, mix
. aliquot in protions of 25 ml in sterile 50 ml blue cap tubes
. store in the dark at RT for max 6 months

0.5 M EDTA pH=8.0
. add 93.05 g EDTA to 400 ml dH2O in 500 ml measure cylinder
. stirr vigorously on magnetic stirrer (EDTA will not completely dissolve)
. adjust pH to 8.0 with 10-11 g of NaOH pellets
. all EDTA should be dissolved and pH should be 8.0
. make 0.5 l with dH2O

L6 buffer (lysis)
. add 120 g GuSCN to 100 ml L1 buffer at 60°C in 250 ml measure cylinder, mix
. add 8.8 ml of 0.5 M EDTA pH=8.0
. pipet approx. 3 ml Triton X-100 into a 15 ml blue cap tube, heat at 60°C for 10 min
. pipet with disposable 5 ml pipet 2.4 ml (=2.6 g) Triton X-100 into L6

solution
. homogenize and make 10 ml aliquots in sterile 15 ml blue cap tubes
. store in the dark at RT for max 6 months

dH2O
. sterilize dH2O in 500 ml bottles by autoclaving
. aliquot in portions of 1 ml in Eppendorf tubes and store at -20°C

TE^{-4} buffer
. add 5 ml 1M Tris-Cl stock and 100 ul 0.5 M EDTA stock to 495 ml dH2O
. aliquot in portions of 600 ul per sterile Eppendorf tube and store at -20°C.

silica
. prepare according to PROTOCOL.059.

[PROTOCOL 059: PREPARATION OF SIZE-FRACTIONATED SILICA-PARTICLES, ENARP.

update: 26-10-1998

material: Silicon dioxide (SiO2), Sigma cat. No. S5631; particle size distribution: 0.5-10 microns; 80% between 1-5 microns.

1. Add 60 g silica to a total volume of 500 ml dH2O in a glass 500 ml measure cylinder (height of aqueous column = 27.5 cm; width = 5 cm).
2. Sediment for 24 hours at RT.
3. Dispose of 430 ml supernatant by suction, until 70 ml is left.
4. Add dH2O up to 500 ml; resuspend by vigorous shaking.
5. Sediment for 5 hours at RT.
6. Dispose of 440 ml supernatant by suction, until 60 ml is left.
7. Check pH: should be 2.0.
8. If necessary: adjust pH with concentrated (37%) HCl using a plastic Pastette. Add HCl drop by drop.
9. Leave at RT for 30 min.
10. Using 25 ml disposable pipette, remove clear supernatant and pipet temporarily into 50 ml blue cap tube.
11. Mix the silica solution and pour into a 100 ml bottle. Rinse the 500 ml measure cylinder with the clear supernatant and add to the 100 ml bottle.
12. Autoclave the silica.
13. Aliquot in portions of 500 ul in Eppendorf tubes and store at -20C°.
14. Write date of silica preparation on top of box containing aliquots. Discard after 6 months.

]

3M NaAc
. add 102 g Na-acetate(.3H2O) to measure cylinder
. add dH2O up to 200 ml

. mix in measure cylinder until dissolved
. adjust pH to 5.2 by adding glacial acetic acid
. add dH2O to 250 ml

ISOLATION LAB

1. Start with PROTOCOL.083 and prepare the laminar flow of the Isolation Lab according to instructions.
2. Put Eppendorf centrifuge time on 1 min (in order to put the display in "sec mode").
3. Vortex Eppendorf tubes and spin down for 3 sec (display: 55 sec).
4. Vortex silica well to get a homogeneous solution (approx. 1 min).
5. Dispense 40 ul silica into tubes 1-5. Do not pipet up and down.
6. Vortex silica again until homogeneous. Dispense 40 ul into tubes 6-10.
7. Vortex sample tubes to get a homogeneous solution.
8. Incubate at RT for 10 min. Every 2 min: mix silica by inverting the tubes, until silica is homogeneously dissolved.
9. During step 8: Label 10 Eppendorf tubes (1.5 ml safe lock) with stickers:
   - 25 ul RNA in 50ul Eth.Abs and 3ul 3M NaAc
   - date isolation
   - LAB ID
10. Put silica sample tubes in centrifuge with opening facing outside.
11. Centrifuge at 12.000 g for 30 sec (display: 25 sec).
12. Change gloves.
13. Aspirate supernatant. N.B. Also check inside lid! **Open 1 tube at the time under all circumstances!** Take new yellow tip for every tube.
14. Add 1 ml Washing buffer (L2) with blue tip pipet (1000 ul). When tip doesn't touch tube, use the same tip for all tubes. Otherwise change tips.
15. Vortex well, untill silica pellet is visually resuspended.
16. Change gloves.
17. Repeat step 10.-16. with L2 buffer.
18. Repeat step 10.-16. 2x with 70% ethanol.
19. Repeat step 10.-16. 1x with acetone.
20. Let 10 ml of dH2O run through the rubber hose of the vacuum system: aceton eats rubber, so remove properly !
21. Dry pellets in heatblock at 56°C: **OPEN TUBES** for 10 min. **NOT LONGER!** (alternatively: dry at RT for 20 min).
22. Start PROTOCOL.082 in Isolation Lab, steps 1.-5.
23. After drying, close tubes. Check if silica pellets are completely dry by tapping firmly with fingers. When silica is dry it's changed into white powder.
24. Add 50 ul TE^{-4} and vortex well. Invert the tubes for proper mixing of silica with TE^{-4}.
25. Incubate 15 min 56°C in thermomixer while gently shaking (position 7).
26. In the meantime prepare RT master mix, according to PROTOCOL.082, steps 6.-9.
27. Put 50 ul of 100% ethanol and 3ul 3M NaAc in the labeled safe lock Eppendorf tubes (step 9. of this protocol).
28. Centrifuge the TE^{-4} tubes containing RNA isolate for 2 min at 12.000 g, opening facing outside.
29. Transfer of each RNA isolate 10 ul supernatant to the corresponding RT-mix PCR Eppendorf tubes and start RT reaction according to PROTOCOL.082, steps 10.-12.
30. Transfer of each RNA isolate 25 ul supernatant to the corresponding safe

lock Eppendorf tubes. (which contain 50 ul 100% ethanol and 3 ul 3M NaAc pH5.2)
31. Rinse tip by pipetting up and down 3 times. Vortex briefly.
32. Spin down all tubes for 3 sec 12.000 g.
33. Spray safe lock Eppendorf tubes containing RNA/ethanol/NaAc with Ethanol 70% and put in clean rack.
34. Store RNA/ethanol/NaAc mixes at -80°C **ASAP!** Freezer 1, Holder AA, boxes 1-11.
35. Administrate the positions on your DNA Sequence Flow Chart
36. Immediatly after storing RNA in -80°C, go back to Isolation Lab and **CLEAN THE ISOLATION FLOW!** according to PROTOCOL.082 step 13.
37. Continue PROTOCOL.082 at step 14.
38. Enter positions of RNA/ethanol mixes in FAME.

39. For the isolation of the RNA from the -80°C.
40. Take the RNA/ethanol/NaAc mixer from the -80°C and spin down for 30 min at 12.000 rpm.
41. Aspirate supernatant and add 500ul cold 70% ethanol, and spin down for 10 min.
42. Aspirate supernatant and dry pellet on the air for 20 min.
43. Solve pellet in 10ul TE^{-4} and start RT reaction according to PROTOCOL.082, steps 10.-12.

PROTOCOL 082: RT-PCR AIDED V3 SEQUENCING SAMPLE PREPARATION, ENARP.

update: 17-12-1998

Notes: The cDNA synthesis kit of Promega is used. Estimated RNA yield from 100 ul serum: <100 ng

stock solutions:

10 mM dNTP mix: 100 ul 100 mM dATP (Promega U 1240)
100 ul 100 mM dCTP (Promega U 1240)
100 ul 100 mM dGTP (Promega U 1240)
100 ul 100 mM dTTP (Promega U 1240)
600 ul dH2O

. aliquot in 25 ul quantities in 40 Eppendorf tubes

ED33 primer (10 pmol/ul): e.g. 100 nmol in 10 ml TE^{-4}
Mw = 6,658 118.19 nmol in 11.819 ml TE^{-4}
1000 ug/6,658 x 1000= 150 nmol

. aliquot in 50-100 ul quantities in 120-250 Eppendorf PCR tubes (0.5 ml) and store at -20°C freezer in airlock Isolation Lab.

ED31 primer (10 pmol/ul)
Mw = 8,786

5'V3 M13 F primer (10 pmol/ul) **17/12/98: 99.56 nmol in 9.956 ml**
Mw = 10,849

3'V3 M13 R primer (10 pmol/ul) **17/12/98: 98.32 nmol in 9.832 ml**
Mw = 13,343

I. RT (Reverse Transcription)

OFFICE

1. Fill in Worksheet Isolation / Reverse Transcription. Collect ice from kitchen.

ISOLATION LAB

2. Power on water bath at 42°C.
3. During step 21. of PROTOCOL.011, thaw and homogenise the following items:

- MgCl2, 25 mM
- RT buffer, 10x
- 3' primer ED33, 10 pmol/ul

4. Put enzymes on ice:
   - AMV-RT, 10 U/ul
   - RNAsin, 40 U/ul

5. Label 10 Eppendorf PCR tubes (0.5 ml): 1-10.
6. During step 25. of PROTOCOL.011, prepare RT-master mix according to Worksheet Isolation/Reverse Transcription.
7. Put enzymes, RT buffer and 3'primer ED33 in -20°C ASAP! Leave $MgCl_2$ and dNTP's on ice.
8. Homogenise RT-Master mix by inverting tube. **DO NOT VORTEX!**
9. Aliquot RT-Mastermix in PCR Eppendorf tubes, 10 ul each and put on ice.
10. Transfer of each RNA isolate 10 ul supernatant into the corresponding RT-mix PCR Eppendorf tubes.
11. Rinse tip by 3x pipetting up and down. Spin for 3 sec **DO NOT VORTEX!**
12. Start RT incubation at 42°C for 45 min (set timer).
13. **CLEAN FLOW:** - discard all unnecessary items
- (silica/L2/ethanol70%/acetone/low TE buffer, etc.)
- discard plastic bag out of sharp container
- spray centrifuge (also **inside**!) with ethanol 70%, clean and dry with tissue
- spray all items left in flow with ethanol 70%, clean and dry with tissue
14. Set heat block (inside flow) at 95°C.
15. Prepare 1st PCR master mix according to 1st PCR Worksheet and put on ice (inside flow).
16. **CLEAN MIX SIDE OF LAB!** (incl. pipets etc.) using ethanol 70% and bleach 10%.
17. After completing RT reaction (45 min): centrifuge RT tubes for 3 sec and inactivate by putting tubes at 95°C for 5 min in heat block.
18. Centrifuge RT tubes (3 sec) and immediately put on ice for 5 min.

II. First PCR

19. Centrifuge RT tubes (3 sec).
20. Mix 1st PCR master mix by inverting tube.
21. Add 80 ul of 1st PCR to each RT tube (don't pipet up and down). Total volume is now: 100 ul (20 ul RT + 80 ul 1st PCR).
22. Overlay 1st PCR tubes with 2 drops of mineral oil (Pastette).
23. Centrifuge 1st PCR tubes (3 sec), transfer to a clean rack, outside flow.
24. Spray flow briefly, incl. pipets/centrifuge/vortex etc. using ethanol 70%.
25. Transfer rack + 1st PCR tubes to Molecular Biology Lab 2.

MOLECULAR BIOLOGY LAB 2

26. Perform 1st PCR. This will take approximately 2.5 hours.

Amplification conditions: (file 94/95/98/99)

94°C	1 min		
55°C	1 min	5x	(file: 94)
72°C	1 min		
94°C	15 sec		
55°C	45 sec	30x	(file: 95)
72°C	1 min		
72°C	5 min	1x	(file: 98)
4°C	soak file		(file: 99)

27. Take 1st PCR tubes out of machine after program has finished. Although PCR can be run overnight, don't leave machine soaking unnecessarily. Put tubes in fridge ASAP.
28. Load 1st PCR products on agarose gel the same day. Use PROTOCOL.054 for preparing agarose gel.
29. Label 1st PCR product tubes using stickers with: Lab ID, date, 1st primers and store at 4°C.
30. Only when 2nd PCR (see below) is OK, store 1st PCR products at the -20°C freezer in Molecular Biology Lab 2. Write tube positions on DNA Sequence Flow Chart and store in box file.

III. Agarose Gel Electrophoresis

Molecular Biology Lab 2

31. Prepare agarose gel according to PROTOCOL.054 description.
32. Vortex 1st PCR tubes well.
33. Centrifuge 1st PCR tubes for 15 sec.
34. Pipet (x times) 10 ul of dH2O in 96 well plate; x = number of samples to be analysed. Take plate to kitchen.

Kitchen

35. In UV cabinet: take 10 ul 1st PCR samples, using filter tips and pipet in the corresponding 96-well micro titer plate wells.
36. Rinse tip by 3x pipetting up and down. Take plate to Molecular Biology Lab 2.

Molecular Biology Lab 2

37. Add 5 ul loading dye (6x). Don't pipet up and down.
38. Using P20 pipet, load 20 ul per well on agarose gel, according to Worksheet 1st PCR. Mix dH2O/1stPCR/Loading dye by pipetting up and down at least 3 times, before loading.
39. Prepare molecular weight marker: make sure you load 1.5 ug of marker. Marker stock: 1 ug/ul. Dilute: 6 ul marker (stock) + 14 ul dH2O. Prepare to load: 15 ul dH2O + 5 ul diluted marker + 5 ul loading dye. Load 20 ul of this mix per slot.
40. Run at 100 mA (limiting factor) for 1 hour. Voltage = approx. 150 V
41. Store remaining 1st PCR product at 4°C.

Dark room

42. Make Polaroid picture and check Mw of 1st PCR product. If visible: should be 564 bp.

IV Second PCR

Isolation Lab

43. Prepare 2nd PCR master mix according to 2nd PCR Worksheet.
44. Aliquot into 0.5 ml PCR tubes: 90 ul each, overlay with 2 drops of mineral oil.

45. Leave tubes in fridge Isolation Lab.

Molecular Biology Lab 2

46. Centrifuge 1st PCR products for 3 sec.

Kitchen

47. Take 2nd PCR-mix tubes and 1st PCR products to hood in kitchen.
48. Pipet 10 ul 1st PCR product through the mineral oil into the corresponding 2nd PCR-mixes. N.B. If the 1st PCR product was clearly positive on agarose gel, use 2-5 ul of it and add 8-5 ul dH20 to make total input volume 10 ul.

Molecular Biology Lab 2

49. Centrifuge 2nd PCR tubes for 3 sec
50. Perform 2nd PCR: this will take approximately 1.5 hours.

Amplification conditions: (file 96/97/98/99)

94°C	1 min		
55°C	1 min	5x	(file: 96)
72°C	1 min		
94°C	15 sec		
55°C	45 sec	25x	(file: 97)
72°C	1 min		
72°C	5 min	1x	(file: 98)
4°C	soak file		(file: 99)

51. Take 2nd PCR tubes out of machine after program has finished. Although PCR can be run overnight, don't leave machine soaking unnecessarily. Put tubes in fridge ASAP.
52. Load 2nd PCR products on agarose gel the same day. Use PROTOCOL.054 for preparing agarose gel.
53. Label 2nd PCR product tubes using stickers with: Lab ID, date, 2nd primers and store at 4°C.

V Agarose Gel Electrophoresis

Molecular Biology Lab 2

54. Prepare agarose gel according to PROTOCOL.054 description.
55. Vortex 2nd PCR tubes well.
56. Centrifuge 2nd PCR tubes for 15 sec.
57. Pipet (x times) 15 ul of dH2O in 96 well plate; x = number of samples to be analysed. Take plate to kitchen.

Kitchen

58. In UV cabinet: take 5 ul 2nd PCR samples, using filter tips and pipet in the corresponding 96-well micro titer plate wells.
59. Rinse tip by 3x pipetting up and down. Take plate to Molecular Biology Lab 2.

Molecular Biology Lab 2

60. Add 5 ul loading dye (6x). Don't pipet up and down.
61. Using P20 pipet, load 20 ul per well on agarose gel, according to Worksheet 2nd PCR. Mix dH2O/1stPCR/Loading dye by pipetting up and down at least 3 times, before loading.
62. Prepare molecular weight marker: make sure you load 1.5 ug of marker. Marker stock: 1 ug/ul. Dilute: 6 ul marker (stock) + 14 ul dH2O. Prepare to load: 15 ul dH2O + 5 ul diluted marker + 5 ul loading dye. Load 20 ul of this mix per slot.
63. Run at 100 mA (limiting factor) for 1 hour. Voltage = approx. 150 V
64. Store remaining 2nd PCR product at 4°C.

Dark room

65. Make Polaroid picture and check Mw of 2nd PCR product. If visible: should be 350 bp.
66. Interpret results according to Worksheet 2nd PCR: negative, weak positive, positive and strong positive.
67. If cycle sequence is performed within 3 days, temporarily store 2nd PCR products in fridge, 4°C (Mol Biol Lab 2).
68. If cycle sequence is performed later: store 2nd PCR products at -20°C, freezer (Mol Biol Lab 2) and write tube positions on DNA Sequence Flow Chart.
69. Perform cycle sequence with optimal DNA input according to results interpreted on Worksheet 2nd PCR.

- switch on light, electricity, ventilation and put glass door in correct position
  - spray inside of flow with Ethanol 70%
  - put plastic bag in yellow sharps container
  - paper (in case of any spilling and for cleaning afterwards)
  - P200 pipet and tips (FILTER!)

2. Collect the plasma samples to be tested from the -80°C freezers, according to the DNA Sequence Flow Chart.
3. Thaw plasma in flow Serology Lab (in rack of Serology lab). This takes approx. 10 min, in the meanwhile do the following:

Isolation Lab

4. Take 2 racks out of bleach 10%, rinse with tap water, spray with Ethanol 70%.

Mix side (left side of lab, clean side)

5. Power on water bath at 42°C.
6. Put a new white absorbing diper on bench.
7. Put plastic bag in sharp container.
8. Put RNAse free pipets (P10, P20, P200, P1000) on diper and spray with Ethanol 70%.
9. Take 10 Eppendorf tubes (safe lock, 1.5 ml) and close them ASAP when on the bench.
10. Number tubes from 1 - 10.
11. Pipet 900 ul Lysis Buffer (L6) into each tube.
12. Add 100 ul dH2O to tubes 5 and 10 (blanco's), rinse tip by pipetting up and down 3 times.

Isolation side (right side of lab, dirty side)

Prepare laminar flow
- switch on light and electricity, **DO NOT SWITCH ON** ventilation!
- spray inside with 70 % Ethanol
- put plastic bag in yellow sharps container
- get paper (in case of any spilling and for cleaning aferwards)
- get the appropriate DIRTY pipets (P10, P20, P200, P1000) and tips (FILTER!)
- get all items needed for isolation:

- Silica, 500 ul (from -20°C)
- Washing buffer, (L2) 25 ml (room temp)
- Ethanol 70%, 25 ml (room temp)
- Acetone, 15 ml (room temp)
- Low TE Buffer (TE^{-4}) 600 ul (-20°C)
- Ethanol Absolute, 1 ml (-20°C)

- Check whether waste flask is filled with 100 ml 10 N NaOH (to neutralize GuSCN/possible basic reagents -> gas).

13. Take rack with the 10 tubes to **Serology lab** and place inside flow.

Serology lab

14. Put a spare set of clean gloves in pocket of your white coat.

15. Vortex 5 sec plasma samples, make sure they are completely thawed.
16. Open 1 plasma sample at a time, put the lid upside down in the flow.
17. Pipet 100 ul plasma into appropriate Eppendorf tube, according to DNA Sequence Flow Chart.
18. Rinse tip by pippetting up and down 3 times. Be careful: Lysis Buffer (L6) intends to foam easily! Be careful with placing the lid of the plasma tube back on. If in doubt of touching plasma with gloves:

- spray the tube with Ethanol 70% immediately
- change gloves immediately (from your pocket)
- never put notes inside flow, place on a chair next to you outside

19. When all plasma samples are added to Lysis Buffer (L6), vortex 5 sec.
20. Spray all plasma tubes, Eppendorf tubes, racks, pipets, outside boxes of filter tips and the vortex with Ethanol 70%.
21. Put all sprayed items on edge of flow, just outside the glass window.
22. Discard the plastic bag from the sharps container.
23. Spray sharps container and bench of flow with Ethanol 70%.
24. Discard your gloves !!!
25. All sprayed items are now safe to touch with bare hands, but in order to keep everything as RNAse free as possible, put one clean glove on and touch all items with that glove only.
26. Take rack with L6 lysed samples to **Isolation room**, spray again with Ethanol 70% before entering, just to be sure and place in flow.
27. Take rack with plasma samples and place them back in the -80°C freezers in appropriate positions, according to DNA Sequence Flow Chart. Do this **ASAP**!
28. Put rack in Ethanol 70% container in Serology lab.
29. Turn off light and electricity and ventilation of flow in Serology lab, pull down the glass untill the flow is closed.

PROTOCOL 054: AGAROSE GEL ELECTROPHORESIS, ENARP.

update: 28-12-1998

reagents:

0.5 M EDTA pH 8.0:
. dissolve 93.05 g EDTA in 400 ml demi H2O
. stir vigorously on magnetic stirrer
. Adjust pH to 8.0 with +/- 10 g of NaOH pellets
. make 0.5 l with demi H2O
. sterilise by autoclaving

50x TAE buffer:
. dissolve 242 g Tris in 800 ml dH2O
. add 57.1 ml glacial acetic acid
. add 100 ml 0.5 M EDTA (pH=8.0)
. add dH2O to 1000 ml

6x loading buffer:
. 1.5 g Ficoll (type 400, Pharmacia)
. 1 ml 0.5M EDTA, pH=8.0
. 0.025 g bromophenol blue
. 0.025 g xylene cyanol
. add dH2O to 10 ml

20.000x EtBr
. add 0.1 g of EtBr to 10 ml of H2O
. mix for 2 hours until dye has completely dissolved
. wrap container in aluminium foil to protect from light
. store at RT

1. Add required g agarose to Erlenmeyer of correct size (see table).
2. Mix 3 ml 50x TAE with 147 ml dH2O in correct size measure cylinder.

1 % agarosegel			
	100 ml	150 ml	200 ml
g agarose	1	1,5	2
ml 50 x TAE	2	3	4
ml H_2O	98	147	196
µl EtBr	5	7,5	10
2 % agarosegel			
	100 ml	150 ml	200 ml
g agarose	2	3	4
ml 50 x E	2	3	4
ml H_2O	98	147	196
µl EtBr	5	7.5	10

3. Add 75 ml mixed 1x TAE to the agarose powder, mix again.
4. Heat the solution 2.5 min at 360 Watt in microwave, check whether all agarose has dissolved. Cool down by adding the other 75 ml 1xTAE.
5. Prepare tray using tape and put the combs in correct positions.
6. Prepare 1.5 l of 1x TAE: 30 ml 50x TAE and dH_2O up to 1.5 l.
7. Add 75 ul of EtBr to solution and mixed thoroughly. Use always gloves, when handling EtBr (carcinogenic !!!).
8. Pour gelmix in the tray and let cool for 30 min.
9. Pour some 1x TAE buffer on top of the gel and let stand 5 min.
10. Carefully remove combs and add 1x TAE buffer until gel is submerged.
11. Add samples to slots and run the gel on max. 100 mA for 30-60 min.

PROTOCOL 084: CYCLE SEQUENCE REACTIONS AND GEL RUNNING, ENARP.

update: 14-4-1999

in general:
- wear gloves
- use autoclaved pipet tips and eppendorf tubes
- use the Thermo Sequenase fluorescent labelled primer cycle sequencing kit. Amersham RPN 2536.

Perform, as much as possible, the following steps on ice:

1. Make a premix of primer in a 0.5 ml Eppendorf tube:
   . for 10 samples: (10x 2.1 ul) = 21 ul primer
   (10x 14.7 ul) = 147 ul milli Q H2O
   . for 5 samples: (5x 2.1 uL) = 10.5 ul primer
   (5x 14.7 ul) = 73.5 ul milli Q H2O
2. Mix premix well before use, by pipetting up and down.
3. Label Eppendorf tubes (e.g. 1-10) and add 16 ul premix to each tube.
4. Dependent on 2nd PCR result as analysed by agarose gel electrophoresis, add 5 ul sample to the appropriate tubes. So: either 5 ul undiluted, or 5 ul diluted 1/2, 1/10, 1/20. Optimal input DNA: approximately 50-100 ng.
5. Vortex 5 sec, centrifuge 3 sec.
6. This mix is called: Y.
7. Take 4 Eppendorf tubes (0.5 ml) for every sample and mark them with reaction number + A, C, G, T, respectively (e.g. 3A, 3C, 3G, 3T).
8. Pipet 2 ul A-mix from the kit in all A-marked tubes.
9. Do the same for C-mix, G-mix and T-mixes.
10. Turn the tubes, so that the previously added 2 ul ACGT-mixes are on the other side of the tube.
11. Add, without touching the 2 ul A-, C-, G-, or T-mixes, 5 ul of Y-mix (see step 6.) in the four ACGT-marked tubes of the correct reaction number
12. Add 1 drop of mineral oil (Sigma).
13. Vortex 5 sec, centrifuge 3 sec.
14. Run PCR program: 30" 95°C; 30" 55°C; for 25 cycles. This will take approximately 1 hour. File 28 on both PCR machines, but please always check.
15. Add to each sample 4 ul loading dye (total volume is now 11 ul).
16. Centrifuge for 3 sec.
17. Denature for 5 min at 95°C (File 17 on both PCR machines).
18. Centrifuge for 3 sec.
19. Put samples on ice, samples are ready to use, load 8 ul to the sequence gel.
20. Load the 2 most outer wells of the ALF with dye only, to avoid smiling effect.
21. Run gel at: 1200 V, 40 mA, 36 W, 40°C, laser power 5 mW, running time 600 min. For further information regarding alternative gel running conditions: ALF DNA sequencer user manual, section 5 "operation".