MOLECULAR DIAGNOSIS AND TYPING OF HTLV-I IN KWAZULU/NATAL

by

Michelle Lucille Tarin

Submitted in partial fulfilment of the requirements

of the degree of

MASTER OF MEDICAL SCIENCE

in the

Department of Molecular Virology
Faculty of Medicine
University of Natal
Durban

January 1998

To Kevin and Deidré, with love.

Contents

Declaration

This study presents original work by the author and has not been submitted in any form to this or any other university. The research described in this thesis was carried out in the Dept of Molecular Virology, Faculty of Medicine, University of Natal, under the supervision of Dr DF York and the co-supervision of Professor AI Bhigjee. Where use is made of the work of others, it has been duly aknowledged in the text.

_____

Michelle Tarin

Acknowledgments

To my supervisors, Dr Denis York and professor Ahmed Bhigjee for their critique of this manuscript.

To the staff of the Dept of Virology for their encouragement and advice.

A special thanks to Dr York for his advice, time and patience.

To Lorna for her advice and support. You were always ready to help.

To professor Bhigjee for his kindness and encouragement.

To Wendy for her time, encouragement and friendship.

To Shyamala for her instruction on the A.L.F. Express automated sequencer.

To the staff of the Biomedical Resource Centre, for the use of the A.L.F. Express automated sequencer.

To Fanny and Karen for their assistance with the cycle sequencing for the ABI automated sequencer.

To the staff of the S.A. Sugar Research Institute for running the samples on their ABI automated sequencer.

To Albert from MMS, for his assistance with the photographs.

To Fr Allan Moss and Claudette of Christ the King Parish for the use of the computer and laser printer.

To my family, especially my grand-parents, for supporting me during my studies.

To Kevin, the calming influence in my life particularly when I was frustrated.

A special thanks for all the hard work during the final preparation of the thesis.

To God, who makes everything possible.

Abbreviations

A.L.F.: Automated Laser Fluorescent

ABI: Applied Biosystems Incorporated

Amino Acids:

A: Alanine

R: Arginine

N: Asparagine

D: Aspartic acid

B: Asparagine or aspartic acid

C: Cysteine

Q: Glutamine

E: Glutamic acid

Z: Glutamine or glutamic acid

G: Glycine

H: Histidine

I: Isoleucine

L: Leucine

K: Lysine

M: Methionine

F: Phenylalanine

P: Proline

S: Serine

T: Threonine

W: Tryptophan

Y: Tyrosine

V: Valine

AMP: ampicillin

ASAP: AmpliSensor Analysis Program

ATL: adult T-cell leukaemia

ATP: adenosine 5'-triphosphate

ATP: adenosine triphosphate

BCL: bootstrap confidence level

bp: base pair

CDC: Centre for Disease Control

CNS: central nervous system

CSF: cerebrospinal fluid

d: number of nucleotide substitutions per site

DNA: deoxyribonucleic acid

dNMP: deoxyribonucleoside monophosphate

dNTPs: deoxyribonucleoside triphosphates

EBF: electrochemiluminescent hybridisation assay

ELH: enzyme-linked colourimetric hybridisation assay

EDTA: ethylenediaminetetra-acetate

EIA: enzyme immunoabsorbant assay

env: envelope proteins

EtBr: ethidium bromide

gag: group specific antigen

gp21: transmembrane envelope protein

gp46: outer membrane envelope protein

HAM: HTLV-I-associated Myelopathy

HBV: hepatitis B virus

HCV: hepatitis C virus

HIV: human immunodeficiency virus

HTLV-I: human T-cell leukaemia/lymphotropic virus type I

HTLV-II: human T-cell leukaemia virus type II

ICP: Internal Control Program

IFA: immunofluorescence assay

IPTG: isopropylthio-B-D-galactoside

JC: Jukes Cantor

KZN: KwaZulu/Natal

L: tree length

LTR: long terminal repeat

$\mathbf{L_u}$: upperbound of tree length

MEGA: Molecular Evolutionary Genetics Analysis

mg/ml: milligram/s per millilitre

mg: milligram/s

$\mathbf{Mg^{2+}}$: magnesium chloride

MHC: major histocompatibility complex

ml: millilitre

mM: millimolar

MP: Maximum Parsimony

pol: reverse transcriptase/polymerase

R: repeat

RIPA: radioimmunoprecipitation assay

RNA: ribonucleic acid

rpm: revolutions per minute

STLV-I: simian T-cell leukaemia virus

SU: external surface membrane

Taq: *Thermus aquaticus*

TE: tris EDTA

TM: transmembrane protein

Tris: 2-amino-2-hydroxymethyl propane-1,3-diol

TSP: Tropical Spastic Paraparesis

ug/ml: microgram/s per millilitre

ug: microgram/s

ul: microlitre

uM: micromolar

UPGMA: Unweighted Pairgroup Method of Assortment

UVP: Ultra-Violet-Products

V: volts

W: watts

WB: Western blot/immunoblot

X-gal: 5-bromo-4-chloro-3-indoly-B-D-galactosidase

3': three primed

5': five primed

List of Figures

Abstract

Two areas of the HTLV-I genome were targeted for an in-house molecular diagnostic test, namely the *pol* and *env* regions. The *pol* primers proved the most sensitive (100%) and specific (100%). Amplification using the *env* primer pair was not reproducible, and was not pursued further. The AmpliSensor assay (Acugen Systems, Lowell, MA) was also tested. The assay was very specific, but not as sensitive as our in-house PCR.

To investigate the predominant HTLV-I subtype in the region, a 1535 bp *env* gene was isolated from peripheral blood obtained from five local HTLV-I seropositive patients. Four of the patients presented with HAM/TSP, and the fifth presented with a skin disease. Nucleotide sequencing of the amplified products revealed the local strains to be very conserved, differing by 0.1% to 0.9% among themselves. No apparent difference was noted for the two clinical manifestations. Phylogenetic analysis was performed using repesentative strains from around the world. The local strains clearly fell within the cosmopolitan subtype. The local strains were most closely related to the North American strains suggesting an unexpected link between the two countries.

Chapter 1

General Introduction and Literature Review

1.1 **Introduction**

The ability to maintain "long term" cultures of lymphocytes, and the subsequent description of adult T-cell leukaemia (Takatsuki *et.al.*, 1977), were both breakthroughs that laid the foundation for the isolation of the first human retrovirus from a patient with cutaneous T-cell lymphoma by Poiesz and co-workers (1980). The virus was called human T-cell leukaemia/lymphoma virus type I (HTLV-I). Patients suffering from tropical spastic paraparesis (TSP), a neurodegenerative disease, were also found to have antibodies to HTLV-I (Gessain *et.al.*, 1985). Independently, Osame and co-workers (1986) described a chronic spastic paraparesis which they called HTLV-I-associated myelopathy (HAM). In 1988, Roman and Osame concluded that HAM and TSP are essentially one and the same disease. This condition is now referred to as HAM/TSP.

1.2 **Virological Characteristics**

HTLV-I is a type C retrovirus of the subfamily *Oncovirinae* (Poiesz *et.al.*, 1980). Mature virions are 110-140 nm in diameter, characterised by a spherical, centrally located nucleoid enveloped by a glycoprotein membrane with short spikes. The core consists of a diploid RNA genome of high molecular weight (approximately 9 056 bp long) with structural features common to all retroviruses, namely the genes for group specific antigen (*gag*), reverse transcriptase (*pol*), and envelope proteins (*env*), and flanked by long terminal repeat (LTR) sequences on either end (Fig. 1.1).

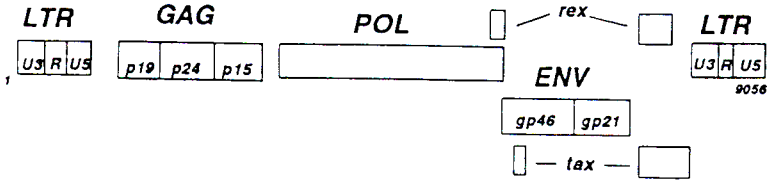


Fig. 1.1: A schematic map of the HTLV-I genome (Cann *et.al.*, 1990).

The HTLV-I envelope is a lipid bilayer in which the smaller transmembrane viral protein (gp21) and the larger outer membrane protein (gp46) are anchored (Haseltine *et.al.*, 1985). The long terminal repeat (LTR) comprises three distinct domains, U3, R and U5, in a 5' to 3' direction (Haseltine *et.al.*, 1985). The length of the repeat (R) sequence (228 bp) of the LTR region is longer than in other retroviruses. HTLV-I also contains unique regulatory genes encoding for transactivating proteins (*tax* and *rex*) at the 3' untranslated end of the provirus. The *tax* gene stimulates transcription of all genes from the 5' LTR sequences (Cann *et.al.*, 1985). The *rex* gene is a positive post-transcriptional regulator for *gag* and *env* expression (Gallo *et.al.*, 1988) and is also a negative regulator capable of inhibiting expression and replication of HTLV-I *in vivo* (Yoshida, 1989).

1.3 **Clinical Features of HTLV-I Infection**

The virus is the causative agent of two clinically distinct diseases: adult T-cell leukaemia (ATL) (Yoshida *et.al.*, 1982) and HAM/TSP (Gessain *et.al.*, 1985; Osame *et.al.*, 1986). The majority of HTLV-I infected individuals, however, are asymptomatic carriers with only 1-5% of infected persons developing either ATL of HAM/TSP in their lifetime (Hollsberg and Hafler, 1993). Onset of either disease is usually late in adult life, usually after 40 years of age (Takatsuki *et.al.*, 1977; Roman and Roman 1988). However, it is possible that this figure is biased towards sexually active individuals, rather than latency of the virus. The different disease manifestations are described below.

1.3.1 ATL

ATL is a malignant form of lymphoma, which can be classified into 5 types (Takatsuki *et.al.*, 1977): smoldering, chronic, crisis, acute, and lymphoma. Smoldering ATL is

characterised by skin lesions without visceral involvement, a minority of circulating leukaemic cells and a normal white-cell count (Dalgleish *et.al.*, 1988). Smoldering ATL may progress to either chronic or acute disease. In chronic ATL, there is visceral involvement with lymphadenopathy, hepatosplenomegaly and an elevated leukocyte count, indicating an increase in circulating leukaemia cells. Patients with either chronic or smoldering forms may go into crisis and progress to the acute phase. Patients with acute ATL have circulating lymphocytes that are mature, activated $CD4^+$ T-cells. Clinical presentations are lymphadenopathy, hepatosplenomegaly and skin lesions due to infiltrating leukaemia cells (Yoshida *et.al.*, 1984). Approximately 50% of patients with acute ATL have hypercalcemia (Yoshida *et.al.*, 1984).

ATL may also present as a diffuse non-Hodgkin's lymphoma with leukaemic cells absent in the peripheral blood (Takatsuki *et.al.*, 1977).

Involvement of the nervous system in ATL patients has been demonstrated in more than 60% of autopsy cases (Tara *et.al.*, 1989). The most common involvement is leukaemic cell infiltration of the brain and retina, accompanied occassionally by hemorrhage and necrosis. Hypercalcemia may cause an altered mental state in more than 25% of patients. Opportunistic cryptococcal infections are also common. Ten percent of fatal cases have leukaemic infiltration of the peripheral nerves.

1.3.2 HAM/TSP

The clinical, neurophysiological, and laboratory results in patients with HAM are identical to those of TSP patients (Osame *et.al.*, 1986). The main neurologic manifestations are chronic and slowly progressive spastic paraparesis with back pain,

macrophages, and show the added feature of a system degeneration (Moore *et.al.*, 1989).

1.5 **Pathogenesis**

The pathogenesis of HTLV-I infection is poorly understood, probably because of the long latent period between exposure and disease manifestations, coupled with the fact that patients are only autopsied after a considerable number of years of having the disease. This has particularly hindered the identification of sites of HTLV-I persistence in the central nervous system (CNS).

HTLV-I does not carry host-derived oncogenes nor has it been reported to activate cellular proto-oncogenes by insertion mutation, as found in other oncogenic animal viruses (Varmus 1988). HTLV-I pathogenesis in ATL is therefore unknown. A proposed model of HTLV-I tumour induction suggests that randomly integrated HTLV-I induces resting T-cells to proliferate followed by "secondary" signals that produce lymphocyte clones (Green and Chen, 1990). After HTLV-I begins to replicate, the *tax* gene product may induce proliferation, which may cause the infected cell to be more susceptible to further signals of transformation (Green and Chen, 1990; Dodon *et.al.*, 1989).

There are two major hypotheses regarding pathogenesis of HTLV-I in HAM/TSP. The first suggests that $CD4^+$ T-cells infected with HTLV-I may activate $CD8^+$ cytotoxic T-cells. These T-cells activate uninfected T-cells through T-cell interaction (Wucherpfennig *et.al.*, 1992). The activated $CD8^+$ T-cells migrate to the CNS and cause demyelination by recognising viral antigens from a proposed infected glial cell. This occurs by the induction of class I major histocompatibility complex (MHC) antigen expression. The

secretion of cytokines from cytotoxic T-cells that recognise viral protein products may also play a role in demyelination (Moore *et.al.*, 1989). The second hypothesis suggests that infected $CD4^+$ T-cells activate autoreactive $CD4^+$ T-cells. The $CD4^+$ T-cells migrate to the CNS and cause demyelination by recognising processed myelin antigens in the context of MHC class II antigens. Data demonstrating the reaction of circulating $CD8^+$ cytotoxic T-cells with HTLV-I protein products in patients with HAM, but not in a carrier or an ATL patient (Jacobson *et.al.*, 1990), supports the first hypothesis. Jacobson and colleagues (1992) later showed that $CD8^+$ cells predominated in the cerebro-spinal fluid (CSF) of HAM/TSP patients and that these cells were cytotoxic to cells infected with HTLV-I. A report by Kira *et.al.* (1994), who were able to identify HTLV-I sequences in brain and spinal cord tissue, is also in accord with the proposal that direct infection of glial cells by HTLV-I plays a role in CNS damage.

1.6 **Disease Transmission**

HTLV-I can be transmitted both horizontally and vertically. Major *et.al.* (1993) reported on a familial study where the father presented with ATL, the mother with TSP and the children were asymptomatic carriers, yet all had identical *tax* sequences. This shows the complete conservation of the gene, whether through horizontal or vertical transmission.

1.6.1 Horizontal Transmission

The virus is primarily spread by sexual contact (Hjelle *et.al.*, 1992; Wiktor *et.al.*, 1992). Sexual transmission can occur from male to female, male to male, or female to male, with the receptive partner being more susceptible than the insertive partner (Blattner, 1990). Blood transfusion of unscreened blood (Saxton *et.al.*, 1989), and sharing of

needles by intravenous drug abusers (Lee *et.al.*, 1989), are also modes of transmission. However, due to the cell associated nature of the virus, transfusion acquired infection occurs only with cellular blood products, not plasma (Okochi *et.al.*, 1984; Sandler *et.al.*, 1989).

1.6.2 Vertical Transmission

Transmission from mother to baby is predominantly via breast milk (Sugiyama *et.al.*, 1989). Twenty to 30% of babies fed exclusively on breast milk were infected after birth, while less than 3% that were fed on formula were infected (Ando *et.al,* 1987; Hino *et.al.*, 1989). A possible explanation for the cause of infection in the bottle fed babies is that infection occurred *in utero*. There is currently controversy over this point. Polymerase chain reaction (PCR) of cord blood samples of infected babies have been negative, and Katamine *et.al.* (1994) have shown that infected cells in the babies' circulation have lost their infectivity, and therefore do not infect neighbouring cells. However, conflicting evidence exists where, for example, an infected baby had shown signs of infection from the fetal stage (Bhigjee *et.al.*, 1995). Fetal movements and walking were delayed, although the baby's mental state was normal. As the mother only became symptomatic during pregnancy, the pregnancy-related immunosuppression could have resulted in a decrease in the passage of maternal antibodies to the infant, with a corresponding increase in the viral load to the baby.

1.7 **Treatment of HTLV-I Infection**

Treatment of the two clinical manifestations of HTLV-I infection have required different approaches. Corticosteroids are generally used to treat HAM/TSP patients (Osame *et.al.*,

1987), although their effect has been reported to be transient (Duncan and Rudge, 1990). ATL is generally treated with combination chemotherapy, although long-term success has been limited, especially for the more aggressive acute ATL (Kawano *et.al.*, 1985). A combination of zidovudine and alpha-interferon has shown a high response rate with prolonged survival (Bazarbachi and Hermune, 1996). Recently, Borg and colleagues (1996) have reported a successful allogenic bone marrow transplant in a patient with acute ATL. The patient was reported to be in remission and free of HTLV-I infection 23 months after surgery.

Gene therapy is slowly emerging as a viable option for treatment of HTLV-I infection. The approaches for treating ATL and HAM/TSP patients are once again different. For ATL gene therapy, corrective mechanisms directed at regulatory sequences rather than viral sequences may be more beneficial (Essex *et.al.*, 1995). This is because the level of virus replication becomes unimportant after the tumor is growing autonomously (Korber *et.al.*, 1991). For HAM/TSP, gene therapy directed to control virus replication may be more useful (Essex *et.al.*, 1995).

1.8 **Detection of HTLV-I Infection**

1.8.1 Serological Testing for HTLV-I

Viral infections are generally countered by the host immune response, which includes both humoral (the production of circulating antibodies) and cell-mediated (cytotoxic lymphocytes) immune responses. A variety of techniques are available to detect the presence of antibodies to HTLV-I, namely, enzyme immunoabsorbant assays (EIA) (Papsidero *et.al.*, 1990), western blot (or immunoblot) (WB) (Burnette *et.al.*, 1981),

immunofluorescence assay (IFA) (Hinuma *et.al.*, 1981), radioimmunoprecipitation assay (RIPA) (Schupbach *et.al.*, 1983) and particle agglutination (PA) (Verdier *et.al.*, 1990). Antibodies that are measured in these assays are usually the *gag*-encoded proteins (p19, p24, p15) and *env*-encoded glycoproteins (gp46 and gp61/68) (Verdier *et.al.*, 1990).. Enzyme immunosorbent assays are rapid colorimetric tests that can easily test large numbers of samples, and is the method of choice for screening donated blood (Khabbaz *et.al.*, 1990). Additional, more specific tests (for example, the WB assay) are suggested for confirmation of EIA positive results. The diagnostic criteria for a positive HTLV-I serologic test is reactivity to both p24 and gp46 or gp61/68 proteins (CDC, 1988). The WB assay is sensitive and specific, although the sensitivity of most WB assays is limited to their ability to detect antibodies against HTLV-I envelope proteins. The sensitivity can be improved by "spiking" the WB strips with recombinant-derived HTLV-I *env* antigens (Coates *et.al.*, 1990). The IFA is convenient and rapid, but has not been as widely employed as the WB, partly because it will not discriminate antibodies to a particular gene product, which is also the case for the PA assay (Khabbaz *et.al.*, 1990). The RIPA has good sensitivity and specificity, but is labour intensive and expensive, therefore limiting its use in routine diagnostic testing (Hartley *et.al.*, 1990).

1.8.2 Detection of Nucleic Acids

Once HTLV-I has integrated into the host cell genome, it often becomes latent, and may remain undetected as free viral particles. It is therefore important to be able to test for the presence of viral nucleic acids within infected cells. Techniques to detect relatively large copy numbers of viral nucleic acid include Southern blot assays and *in situ* hybridisation. Southern blots and *in situ* hybridisation procedures use labelled probes for detection from

blotted membranes or specifically treated slides of cells or tissues, respectively (Southern 1975; Pardue *et.al.*, 1970). The sensitivity and specificity of the assays are dependent on the sensitivity and specificity of the probe used to detect the proviral DNA. The application of the PCR for the detection of HTLV-I has allowed more sensitive and specific detection. This procedure will therefore be discussed in more detail.

1.8.2.1 PCR

PCR is an *in vitro* technique which allows the amplification of a specific DNA region that lies between two regions of known DNA sequence, in an exponential manner. It was invented in 1985 by Kary Mullis (Saiki *et.al.*, 1985) then working for Cetus Corporation in California. The original protocols for PCR amplification (Saiki *et.al.*, 1985; Mullis *et.al.*, 1987) used the Klenow fragment of *E. coli* DNA polymerase I to catalyse the oligonucleotide extension. However, this enzyme was thermally inactivated during the denaturation step of a PCR cycle, and fresh enzyme had to be added after each cycle. This hindered its implementation as a routine technique. However, this was overcome by the introduction of a thermostable polymerase, the first being *Thermus aquaticus (Taq)* DNA polymerase (Saiki *et.al.*, 1988) in the PCR as well as the automation of the reaction by the development of thermal cyclers (PCR machines).

Although extremely efficient, amplification of the target sequence in an exponential manner is not unlimited. Normally, the amount of enzyme becomes the limiting factor after 25 - 30 cycles (approximately 10^6 fold amplification). This is due to molar target excess, or due to thermal danaturation of the enzyme during the PCR process. Reannealing of target strands as their concentration increases, competes with primer annealing and also decreases efficiency.

Despite these limitations, PCR has already found extensive application in the diagnosis of genetic disorders (Engelke *et.al.*, 1988), detection of nucleic acid sequences of pathogenic organisms in clinical samples (Kwok *et. al.*, 1987), the genetic identification of forensic samples (Higuchi *et.al.*, 1988), and the analysis of mutations in activated oncogenes (Bos *et.al.*, 1987). PCR is also useful in the generation of specific sequences of cloned single stranded DNA for use as probes, the generation of libraries of cDNA from small amounts of mRNA, and the generation of large amounts of DNA for sequence analysis of mutations.

Similarly, PCR is an important tool in the research of HTLV-I. Amplification of the proviral DNA, using specifically designed primers can provide a rapid diagnostic test. If used in conjunction with the sequencing of amplified products it can provide valuable information on strain variation and origin. This will assist in molecular epidemiological studies aimed at clarifying the global distribution of HTLV-I.

PCR is also useful for seroindeterminate or seronegative subjects at risk of HTLV-I infection (Heneine *et.al.*, 1992). However, not all seropositive patients are found positive by a single step PCR, possibly due to low levels of infected lymphocytes. In such instances, a second amplification step ("nested" PCR), using internal primers, is used to improve the sensitivity of the PCR reaction (Matsumoto *et.al.*, 1990).

1.8.2.2 Quantitative PCR

The ability to measure and relate the final PCR amplified target to the amount in the original source material is refered to as quantitative PCR. There are numerous quantitative PCR methods currently available, such as competitive PCR (Gilliand *et.al.*,

Japanese Americans in the Hawaiian Islands (Blattner *et.al.*, 1986), among some remote groups in Papua New Guinea and the Solomon Islands (Yanagihara *et.al.*, 1990, and 1991; Nicholson *et.al.*, 1992) and some Aboriginal groups in northern and central Australia (Asher *et.al.*, 1988; Bastian *et.al.*, 1993). Locally, the virus is endemic in KwaZulu/Natal (KZN) (Bhigjee *et.al.*, 1993; 1994).

1.10 **Seroprevalence**

In Japan, the zones of endemic HTLV-I are spread heterogeneously (approximately 1 % of the population are HTLV-I carriers [Tajima *et.al.*, 1990]), with clusters in the Kyushu and Okinawa prefectures. Some of the districts in these areas have a seroprevalence of more than 30% (Tajima *et.al.*, 1990). In the United States, the prevalence is 0.025% among asymptomatic blood donors (Williams *et.al.*, 1988). In a study of the seroprevalence of HTLV-I in parts of Africa (Goubau *et.al.*, 1993), an overall seroprevalence of 4% was found. This can be broken down into a 1-15% prevalence in Zaire, 0.9% in Ghanian refugees in Belgium and 0.2% in pregnant women in South Africa. The figures for South Africa do not reflect the high seroprevalence in endemic regions, particularly KZN. Small studies done in South Africa have shown a seroprevalence of 0-5.2% in different ethnic and geographical areas (Bhigjee *et.al.*, 1993; Botha *et.al.*, 1989; van der Ryst 1996 [a]). Recently, the first cases of TSP in the Orange Freestate was shown by van der Ryst and colleagues (1996 [b]), where 18% of HTLV-I positive patients with spastic myelopathy had TSP. The seroprevalence of HTLV-I in asymptomatic groups in the Freestate ranged from 0-2%, and in patients with unexplained spastic myelopathy, the seroprevalence was 33.3% (van der Ryst *et.al.*, 1996 [a]). Most of the work done in KZN has been carried out by Bhigjee and colleagues. In a study

carried out in the Ngwelezane district in 1993, a 2.6% seroprevalence (95% CI 1.62-3.58) was found. The seropositivity rose with age to 6.1% in those over 55 years of age. This was possibly related to many of the individuals acquiring the infection only once sexually active. A follow-up study at Ubombo showed a seroprevalence rate of 3.33% (95% CI: 1.19-5.475). This emphasised the fact that HTLV-I is endemic in KZN.

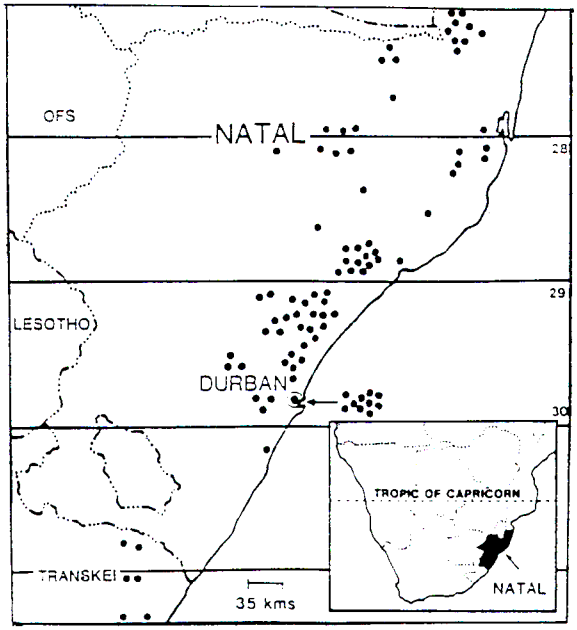


Fig. 1.2: Distribution of HAM/TSP cases in the KwaZulu/Natal region and the Transkei (Bhigjee *et.al.*, 1993).

1.11 **Influence of Geographical Area on Subtype**

The degree of sequence similarity between HTLV-I strains appears to depend on geographic origin (Komurian *et.al.*, 1991; Paine *et.al.*, 1991). Yanagihara *et.al.* (1992), and later Nerurkar *et.al.* (1994), divided the strains into two major geographic subtypes, the cosmopolitan strains of Japan, the Caribbean, the Americas and Africa, and the Australo-Melanesian strains of Papua New Guinea, the Solomon Islands and Australian Aboriginals. The cosmopolitan strains exhibit more than a 98.5% sequence similarity among themselves (Malik *et.al.*, 1988; Gessain *et.al.*, 1992), while Australo-Melanesian strains have only a 93% similarity with the cosmopolitan strains (Gessain *et.al.* 1993;

Bastian *et.al.*, 1993). The Australo-Melanesian strains share multiple nucleotide substitutions not found in the cosmopolitan strains (Gessain *et.al.*, 1992; Yanagihara *et.al.*, 1994). For example, at the 5' end of the *env* gene at base position 5251-5263, multiple base substitutions (some resulting in amino acid changes) are found (Sherman *et.al.*, 1992; Bastian *et.al.*, 1993; Gessain *et.al.* 1993). These substitutions serve as a signature sequence for the Australo-Melanesian strains (Yanagihara *et.al.*, 1994). Similarly, nucleotide substitutions occur in base positions 6113-6128 and 6410-6434 (Nerurkar *et.al.*, 1994; Yanagihara *et.al.*, 1994). Primers derived from sequences unique to these variants have been used to discriminate HTLV-I strains into the two geographic-specific subtypes (Nerurkar *et.al.*, 1994). Yanagihara *et.al.* (1994) further divided the Australo-Melanesian strains into the Papau New Guinea, Solomon Island and Australian-Aboriginal strains. These strains differed by approximately 3.8% between themselves. Similarly, the cosmopolitan strains were divided into the Zairian (Central African) and Afroindoamerasian strains. More recent reports have separated the Central African strain from the cosmopolitan strains. This gave a total of three HTLV-I subtypes: 1a (cosmopolitan), 1b (Central African), 1c (Melanesian) (Vandamme *et.al.*, 1994; Lui *et.al.*, 1996). Focusing on a 673 bp region of the LTR gene, Lui *et.al.* (1996) found the Central African strain to differ from the prototype cosmopolitan strain (ATK-1) by 5%, with the rest of the cosmopolitan strains differing by only 0.6% -2.6%.

Recently, Mahieux *et.al.*, (1997) have discovered a new subtype isolated from pygmies in Central Africa. Using 58 new HTLV-I African strains (18 from West Africa, 36 from Central Africa and four from South Africa), the authors showed that while the West and South African strains fell within the cosmopolitan subtype, three of the 36 Central African strains did not cluster with the Central African subtype. They

called this new subtype subtype D (Mahieux *et.al.*, 1997).

On the other hand, a report by Vidal *et.al.* (1994) also suggested that the HTLV-I strains be divided into five major subtypes: cosmopolitan, Japanese, West African, Central African and Australo-Melanesian. They report the West African and Japanese strains to form separate clusters within the cosmopolitan subtype.

1.12 **Phylogeny**

Reconstruction of the evolutionary history of genes and species is one of the most important subjects in the current study of molecular evolution. Phylogenetic relationships of genes or species are usually presented in a treelike form with a root ("rooted" tree) or a tree without a root ("unrooted" tree). There are many possible rooted and unrooted trees for a given number of samples, with the number rapidly increasing as the number of samples increase (Cavalli-Sforza and Edwards 1967). However, there is only one true tree. Great care should be taken so that reliable phylogenies are produced as these can shed light on the sequence of evolutionary events that generated the present day diversity of genes and species. This will help us to understand the mechanisms of evolution as well as the history of organisms.

There are numerous methods that can be used to construct phylogenetic trees from molecular data (Felsenstein 1989, Miyamoto and Crucraft 1991). These methods can be classified into distance methods and discrete-character methods.

In distance methods, a pair-wise evolutionary distance is computed for all Operational Taxonomic Units (OTUs) to be studied, and a phylogenetic tree is constructed by certain principles and algorithms. Nei (1991) has shown that one of the most efficient distance

methods in recovering the correct topology is the Neighbour-joining method (NJ) proposed by Saitou and Nei (1987). In discrete-character methods (for example, Maximum Parsimony [MP]), data with discrete character states such as nucleotide states in DNA sequences are used, and a tree is constructed by considering the evolutionary relationships of OTUs in DNA sequences at each character or nucleotide position (Nei 1991).

The evolutionary distances between sequences are usually measured by the number of nucleotide and amino acid substitutions between them. The simplest method of distance estimation is the *p*-distance method. Others include the Kimura-2-parameter model the Jukes Cantor method, the Poisson-correction and Gamma distance methods. Various computer packages are available for phylogenetic analyses. These include the Phylogeny inference package (PHYLIP, Felsenstein 1989), TREECON (Van de Peer and De Wachter, 1993) and Molecular Evolutionary Genetics Analysis (MEGA, Kumar *et.al.*, 1993). The computer packages incorporate many of the tree drawing and distance estimation methods, so that various combinations can be used and the trees compared.

The sequence variations of HTLV-I serve as biological markers for early and recent migrations of human populations (Yanagihara *et.al.*, 1994). Sequence and phylogenetic evidence suggest that HTLV-I originated from its simian counterpart simian T-cell leukaemia virus type I (STLV-I) (Koralnick *et.al.*, 1993; Vandamme *et.al.*, 1994; Yanagihara *et.al.*, 1994). Some African STLV-I strains are genetically indistinguishable from African HTLV-I strains (Koralnick *et.al.*, 1993), while STLV-I strains from Japan and India show only a 90% homology to HTLV-I strains of the same area (Song *et.al.*,

1994). This suggests several interspecies transmission events, with most of the evidence pointing to Africa as the place where the original STLV-I/HTLV-I transmission event occurred (Gessain *et.al.*, 1992; Vandamme *et.al.*, 1994). Conflicting theories suggest an Asian origin (Saksena *et.al.*, 1992; Lui *et.al.*, 1996). As both Asia and Africa harbour very divergent HTLV-I/STLV-I strains, the origin of HTLV-I/STLV-I will be biased towards the most divergent strain used in the analysis (Lui *et.al.*, 1996).

HTLV-I sequence data from the southern part of Africa is not as well documented as the rest of Africa. Sequencing of the local HTLV-I strain will therefore provide useful information on the subtype prevalent in this region. In this study, an in-house molecular diagnostic test is developed for the detection of HTLV-I. In-house PCR results are compared with the Acugene AmpliSensor assay (a novel qualitative and quantitative technique) using the Western bolt assay as the "Gold" standard. The *env* region of the HTLV-I genome is amplified from five local HTLV-I positive patients. The amplicons are cloned and sequenced using the Automate Laser Fluorescent (A.L.F.) Express and Applied Biosystems Incorporated (ABI) Automated Sequencers. Phylogenies are constructed using selected sequences from around the world to determine the identity of the local subtype.

1.13 **Aims and Objectives**

1. To develop and optimise a molecular diagnostic assay for the detection of HTLV-I.
2. To clone and sequence segments of HTLV-I from five local isolates.
3. To establish the predominant strain of HTLV-I in KZN.

Chapter 2

Molecular Diagnosis of HTLV-I

2.1 **Introduction**

There is a general trend towards the use of molecular techniques for diagnosis, particularly where serological results are inconclusive. PCR has an increased sensitivity, able to amplify as few as 10 copies of the viral genome. The first PCR detection of HTLV-I proviral sequences was accomplished by Kwok *et.al.* (1988). The use of PCR for diagnostics, however, has been hindered by the lack of a convenient, sensitive and quantitative method for detecting amplified products. Gel-based methods of detection are time-consuming and expose workers and equipment to contamination with amplified DNA. Other nonisotopic detection methods, based on enzyme-linked assays (Loeffelholz *et.al.*, 1992) or electrochemiluminescence (DiCesare *et.al.*, 1993), represent major improvements on gel-electrophoresis, but still require post-amplification signal-development steps and direct manipulations of amplified DNA, which can lead to amplicon contamination. The novel AmpliSensor assay (Biotronics) utilises a fluorescence energy transfer based process, where detection occurs concurrently with the final stages of amplification (Wang *et.al.*, 1995). The AmpliSensor assay also has the option of being quantitative. Viral quantitation is increasing in popularity in molecular diagnosis (particlarly for HIV) as it indicates the extent of infection, and can be used as a prognostic indicator of disease progression. More relevant though, is its use in the management of chronic viral infections such as HIV, HTLV, HCV, HBV and others.
In this chapter, an in-house PCR diagnostic test for HTLV-I is optimised. Primers were designed, targeting two regions of the genome, the *pol* and *env* genes. The *pol* primers (SK110 and SK111) were designed on the reported primer sequences of Kwok *et.al.* (1988), which proved to be very reliable (100% sensisivity and specificity). The specificity of the PCR assay was confirmed using a specific probe SK112 (Kwok *et.al.*,

1988). Primers targeting the more variable *env* gene (6275 [forward] and 6666 [reverse]) amplified a 410 bp product. The PCR product was confirmed by sequencing. Quantitation was performed on the samples using the AmpliSensor assay. A comparison was made between the PCR and WB assay, using the WB assay as the "Gold" standard.

2.2 **Methods**

2.2.1 Subjects

Subjects were seropositive patients with HAM-TSP seen in the Department of Neurology, Wentworth Hospital, or patients with infective eczema seen in the Skin Clinic at King Edward VIII Hospital, Durban. A brief clinical profile of the patients can be seen in Table 2.1. Patients' blood samples were collected in heparin free EDTA tubes (to prevent agglutination of the blood) and processed as soon as possible, or stored at 4°C until processed (usually within 48 hours). One-hundred and fourteen WB negative samples were obtained from the Department of Medicine, King Edward VIII Hospital, Durban. These samples were pooled into sets of six, to give nineteen samples (numbered one to nineteen), and tested with the in-house PCR as well as the AcuGene AmpliSensor assay. No history was given for these patients.

TABLE 2.1: Patient data of all patients included in the study. All patients were clinically diagnosed with HTLV-I infection at the Department of Neurology, Wentworth Hospital, and the Skin clinic at King Edward VIII hospital.

No	Sex	Age	Hospital	Diagnosis
23	F	-	W	H
48	F	30	W	H
49	F	60	W	H
50	F	55	W	H
51	F	24	W	H
52	M	38	W	H
57	F	57	K	H
58	M	38	W	H
59	M	-	K	S
60	F	40	K	S
61	F	13	K	S
62	F	41	W	H
64	M	49	W	H
65	F	46	W	H
66	F	50	W	H
67	M	27	W	H
68	M	51	W	H
69	M	35	K	H
70	F	-	K	H
71	F	-	K	H

F=female; M=male; W=Wentworth Hospital; K=King Edward VIII Hospital; H=HTLV-I myelopathy; S=skin disease

2.2.2 DNA Extraction

DNA was extracted using the scale-down protocol of the GENOMIX kit (Talent, Srl, Italy). The Talent GENOMIX kit can be used for the extraction of DNA from whole blood, cells and tissue. An average of 30-50 ug of genomic DNA per ml of whole blood is obtained per extraction.

2.2.2.1 Reagents

Solution 1 (Lysing reagent): provided in the kit -ingredients unknown

Solution 2 (Precipitating reagent): provided in the kit -ingredients unknown

Solution 3 (Exchange reagent): provided in the kit -ingredients unknown

3M Sodium acetate: 40.81 g of sodium acetate $\cdot 3H_2O$ was dissolved in 80 ml of water. The pH was adjusted to 7.0 with dilute acetic acid. The volume was adjust to 100 ml with water. The solution was autoclaved. Unless otherwise stated, all water refers to deionised water.

70% ethanol: (Appendix I)

95% ethanol: (Appendix I)

2.2.2.2 Procedure

The three solutions were aliquoted into separate containers before beginning the extraction to prevent any cross-contamination by the DNA. Disposable pipettes were also used as often as possible to prevent cross-contamination. Six hundred ul of the Lysing reagent (solution 1) was added to each 2 ml eppendorf tube. Three hundred ul of blood was transferred into each tube and mixed. The tubes were heated for two minutes in a water bath at 68°C. Care was taken that the time did not exceed two minutes, as this would have resulted in a lower yield. Nine hundred ul of chloroform was added to the lysate and thoroughly mixed. The tubes were centrifuged for three minutes in a Sorval microcentrifuge (DuPont model MC12V) at 12500 x g. After centrifugation, each aqueous phase was placed in second eppendorf, to which 900 ul of precipitating reagent (solution 2) had been added. The tubes were mixed several times until the precipitate of DNA could be seen, The tubes were then centrifuged for two minutes at 5200 x g. The liquid

phase from each tube was carefully removed and 500 ul of exchange reagent (solution three) was added to resuspend the pellet. One ml of 95% ethanol was added to re-precipitate the DNA (this was done by gently inverting each tube). The tubes were centrifuged for two minutes at 3100 x g. The supernatant from each tube was carefully removed, each precipitate was washed with 500 ul of 70% ethanol. The tubes were centrifuged for another two minutes at 3100 x g. The supernatant was removed and each tube was placed inverted on paper towel until all the ethanol had evaporated. Each pellet was resuspended in 100 ul of distilled water.

An additional precipitation step was performed to further purify the DNA. The DNA was precipitated as follows: Ten ul of 3 M sodium acetate (0.3 M final concentration) and 255 ul of 95% ethanol (2.5 volumes) were added to the DNA. Following one hour at -20°C, the tubes were centrifuged for 15 minutes at 5200 x g, after which, the supernatants were carefully removed. The pellet was washed in 100 ul of 70% ethanol and centrifuged for five minutes at 3100 x g. The supernatants were removed and each tube was placed inverted on paper towel until all the ethanol had evaporated. Each pellet was resuspended in 100 ul of distilled water.

2.2.3 Measurement of DNA Concentration

The DNA concentration of the target template is an important component of the PCR reaction. After each extraction, the concentration was measured. The procedure was as follows: A 1:100 dilution (1 ul DNA and 99 ul water) of extracted DNA was pipetted into a quartz cuvette. The absorbance was read at 260 nm and 280 nm using a GeneQuant spectrophotometer (Pharmacia, Biotech), and water as the blank. An optical density (OD)

of one corresponded to approximately 50 ug/ml of double-stranded DNA. The concentration was automatically calculated by the GeneQuant. The ratio between the readings at 260 nm and 280 nm provided an estimate of the purity of the DNA (pure preparations have a ratio of 1.8). The concentration was measured in duplicate and the average concentration was used.

2.2.4 Primer Design

Primers were designed to target two areas of the genome, namely the *pol* and *env* regions. The *pol* primers targeted the region from 4757 to 4942 (Kwok *et.al.*, 1988), giving a 186 bp product. The primer sequences were:

SK110: 5` CCC TAC AAT CCC ACC AGC TCA G 3'

SK111: 5' GTG GTG GAT TTG CCA TCG GGT TTT 3'
5` biotinylated

The env primers targeted the area from 6275 to 6666 (Genebank accession no: J02029), giving a 410 bp product. The primer sequences were:

6275 (forward): 5' CTC AAG CAA TAG TCA AAA 3'

6666 (reverse): 5' ATT GCG TGC TTG GTT TAC 3'

The position of the primers in relation to the genome can be seen in Fig. 2.1.

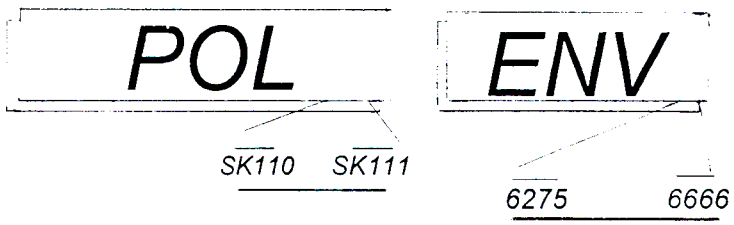


Fig. 2.1. Design of primers targeting two areas of the HTLV-I genome (*pol*: 4757 to 4942 and *env*: 6275 to 6666).

2.2.5 Dilution of Primers

Primers were diluted to give a 20 pmol/ul working solution and stored at 4°C. The dilution was calculated using the following formula:

$z = {}^{x}/_{y}$,

(x = OD of primer x 30)

(y = length of primer x 0.33 [average molecular weight of a dNMP X 10^{-3}])

For 20 pmol/ul (p):

$p = {}^{z}/_{20}$

For SK110:

OD = 122; length = 22

Therefore dilution = 1 ul + 24.2 ul H_2O

For SK111:

OD = 118; lenghth = 24

Therefore dilution = 1 ul + 213.5 ul H_2O

For 6275:

OD = 22.3; length = 20

Therefore dilution = 1 ul + 4.07 ul H_2O

For 6666:

OD = 18.7; length = 20

Therefore dilution = 1 ul + 3.25 ul H_2O

2.2.6 Optimisation of the PCR Reaction

The PCR reaction is influenced by the concentrations of the components of the reaction. These components were addressed one at a time until the reaction was optimal and reproducible.

2.2.6.1 Reagents

Boehringer Mannheim 10 X buffer: (Appendix I)

DNTPs: (Appendix I)

Taq polymerase: (Appendix I)

Mg^{2+}: (Appendix I)

2.2.6.2 Procedure

The reaction mixture was prepared as follows: 5 ul of Boehringer Mannheim 10 X buffer; 250 uM dNTPs (Pharmacia); 20 pmol primers; one unit of *Taq* (Boehringer Mannheim). The magnesium chloride (Mg^{2+}) concentration was titrated from 1.5 mM to 4 mM (2.0 mM, 2.5 mM, 3.0 mM, 3.5 mM, 4.0 mM). The PCR was performed in a Perkin Elmer 2400 thermocycler according to the following conditions: 94°C for five minutes; 40 cycles of 94°C for one minute, 60°C for 30 seconds and 72°C for 30 seconds; 72°C for seven minutes; and held at 4°C.

Once the optimum Mg^{2+} concentration was determined, this was used in the optimisation of the primer concentrations, which were titrated as 5 pmol, 10 pmol, 20 pmol, 50 pmol and 100 pmol each for SK110/111, and 40 pmol, 60 pmol, 80 pmol, 100 pmol, 120 pmol each for the *env* primers.

2.2.7 Detection of PCR Products

2.2.7.1 Agarose Gel Electrophoresis

2.2.7.1.1 Principle

Agarose gels are capable of separating and resolving DNA fragments from 200 bp to approximately 50 kb in length. The technique is rapid, simple, and the location of the DNA can be determined directly by staining with low concentrations of the fluorescent intercalating dye, ethidium bromide (EtBr) (Sharp *et.al.* 1973), and viewing under ultraviolet light. The size of the DNA is determined by comparison to a DNA marker containing DNA fragments of known base pairs. Agarose gels are prepared by melting the agarose in the presence of the desired buffer until a clear, transparent solution is obtained. This is then poured into a casting tray and allowed to harden. With the application of an electric field across the gel, the DNA migrates toward the anode. The migration of the DNA depends primarily on the size of the DNA fragment, as larger molecules migrate more slowly because of greater frictional drag and are less efficient in moving through the gel matrix (Helling *et.al.* 1974).

2.2.7.1.2 Reagents

1 X TBE: 200 ml of 5 X TBE (Appendix I) was diluted with 800 ml water.

EtBr: 1 g of EtBr was diluted in 100 ml distilled water, to give a 10 mg/ml stock. The stock solution was aliquoted into 2 ml eppendorf tubes. The tubes were covered with aluminium foil and stored at room temperature.

Gel loading buffer (6 X): 25 g of bromophenol blue, 25 ml xylene cyanol FF and 15 ml Ficoll was made up to 100 ml with water. The loading buffer was aliquoted in bijou bottles and stored at room temperature.

2.2.7.1.3 Casting of Gels

One gram of agarose per 100 ml of 1 X TBE was heated in a microwave until completely mixed. When cool to the touch, 5 ul EtBr (10 mg/ml) was added and the gel poured into perspex casting trays. The required number of wells was obtained using a comb with the appropriate number of teeth. Two combs could be used per tray. After the gel was completely set (30-45 minutes at room temperature), the comb was removed and the gel was ready to be placed in the electrophoresis tank.

2.2.7.1.4 Running of Gels

The gels were placed in a horizontal electrophoretic apparatus (Hybaid-Electro-4) containing electrophoresis running buffer (1 X TBE). Ten ul of the PCR product was mixed with 2 ul of the 6 X gel loading buffer. The samples were carefully loaded into the wells. The dense Ficoll prevented the sample from floating out of the wells once pipetted, and the two dyes served as indicators to monitor the migration of the samples. The gel was run at 100 volts (Hoefer Scientific).

A molecular weight marker was run concurrently to calculate the size of the amplicons. After approximately one to two hours the gel was viewed.

2.2.7.1.5 Imaging of Gels

The gels were viewed on a UVP (Ultra-Violet Products) transilluminator (312 nm), with the DNA visible as a fluorescing orange band. The images were captured using a UVP camera connected to the UVP image store 5000 system. The image was printed (Sony video graphic printer), or stored on a computer disk.

2.2.7.2 Southern Blotting

Where detection on an agarose gel was not satisfactory, the PCR products were transferred from the gel onto a nitrocellulose membrane, and detected using a specific probe (SK112). The non-radioactive DIG oligo 3'-end labelling kit was used to label the probe. The kit is faster and safer than radioactive procedures.

Base sequence of the probe (SK112):

5' GTA CTT TAC TGA CAA ACC CGA CCT AC 3'

The probe was diluted to a 20 pmol/ul, and stored at -20°C. The dilution was calculated according to the formula described in 2.2.5.

OD = 23.9; length = 30

Therefore, the dilution was: 1 ul probe + 2.62 ul water.

2.2.7.2.1 Reagents

Reagents from the DIG oligonucleotide 3'-end labeling kit (Boehringer Mannheim):

Vial 1 (Reaction buffer [5X]): 1 M pottassium cacodylate, 0.125 M tris HCl, 125 mg/ml bovine serum albumine, pH 6.6.

Vial 2 ($CoCl_2$ solution): 25 mM $CoCl_2$.

Vial 3 (DIG-ddUTP solution): 1 mM DIG-ddUTP in redistilled water.

Vial 4 (Terminal transferase): 50 U/ul in 0.2 M pottassium cacodylate, 1 mM EDTA, 200 mM KCl, 0.2 mg/ml bovine serum albumine, pH 6.5, glycerol 50% (v/v).

Additional Reagents:

Denaturation buffer: 300 ml of 5 M NaCl (Appendix I) and 500 ml 1 M NaOH (Appendix I) were diluted to one litre with water and autoclaved. (Final concentrations:

1.5 M NaCl and 0.5 M NaOH.)

Neutralisation buffer: 300 ml 5 M NaCl (Appendix I) was added to 700 ml of 1 M Tris (Appendix I).

20 X SSC: 175.3 g of NaCl and 88.2 g of sodium citrate were dissolved in 800 ml of water. The pH was adjusted to 7.0 with a few drops of NaOH. The volume was adjusted to 1 litre with water. The solution was autoclaved and stored at room temperature.

2 X SSC: 20 ml of 20 X SSC was diluted with 180 ml of water.

10% Sodium dodecyl sulfate (SDS): (Appendix I)

Wash 1: 2 ml of 10% SDS (Appendix I) was diluted with 198 ml of 2 X SSC.

Buffer 1: 1 g of maleic acid and 8.7 g of NaCl were dissloved in 800 ml of water. The pH was adjusted to 7.5 with solid NaOH, and the volume was adjusted to one litre with water. (Final concentrations: 0.1 M maleic acid and 0.15 M NaCl.)

Stock blocking solution: 10 g of blocking powder were dissolved in 90 ml of buffer 1.

Buffer 2: 10 ml of stock blocking solution was diluted with 90 ml of buffer 1.

Wash 2: 300 ul Tween 20 was diluted to 100 ml with buffer 1.

Buffer 3: 10 ml of 1 M tris HCl (Appendix I), 2 ml of 5 M NaCl (Appendix I), and 500 ul of 1 M Mg^{2+} (Appendix I) were diluted with 80 ml of water. The pH was adjusted to 9.5 with NaOH. The volume was adjusted to 100 ml with water. (Final concentrations: 0.1 M tris HCl, 0.1 M NaCl, 50 mM Mg^{2+}.)

Reagents from the Anti-digoxigenin-detection kit (Boehringer Mannheim):

Anti-digoxigenin-AP conjugate: FAB frgaments from an anti-digoxigenin antibody from sheep, conjugated with alkaline phosphatase.

NBT solution: 75 mg/ml nitroblue tetrazolium salt in 70% dimethyl formamide (v/v).

X-phosphate solution: 50 mg/ml 5-bromo-4-chloro-3-indolyphosphate, toluidinium salt in

100% dimethylformamide.

2.2.7.2.2 Procedure

The gel was soaked in denaturation solution for 30 minutes. It was then placed in neutralisation buffer for ten minutes. After draining off the buffer, the gel was placed inverted between pre-cut nylon membrane (MSI, Amersham) and subsequent layers of 3mm whatman paper and tissue paper. A weight was placed on the tissue to facilitate capillary transfer, which took place overnight. The following day, the DNA was covalently linked to the nylon membrane by exposing the nylon (DNA face down) to ultra-violet light for five minutes.

Labelling the probe: The probe was 3' end labelled with DIG-ddUTP using a 3' end labelling kit from Boehringer Mannheim. Briefly, 4 ul of reaction buffer (vial 1), 4 ul $CoCl_2$ (vial 2), 100 pmol probe (SK112), 1 ul DIG-ddUTP (vial 3) and 1 ul terminal transferase (vial 4) was brought up to 10 ul with water. This was incubated for 15 minutes at 37°C .

While the probe was being labelled, the membrane was pre-hybridised in hybridisation buffer (DIG Easy Hyb) for ten minutes at 55°C.

Ten ul of the 3' DIG labelled probe was added to the hybridisation buffer and incubated for one hour at 55°C. After hybridisation, the filter was washed (while being shaken) to remove unbound probe. This was done four times with wash 1, three minutes each, at room temperature. This was followed by one wash at 55°C. The membrane was rinsed in buffer 1 then blocked in 20 ml of buffer 2 for 30 minutes.

Four ul of the conjugate (anti-digoxigenin-AP, Boehringer Mannheim) was added to buffer 2 for 30 minutes.

After conjugation, the membrane was washed twice, for 15 minutes each, with wash buffer 2. The membrane was then equilibrated for two minutes in 10 ml of buffer 3. Forty-five ul of NBT and 35 ul of X-phosphate solution were added to buffer 3 and incubated overnight in the dark.

The following day, the reaction was stopped with 50 ml of water.

2.2.8 Sequencing of Amplicons

The 410 bp PCR product was cloned and sequenced according to the methods decribed in 3.2.5 to 3.2.15. The sequence was compared to the Japanese prototype sequence ATK-1 (Genebank accession no.:J02029; Seiki *et.al.*, 1983).

2.2.9 Acugene AmpliSensor Assay

2.2.9.1 Principle

The assay is based on the principle that fluorescence energy transfer can be used to detect duplex formation between complementary nucleic acid strands. If the two complementary strands are labelled with donor and acceptor fluorophores, respectively, fluorescence energy transfer between the flourophores will be facilitated when the strands are paired, or eliminated when the paired strands are disrupted. The amount of energy transfer is used to measure the amount of duplex formation between the fluorophore-labelled oligonucleotide complex, thereby relating the extent of duplex formation mediated by the DNA polymerase.

The AmpliSensor assay was performed using DNA extracted (according to the talent protocol described previously) from 14 seropositive HAM/TSP patients and 24 seronegative patients. Details of these patients are given in Table 2.1.

2.2.9.2 Reagents

The AmpliSensor kit was obtained from Biotronics Corporation, Boston MA, USA. The contents of the kit were as follows:

Amplification Mix: Excess primer (0.9 uM) and limiting primer (90.12 uM) in 50 mM Tricine, pH 8.9, 0.2 mM dNTPs, 40 mM KCl, 1 mM DTT, 4 mM Mg^{2+} and detergent.

AmpliSensor Mix: AmpliSensor duplex (0.14 uM) in 50 mM Tricine, pH 8.9, 40 mM KCl, 1 mM DTT, 3 mM Mg^{2+} and detergent.

Bleach Mix: 10 ng/ul of probe complementary to the internal standard.

Positive standard: 100000 copies/ul of plasmid containing target sequence from HTLV-I *pol gene*).

Dilution buffer: 40 ng/ul of yeast transfer RNA

Internal standard: 100000 copies/ul of HTLV-I target DNA.

Mineral oil.

All reagents were prepared according to the protocols decribed by Wang *et.al.* (1995).

Primers used in the amplification reaction were:

Limiting primer (HTLV-I and II): 5' CCC TAC AAT CCA AAC CAG CTC AGG 3'

Excess primer (HTLV-I specific: 5' GAA TGT GTC TCT GGG ATC GGC TGG 3'

All amplifications were carried out in 96-well polycarbonate plates using an AG-9600 Silver Block Thermal Cycler (Biotronics Corporation, Lowell, MA).

2.2.9.3 Procedure

All reagents were thawed, vortexed, spun briefly and placed on ice before proceeding. The protocol followed was as described in the AcuGen HTLV-I All-in-one kit (Acugen

Systems, cat no. 20-1080-70). Briefly, One-hundred-and-ninety-two ul of Amplification Mix was aliquoted in a 2 ml screwcap tube (sufficient for 24 reactions). To this, 0.9 ul of Internal standard (100000 copies/ul) was added (now called reaction mix). The internal standard would later indicate the efficiency of the amplification reaction. Four ul of *Taq* DNA polymerase (10 U/ul) (Boehringer Mannheim) was added to the reaction mix, vortexed and spun briefly. The reaction mix was kept on ice until ready to be used.

A 1:5 serial dilution of the standard control (10^5 copies/ul) was performed (20000 copies/ul to 6.4 copies/ul).

Two ul of each standard (excluding the first dilution) was pipetted into the wells of a polycarbonate plate (i.e. 8000 copies, 1600 copies, 320 copies, 64 copies and 13 copies). Two ul of each sample, and 2 ul of water for the blank, were dispensed into the appropriate wells. An example of the layout is given in Table 2.2. Eight ul of Amplification Mix (without Internal standard and *Taq*) was added to the Apex.

The plate was placed in the AG-9600 Analyzer (Biotronics), which is controlled by a computer workstation using the AmpliSensor Analysis Program (ASAP) software package (Acugene Systems, Lowell, MA). The reaction mix was screwed onto dispensing port one and mineral oil was screwed onto the dispensing port four.

Eight ul of reaction mix was automatically dispensed to each well, except the Apex. All wells were overlaid with mineral oil. The plate was placed in the AG-9600 Silverblock (Biotronics) and amplified as follows: 23 cycles of 94°C for 25 seconds, 60°C for 25 seconds, 72°C for 40 seconds; 72°C for 40 seconds; and 20°C for one minute (to equilibrate the signal).

After the initial amplification, the AmpliSensor was added as a "hemi-nested" primer to

detect duplex formation between complementary nucleic acid strands. Briefly, 96 ul of AmpliSensor was aliquoted into a 2 ml screwcap tube, and screwed onto a dispensing port. Four ul of AmpliSensor was automatically dispensed into each well, except the blank. One more cycle was run and a base reading was taken using the AG-9600 Analyser (Biotronics). The information was interpreted by the ASAP software package (Acugene Systems, Lowell, MA). Readings were taken every three cycles until the required sensitivity was achieved.

To test the efficiency of the amplification reaction, an Internal Control bleach was added to the reaction. Ninety-six ul of Internal Control Bleach was aliquoted into a 2 ml screwcap tube, to which 0.9 ul of Taq (10 U/ul) was added, vortexed and spun. Four ul of Bleach mix was automatically dispensed to each well, except the apex.

One cylce was run and the program was switched to ICP mode.

A base reading was taken, and readings were taken every two cycles until the required sensitivity was achieved.

TABLE 2.2: Example of the sample layout for the Acugene AmpliSensor assay. The samples tested were isolated from the peripheral blood of patients 48, 49, 51, 52, 57, 58, 62, 64, 65, 66, 67, 68, 61, 60 and 69, respectively.

	1	2	3
A	Neg	48	64
B	Std1	49	65
C	Std2	51	66
D	Std3	52	67
E	Std4	57	68
F	Std5	58	61
G	Blank	Apex	60
H	Neg	62	69

2.2.9.4 Data Processing and Interpretation

Raw readings were first corrected for background by subtracting the blank value. All readings were compared to a constant reference point (the Apex value) to correct for well-to-well differences. All the raw data was reduced to detection index, whose magnitude was then in linear proportion to the amount of amplification product. The readings were converted to detection indexes by a mathematical manipulation, which takes into account both the template and primer availabilities as the rate limiting factors of amplification. The corrected signal change was then plotted against the cycle number to reflect a true exponential accumulation process of the amplified product. A standard curve correlating the target dosage to the cycle number was established (Fig. 2.2).

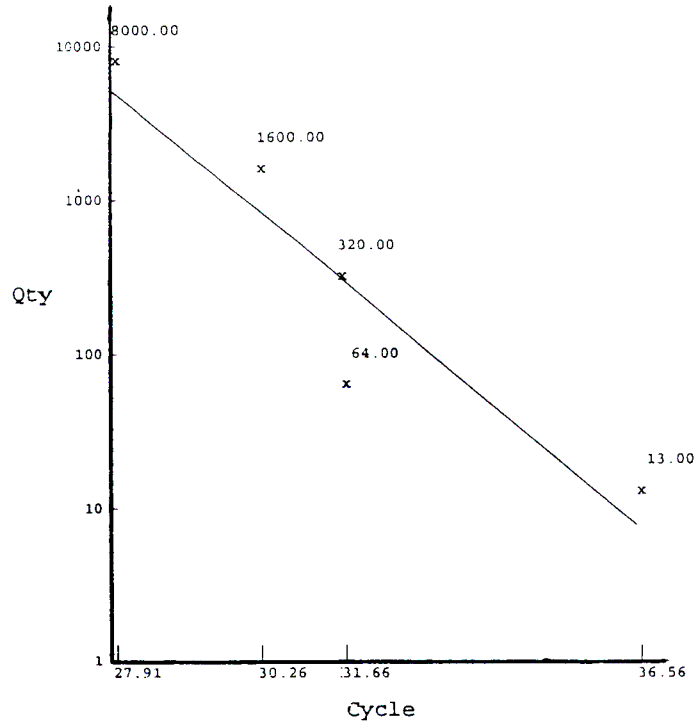


Fig. 2.2: Standard curve correlating the target dosage to the cycle no. in the AmpliSensor assay. The standards were titrated from 8000 copies to 13 copies of template.

For qualitative threshold analysis, a cutoff signal value (the standard with the lowest dilution) was chosen. Samples with detection indexes significantly above the threshold were designated a "+" status, while those below the threshold were designated a "-"

status. For those samples of detection indexes fluctuating within two times of the standard deviation about the threshold, an ambiguity ("?") status was assigned.

For qualitative analysis by linear regression, the detection index was related to the target quantity by an empirical function, which is characterised by a four parameter equation. The equation was simplified for quantitation purposes, by extracting the linear component for each assay cycle. Quantitation was performed by interpolating the standard curve for each cycle. The window for regression analysis was determined by the intercept of the upper and lower limit of the linear part of the standard curve. The quantitation was performed within the window by direct interpolation in a cycle dependent manner.

2.2.10 Western Blot (WB) Assay

Although it was possible that some infected patients were negative by serology, for the purpose of this study the WB result was used as the Gold standard. The WB test was performed by the staff at the Serology laboratory in the Department of Virology, King Edward VIII Hospital, Durban. The HTLV Western Blot kit 2.3 (Diagnostic Biotechnology) was used.

2.2.10.1 Principle

The nitrocellulose strips are labelled with HTLV-I viral proteins derived from native inactivated disrupted viral particles and genetically engineered proteins. Individual nitrocellulose strips are incubated with diluted serum or plasma specimens and controls. Specific antibodies to HTLV-I/II, if present in the specimen will bind to the HTLV-I/II proteins on the strips. The strips are washed to remove unbound materials while antibodies that bind specifically ot the HTLV proteins can be visualised using a series of

reactions with goat anti-human IgG conjugate with alkaline phosphtase and the substrate, BCIP/NBT. This method is sensitive enough to detect marginal amounts of HTLV antibodies in serum or plasma.

2.2.10.2 Reagents

The reagents used for the HTLV WB was provided in kit form. They have been listed here for completion. However, the concentrations and fine details of the reagents are not always indicated, as they were not disclosed by the manufacturers.

Nitrocellulose strips: HTLV-I viral lysate and recombinant *env* antigens.

Non-reactive control: Inactivated normal human serum.

Strong reactive control I: Inactivated human serum with high titred antibodies to HTLV-I.

Strong reactive control II: Inactivated human serum with high titred antibodies to HTLV-II.

Lyophilised stock buffer: Tris buffer with inactivated animal and non-animal proteins. The buffer was reconstituted in 100 ml of water

Wash buffer concentrate (20 X): Tris with Tween-20. One volume of buffer was diluted with 19 volumes of water to give the diluted wash buffer.

Blocking powder: Non-fat dry milk.

Blocking buffer: 1 g of blocking powder was dissolved in 20 ml of reconstituted stock buffer.

Conjugate: Goat anti-Human IgG conjugated with alkaline phosphatase. Ten ul of the conjugate was diluted with 10 ml of blotting buffer to give the working conjugate solution.

Substrate: Solution of 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) and nitroblue

tetrazolium (NBT).

2.2.10.3 Procedure

The method was briefly as follows:

Numbered strips were placed in separate wells, including strips for strong reactive and non-reactive controls. Two ml of diluted wash buffer were added to each well. The strips were incubated for five minutes at room temperature on a rocking platform. The buffer was aspirated. Two ml of blocking buffer were added to each well, followed by 20 ul of each patient's sera, or controls to appropriate wells. The tray was covered and incubated for one hour at room temperature on a rocking platform. After incubation, the wells were aspirated. Each strip was washed three times with diluted wash buffer, allowing a five minute soak between each wash. Two ml of working conjugate solution was added to each well, and incubated for one hour at room temperature. The wells were aspirated. Two ml of substrate solution were added to each well and incubated for ten minutes. The wells were aspirated and rinsed with water to stop the reaction. The strips were removed and dried on paper towel. Dried strips were stored in the dark.

2.3 **Results**

2.3.1 In-house HTLV-I/II PCR Using the *Pol* Primer Pair

There are many parameters that influence whether a PCR works or not. One of the first parameters that are optimised is the divalent cation. In this primer set, Mg^{2+} was the divalent cation that was used. The concentration did have an effect as can be seen in Fig. 2.3.

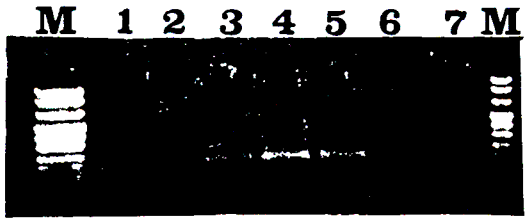


Fig. 2.3: Titration of the Mg^{2+} concentration using the *pol* primer pair (SK110/111). The Mg^{2+} concentration was titrated as follows: 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM, and 4 mM (lanes 1 - 6 respectively). Lane 7 was a negative control. The PCR products were detected on a 2% EtBr stained agarose gel, using marker VIII (Boehringer Mannheim). The optimum Mg^{2+} concentration observed was 3 mM (lane 4).

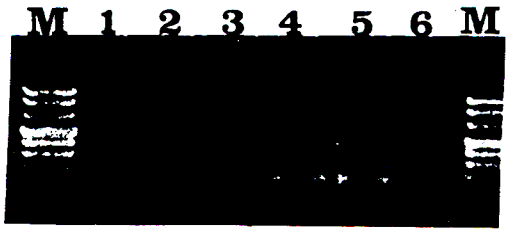


Fig. 2.4: Titration of the primer concentration of the *pol* primer pair (SK110/111). The primers were titrated as follows: 5 pmol, 10 pmol, 20 pmol, 50 pmol and 100 pmol (lanes 1 - 5 respectively). Lane 6 was a negative control. The PCR products were detected on a 2% EtBr stained agarose gel, using marker VIII (Boehringer Mannheim). The optimum primer concentration observed was 20 pmol (lane 3).

The optimum Mg^{2+} concentration for the *pol* primer pair was taken as 3 mM using 20 pmol of primer (Fig. 2.3). The optimum primer concentration using the optimum Mg^{2+} concentration was determined to be 20 pmol (Fig 2.4). The optimised reaction (50 ul total volume) for the *pol* primer pair can be summarised as follows: 5 ul of 10 X reaction buffer with Mg^{2+} (Boehringer Mannheim), 5 ul (250 uM) of dNTPs (Pharmacia), 1 ul (20 pmol) each primer, one unit of *Taq* polymerase (Boehringer Mannheim) and 3 mM final Mg^{2+} concentration. This PCR reaction mix was tested on a total of 37 patients, using the WB result as the "Gold" standard. The test detected 16 out of 16 WB positive patients, giving a sensitivity of 100%. No false positive were detected (23/23 WB negatives), showing the test to be highly specific. An example of a PCR result can be seen in Fig. 2.5. Lanes one, two, three, and five were WB positive samples (65, 66, 67 and 68 respectively). Lane four was a positive control (PCR and WB positive), and lane six was a negative control (no DNA). The gel was blotted and probed (Fig. 2.6)

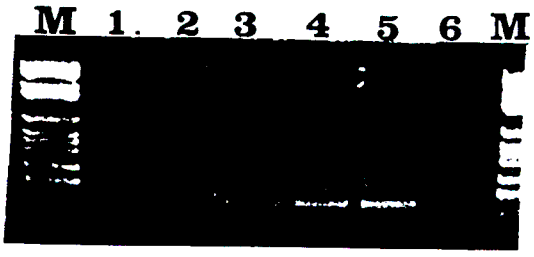


Fig. 2.5: PCR results using the *pol* primer pair. Lanes 1, 2, 3, and 5 were WB positive samples (65, 66, 67 and 68 respectively). Lane 4 was a positive control (PCR and WB positive), and lane 6 was a negative control (no DNA).The PCR products were detected on a 2% EtBr stained agarose gel, using marker VI (Boehringer Mannheim).

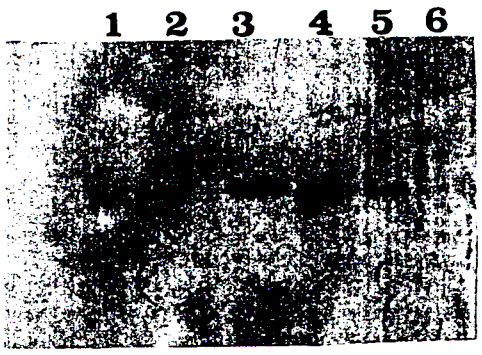


Fig. 2.6: Southern blotting of *pol* PCR products. Lanes 1, 2, 3 and 5 were PCR positive samples (65, 66, 67 and 68 respectively). Lane 4 was a positive control (sample 58), and lane 6 was a negative control (no DNA).

2.3.2 Optimisation of the PCR Conditions for the HTLV-I/II *Env* Primer Pair

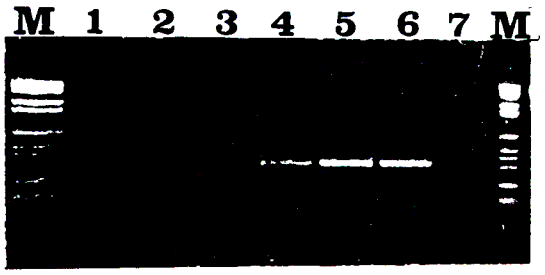


Fig. 2.7: Titration of the Mg^{2+} concentration of the *env* primer pair (6275 and 6666). The Mg^{2+} concentration was titrated as follows: 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM, and 4 mM (lanes 1 - 6 respectively). Lane 7 was a negative control. The PCR products were detected on a 2% EtBr stained agarose gel, using marker VI (Boehringer Mannheim). The optimum Mg^{2+} concentration observed was 3.5 mM (lane 5).

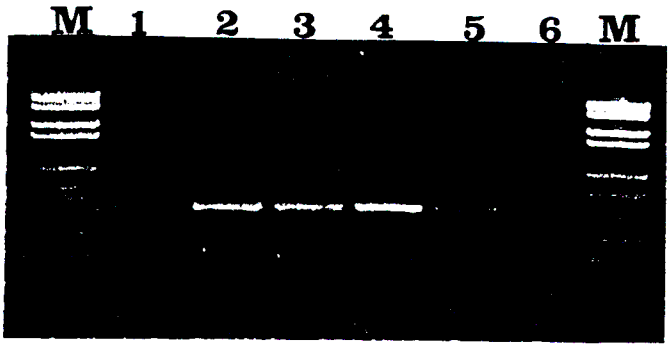


Fig. 2.8: Titration of the primer concentration of the *env* primer pair (6275-6666). The primers were titrated as follows: 40 pmol, 60 pmol, 80 pmol, 100 pmol and 120 pmol (lanes 1 - 5 respectively). Lane 6 was a negative control. The PCR products were detected on a 2% EtBr stained agarose gel, using marker VI (Boehringer Mannheim). The optimum primer concentration observed was 100 pmol (lane 4).

The results of the Mg^{2+} concentration titration can be seen in Fig. 2.7. The optimum concentration was 3.5 mM of Mg^{2+}. For the primer titration, 100 pmol was taken as the optimum (Fig. 2.8). The PCR product of patient 52 was cloned and sequenced (as described in chapter 3) to confirm that the amplified product was HTLV-I. The sequence showed a 98.5% similar to the J02029 Japanese prototype strain (Appendix II). Although the reaction was optimised, it did not prove as sensitive as the *pol* primer pair. It was therefore not used for the HTLV PCR assay.

2.3.3 Acugene AmpliSensor Assay

The sensitivity and specificity of the AmpliSensor assay was assessed on 14 HTLV-I positive samples, 19 negative samples and two samples without DNA. (Samples 70 and 71 were not included in the assay). The AmpliSensor assay did not detect one of the WB

and *pol* positive patients (patient 57), resulting in a sensitivity of 93%. There were no false positives (100% specificity). The standard titration was important in the quantitative aspect of the AmpliSensor assay (Fig. 2.9). Quantitation was performed on 13 positive samples, two negative samples and two samples without DNA. Samples 23, 70 and 71 were not included in the quantitative assay. The quantitation results of the samples are represented in Fig. 2.10, and the raw data can be seen in Appendix III. A comparison of the Acugene AmpliSensor assay with the PCR and WB results can be seen in Table 2.3.

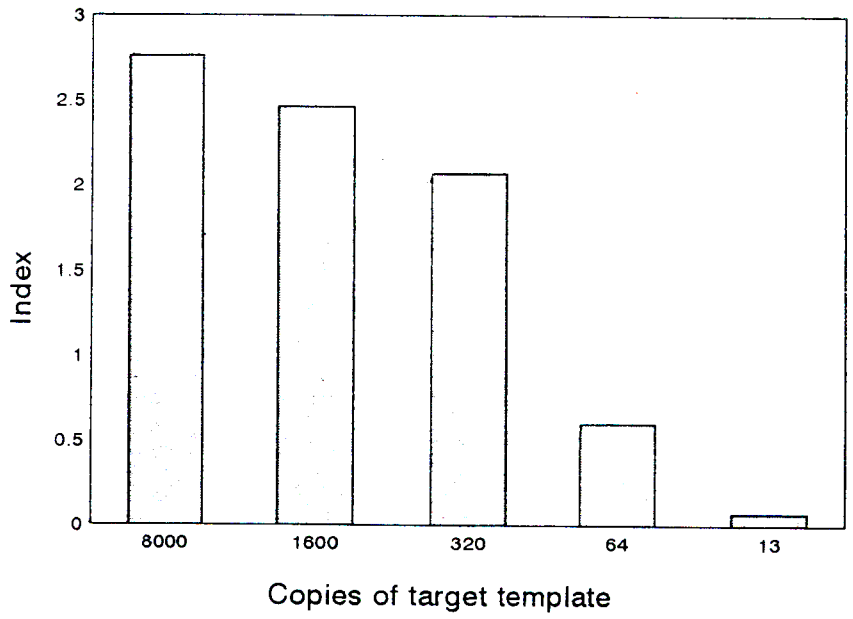


Fig. 2.9: Bar graph representing the titration of the standards in the AmpliSensor assay. Standard one contained 8000 copies of template, standard two contained 1600 copies, standard three contained 320 copies, standard four contained 64 copies and standard five contained 13 copies.

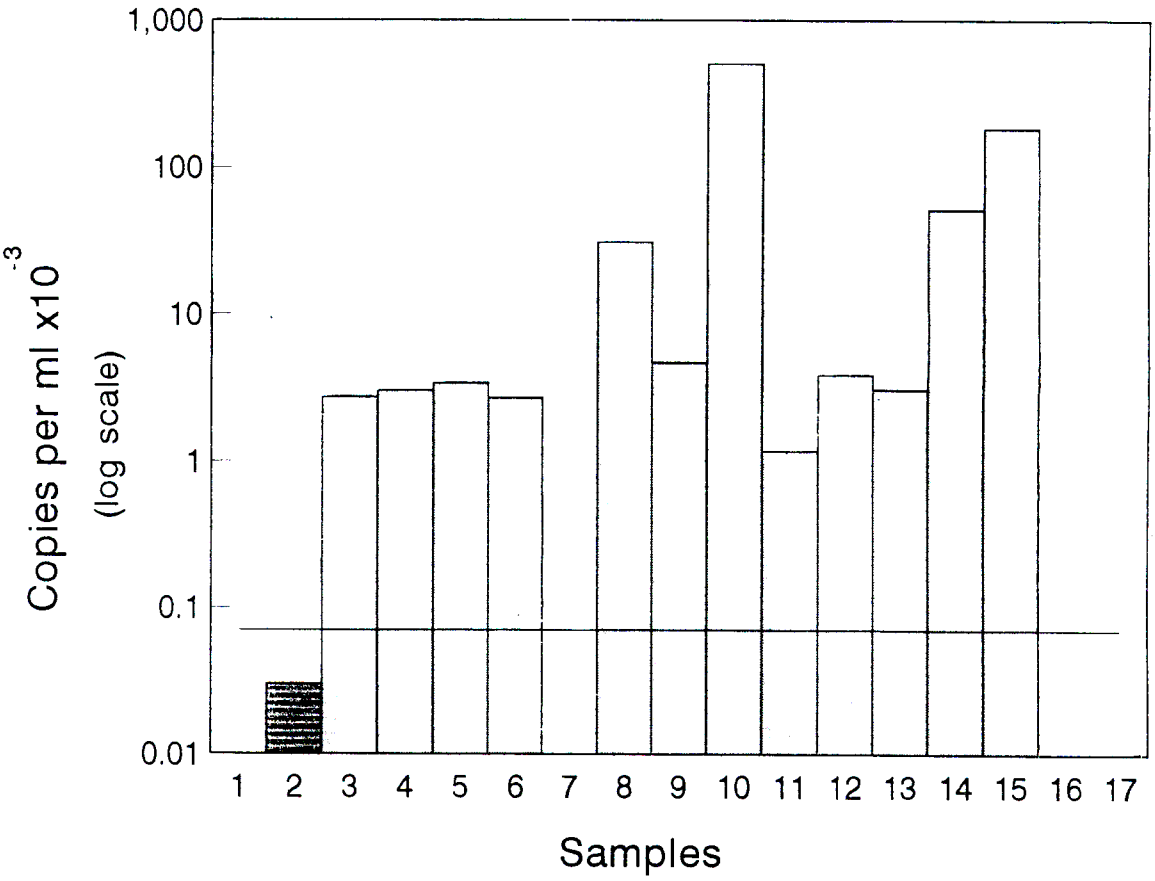


Fig. 2.10: Bar graph representing the detection of 13 positive, two negative samples, and two negative controls using the AmpliSensor assay. Lanes 1 and 2 are negative controls (no DNA) and lanes 3-17 are patient samples. The cut off is represented by the line. The negative controls were negative. The sample in lane 3 (sample 48) had 2 710 copies of viral DNA per ml of whole blood, lane 4 (sample 49): 3 000 copies, lane 5 (sample 51): 3 370 copies, lane 6 (sample 52): 2 680 copies, lane 7 (sample 57): 0 copies, lane 8 (sample 58): 30 700 copies, lane 9 (sample 61): 4 600 copies, lane 10 (sample 62): 502 000 copies, lane 11 (sample 64): 1 170 copies, lane 12 (sample 65): 3 830 copies, lane 13 (sample 66): 3 060, lane 14 (sample 67): 51 300 copies, lane 15 (sample 68): 183 000 copies, lane 16 (sample 60): 0 copies and lane 17 (sample 69): 0 copies.

2.3.4 Comparison with the WB Assay

TABLE 2.3: Comparison of PCR, Acugene AmpliSensor assay and WB results. The WB was used as the Gold standard. AmpliSensor results were not available (NA) for samples 70 and 71.

NO	WB	PCR	ACUGENE
1-19	-	-	-
23	+	+	+
48	+	+	+
49	+	+	+
50	-	-	-
51	+	+	+
52	+	+	+
57	+	+	-
58	+	+	+
59	-	-	-
60	-	-	-
61	+	+	+
62	+	+	+
64	+	+	+
65	+	+	+
66	+	+	+
67	+	+	+
68	+	+	+
69	-	-	-
70	+	+	NA
71	+	+	NA

2.4 **Discussion**

Molecular diagnosis of HTLV-I is rapidly becoming the method of choice for diagnosing HTLV-I infected patients. The PCR has an advantage over serological methods as it can detect as little as ten copies of viral DNA (Heneine *et.al.*, 1992).

The major concerns when using PCR for diagnosis are cross-contamination and non-specific priming, resulting in false positives. To avoid the possiblity of false positives, two areas

of the HTLV-I genome were targeted for the in-house diagnostic test, namely the *pol* (186 bp product) and the *env* (410 bp product) regions. Ehrlich and colleagues (1990) recommend that in diagnostic cases, at least two sets of primer pairs produce a positive result, particularly if the patient is seronegative. Optimum conditions for the amplification of the *pol* product was 20 pmol of each primer and 3 mM of Mg^{2+}. The optimised reaction mix was tested on 16 WB positve and 23 WB negative samples. The *pol* primers proved to be very sensitive (100%) and specific (100%). This is consistent with reports from Kwok *et.al.* (1988) and Vallejo and Garcia-Saiz (1995). The latter found a high correlation between WB and PCR results using the *pol* primer pair.

Amplification using primers targeting the *env* region of the genome (410 bp product) were optimised at 100 pmol of each primer and 3 mM Mg^{2+}. The PCR product cloned and sequenced. Sequencing of the clone confirmed the amplification of HTLV-I (98.5% similarity to the Japanese ATK-1 strain: J02029). The *env* PCR however, was not reproducible, and was no longer pursued as a diagnostic test. It was decided that the *pol* PCR result be used in conjunction with the AmpliSensor and WB assay results for diagnosis.

Qualitative PCR, although an excellent diagnostic tool, is limited in its inablity to predict the progression of disease. Quantitative PCR, on the other hand, is a powerful predictor of disease progression and clinical outcome (Lion, 1994). It can be used to monitor treatment efficacy anf thereby assess the need for changes in antiretroviral therapy (Lion, 1994; Schoolly, 1995). Recent trends in quantitative PCR techniques are towards "real-time" PCR. This technique does not require post-PCR handling, preventing potential PCR

product carry-over contamination (Wang *et.al.*, 1995). The AmpliSensor assay is a novel qualitative and "real-time" quantitative technique (Wang *et.al.*, 1995). The primers specifically target the *pol* region of the HTLV-I genome in a two-step "hemi-nested" reaction. Haff (1994) has reported that nested quantitative PCR provides intrinsic PCR product carry-over protection and generally improves the lower limit of detection of the PCR.

In the first stage of the AmpliSensor assay, asymmetric PCR was carried out for 23 cycles to generate excess amplified product. In the second stage, a fluorescent labelled signal duplex primer (the AmpliSensor) was added. This AmpliSensor primer was complementary to the excess amplified product, for a "hemi-nested" reaction. Further amplification resulted in the disruption of the signal duplex and a decrease in energy transfer. The decrease in energy transfer (proportionate to the target concentration) was monitored. The assay was performed on 14 WB and *pol* PCR positive samples, 23 negative DNA samples and two samples without DNA. The assay proved to be highly specific (100%). This is consistent with data showing that non-specific amplification of products that lack homology to the target duplex did not affect the assay (Wang *et.al.*, 1995). This is apparently due to the terminal intrastrand sequence complementarity of the amplicons. The amplicons form intrastrand loops that decrease their availability for hybridisation with new amplification primers. The kinetics of intrastrand hybridisation would be strongly dependent on the separation of the terminal complementary regions. In this way, intrastrand hybridisation of, for example, primer dimers, would be expected to follow fast "snap-back" kinetics, thereby inhibiting amplification of these products (Wang *et.al.*, 1995). However, the sensitivity of the AmpliSensor assay was lower then the

specificity (93%), with the assay missing one of the WB and PCR positive samples (sample 57). The sensitivity obtained for the assay may not reflect the true sensitivity due to the small sample size.

Quantitation was performed on 13 WB positive samples, four WB negative samples and six samples without DNA. Sample 62 had the highest viral load of 502 000 copies per ml of whole blood, followed by sample 68 (183 000) and 67 (51 300). This correlates well with reported viral loads of 120 000 to 2 700 000 copies per ml of whole blood in HTLV-I infected patients (Kira *et.al.*, 1991). Sample 57 was outside the quantitation range and gave a negative result.

The AmpliSensor assay has an advantage over PCR as the detection process occurs concurrently with amplification. The extent of amplification can also be monitored throughout the reaction. However, quantitation was dependant on the titration of the standards during the assay. It was imperative that four of the standards formed a linear relationship for extrapolation of sample concentrations. Overall, the quantitative AmpliSensor results correlated well with the PCR results. The faint bands that were noted in the in-house PCR corresponded to lower viral loads in those samples. The software was easy to use making it a convenient assay. The technique would also be useful for testing large numbers of samples as the possibility of cross-contamination is greatly reduced (Wang *et.al.*, 1995).

Unfortunately, the small sample size did not allow a true reflection of the sensitivity of the AmpliSensor assay. The small sample size was unavoidable as the number of HAM/TSP patients presenting to the Department of Neurology (Wentworth Hospital,

Durban), which is the major source of blood samples, are small. The assay will be re-evaluated once sufficient samples have been accumulated.

The in-house HTLV-I/II PCR is now used routinely for the diagnosis of HTLV-I and has replaced the WB as the confirming assay. A drug study is also underway where the effect of therapy in HTLV-I viraemia will be monitored using the AmpliSensor assay.

Chapter 3

Partial Sequencing of the Envelope Gene of Five Local HTLV-I Isolates

3.1 Introduction

The first HTLV-I provirus was cloned in 1980 (Miyoshi *et.al.*, 1980). The viral DNA was isolated from an MT-1 cell line derived from the leukaemic cells of a Japanese ATL patient that was designated ATM-I. The LTR region of this clone was sequenced by Seiki *et.al.*, in 1982. Seiki and colleagues (1983) were also the first to completely sequence an integrated provirus, ATK, from a Japanese patient's leukaemic cells. Malik and colleagues, (1988), were the first to completley sequence a proviral HTLV-I clone of non-Japanese origin. This proviral clone, called HS-35, was isolated from a British patient of Caribbean origin. When compared to the isolate of Japanese origin, only a 2.3% divergence was noted, indicating strong sequence conservation. However, Malik *et.al.* (1988) also demonstrated that isolates originating from similar geographical areas only diverged by approximately 1%. The influence of geographical area on sequence diversity was further demonstrated by the discovery of HTLV-I in remote Melanesian populations in Papua New Guinea and the Solomon Islands (Yanagihara *et.al.*, 1991) and among aboriginals in Australia (Bastian *et.al.*, 1993). Unlike HTLV-I strains from Japan, the Caribbean, the Americas and Africa, which differ by less than three percent among themselves (Seiki *et.al.*, 1983; Malik *et.al.*, 1988; Gessain *et.al.*, 1992), the Australo-Melanesian strains differ from the others by approximately 7% (Gessain *et.al.*, 1991; Bastian *et.al.*, 1993). Yanagihara *et.al.* (1994) therefore proposed the classification of HTLV-I into two major subtypes or topotypes: the Australo-Melanesian topotype and the cosmopolitan topotype. Vandamme *et.al.* (1994) and Lui *et.al.* (1996) demonstrated that the Central African strains were distinct from the cosmopolitan strains (5% divergence) and classified them as a separate major subtype. More recently, Mahieux *et.al.*, (1997) have discovered a new subtype ("D") isolated from pygmies in Central Africa.

HTLV-I is endemic in KZN (Bhigjee *et.al.*, 1993; 1994), but limited sequence data is available for the local subtype. In this chapter, the amplification, cloning and sequencing of a 1.5 kb region of the *env* gene of five HTLV-I positive patients from KZN is described. The sequences obtained were then compared with *env* sequences from strains representative of other major subtypes.

3.2 **Methods**

3.2.1 Subjects

Four HAM/TSP patients, who presented at Wentworth Hospital, and one patient who presented at the skin clinic at King Edward VIII Hospital, Durban, were the source of DNA used to isolate the HTLV-I sequences. These patients were local residents and chosen to represent the local variants of HTLV-I.

3.2.2 Primer Design

The ultimate aim, which is broader than this project, is to amplify the entire HTLV-I genome isolated from a locally infected patient. The first objective to this end was to determine the most representative "local strain" by sequencing the *env* region of five local patients. Once the most representative isolate was determined, the DNA from this patient would be used to amplify the entire HTLV-I genome. The primers were designed to overlap and cover the entire genome (Fig 3.1). The primer sequences and dilutions can be seen in Appendix II.

The primers targeting the *env* region were from nucleotide 5146 to 5166 (forward) (5' CAT GCC CAA GAC CCG TCG GAG 3') and from nucleotide 6661 to 6681 (reverse) (5' AGG GGC TGA GAA GGC GAA AGA 3'). An internal primer pair was also

designed for a "nested" PCR reaction: 5166 to 5186 (5' GGC CCC GCC GAT CCC AAA GAA 3') and 6638 to 6658 (5' ACT CAG GTT TTA TAA GAG AGT 3'). The design of the *env* primers are shown in more detail in Fig. 3.2.

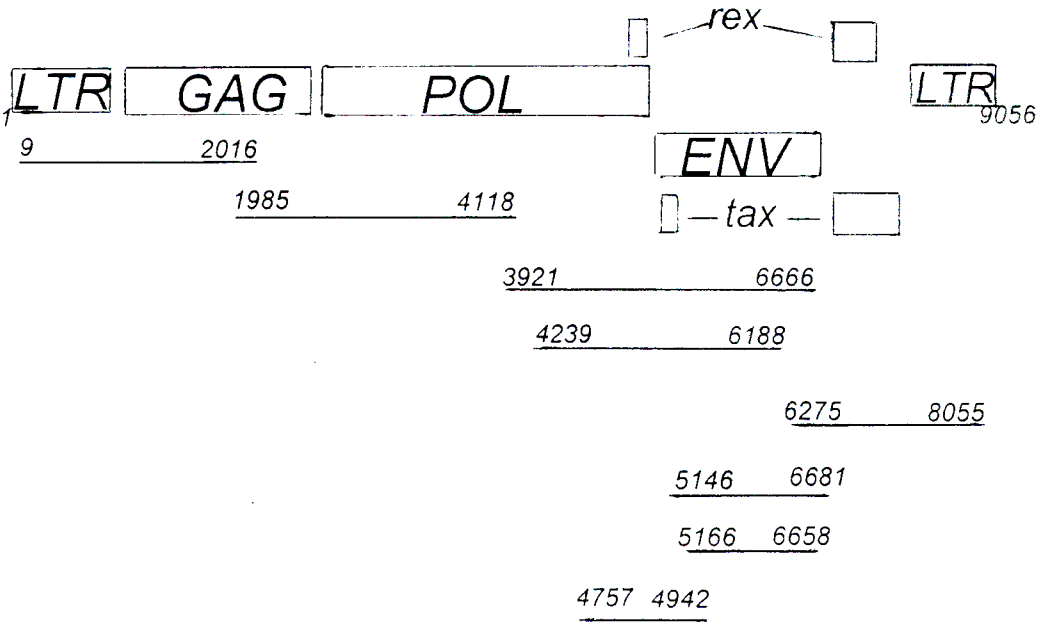


Fig. 3.1: The primers were designed using the Oligo IV software programme (NBI). Primers were designed so that expected amplicons would overlap and cover the entire genome. The primer pairs are 9 and 2016, 1985 and 4118, 4239 and 6188, 3921 and 6666, 5146 and 6681, 5166 and 6658, 6275 and 8055.

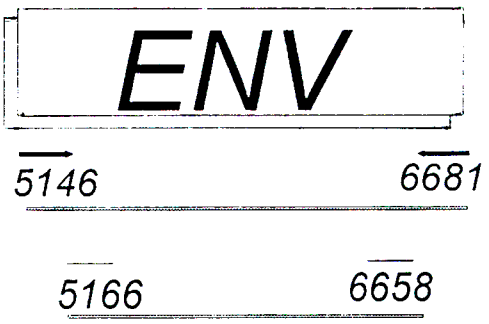


Fig. 3.2: Env primer design. Primers were designed to amplify a 1535 bp product of the *env* genome (5146 and 6681). Internal primers were designed for a "nested" PCR (5166 and 6681).

3.2.3 Primer Dilution

Primers were diluted according to the equation described in chapter 2 to have a final concentration of 20 pmol/ul.

For 5146:

OD = 22.4; length = 20

Therefore dilution = 1 ul + 4.09 ul water

For 6681:

OD = 27.5; length = 20

Therefore dilution = 1 ul + 5.25 ul water

For 5166:

OD = 22; length = 20

Therefore dilution = 1 ul + 4 ul water

For 6658:

OD = 22.9; length = 20

Therefore dilution = 1 ul + 4.2 ul water

The diluted primers were stored at 4°C. Undiluted primers were stored at -70°C.

3.2.4 Optimisation of the PCR Reaction

3.2.4.1 Reagents

Boehringer Mannheim 10 X buffer: (Appendix I)

dNTPs: (Appendix I)

Taq: (Appendix I)

Mg^{2+}: (Appendix I)

3.2.4.2 Procedure

Optimisation of the external primer pair was addressed first. The Mg^{2+} concentration was titrated using 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM and 4 mM Mg^{2+}. The concentrations of the other reagents were: 10 ul of Boehringer Mannheim 10 X buffer (with 1.5 mM Mg^{2+}), 250 uM dNTPs (Pharmacia), two units of *Taq* (Boehringer Mannheim) and 20 pmol of each primer in a total volume of 100 ul. The primer concentration was titrated using 5 pmol, 10 pmol, 20 pmol, 50 pmol and 100 pmol of each primer, and the optimum Mg^{2+} concentration.

The PCR was performed in a Perkin Elmer 2400 thermocycler under the following conditions: 94°C for five minutes; 40 cycles of 94°C for one minute, 60°C for 30 seconds, 72°C for 30 seconds; 72°C for seven minutes; and held at 4°C.

3.2.4.3 Nested PCR

Three ul of the PCR product of the first step was used in the second step of the reaction. The Mg^{2+} concentration was titrated using 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM and 4mM Mg^{2+}. The concentrations of the other reagents were: 5 ul of Boehringer Mannheim 10 X buffer, 250 uM dNTPs (Pharmacia), one unit of *Taq* (Boehringer Mannheim) and 20 pmol of primers in a total volume of 50 ul. The primer concentration was titrated using 5 pmol, 10 pmol, 20 pmol, 50 pmol and 100 pmol of each primer, at the optimum Mg^{2+} concentration (of 2 mM).

The PCR was performed in a Perkin Elmer 2400 thermocycler under the same conditions described for the first step of the reaction. PCR products were visualised on a 1% agarose gel stained with ethidium bromide. The method for making up the gel is described in chapter 2. The image was printed or stored on a computer disc as described in chapter 2.

PCR products were stored at 4°C (temporarily), or at -20°C (long term).

3.2.5 Cloning of PCR Products

The pMOS*Blue* T-vector cloning kit (Amersham) was used to clone all PCR products. This system exploits the template-independent activity of *Taq* polymerases which preferentially add a single adenosine nucleotide to the 3' end of double stranded DNA (Clark, 1988). The PCR products are ligated into a compatible thymidine tailed vector (Marchuk, 1990), for which the pMOS*Blue* T-vector has specifically been constructed.

3.2.5.1 Reagents

Components of the pMOS*Blue* T-vector cloning kit:

pMOS*Blue* T-vector: 50 ng/ul

MOS*Blue* competent cells

SOC media: (Appendix I)

T4 DNA ligase: 4 U/ul

10 x ligation buffer

100 mM DTT

10 mM ATP

Nuclease free water

Additional Reagents:

24:1 chloroform-isoamylalcohol mix: 1 ul of isoamylalcohol was added to 24 ul of chloroform and vortexed.

AMP: (Appendix I)

LB-AMP agar plates: 20 agar tablets (Sigma) were weighed out and brought up to 40 g with agar (Sigma). The agar was brought up to one litre with water and autoclaved. When the agar had cooled to approximately 60°C, 750 ul of AMP (100 mg/ml) was added. Approximately 30 ml of agar was poured per plate. Plates were stored at 4°C for up to two weeks, after which they were discarded.

IPTG (Isopropylthio-B-D-galactoside): 2 g of IPTG was dissolved in 8 ml of distilled water. The volume was adjusted to 10 ml with distilled water, and filtered through a 0.22 um filter. The IPTG was dispensed in 1 ml aliquots and stored at -20°C.

X-gal (5-Bromo-4-chloro-3-indolyl-B-D-galactosidase): 20 mg of X-gal was dissolved in 1 ml of water in a polypropylene tube. The tube was wrapped in foil to prevent damage by light, and stored at -20°C.

3.2.5.2 Purification of PCR product

The PCR product was purified to allow size selection and removal of DNA polymerase. Briefly, equal volumes of the PCR product and a 24:1 chloroform-isoamylalcohol mix were vortexed for one minute then centrifuged at 12500 x g for one minute. Two layers were obtained, with the upper aqueous layer containing the cleaned DNA.

3.2.5.3 Ligation

For optimal cloning efficiency, the vector to insert ratio should range from 1:5 to 1:10. This is calculated according to the following equation:

$$\frac{\text{size insert (bp)} \times \text{amt vector (ng)}}{\text{size vector (bp)}} \times \frac{\text{insert}}{\text{vector}} = \text{ng}$$

Reagents were used directly from the pmos*Blue* cloning kit. Ligation reagents were thawed on ice before use. The reaction mix for each sample contained 1 ul of 10 X ligation buffer, 0.5 ul of 100 mM DTT, 0.5 ul of 10 mM ATP, 0.5 ul of T4 DNA ligase, 1 ul of pMOS*Blue* T-vector, and 4.5 ul nuclease free water. A master mix was made for the required number of samples. Eight ul of the master mix was added to 2 ul of the cleaned product and incubated overnight at 16°C.

3.2.5.4 Transformation

The required number of vials of MOS*Blue* competent cells were thawed on ice. Twenty ul of competent cells were pipetted into pre-chilled polypropylene tubes. Two ul of the ligated product was added and allowed to stand on ice for 30 minutes. The cells were heat-shocked at 42°C for 40 seconds, and placed on ice for a further two minutes. Eighty ul of warmed SOC solution was added to the tubes and shaken for one hour at 37°C at 250 rpm. LB-AMP plates were warmed at 37°C, and overlayed with 20 ul of 100 mM IPTG and 35 ul of 50 mg/ml X-gal. Plates were left to soak for at least 30 minutes prior to plating. Seventy ul of each transformation were spread onto the plates. The plates were inverted and incubated overnight at 37°C.

3.2.6 Screening of Clones

The pMOS*Blue* vector allows for blue-white screening, with recombinant colonies appearing white when plated on X-gal and IPTG indicator plates.

White colonies were spotted onto agar plates that were divided and numbered. Colonies were screened either by PCR, or by plasmid minipreps.

3.2.6.1 PCR

3.2.6.1.1 Reagents

Boehringer Mannheim 10 X buffer: (Appendix I)

dNTPs: (Appendix I)

Taq: (Appendix I)

Mg^{2+}: (Appendix I)

3.2.6.1.2 Procedure

Sterile plastic tips were used to transfer colony samples to PCR tubes containing 50 ul sterile distilled water. The tubes were boiled for ten minutes at 99°C, and centrifuged for three minutes at 12500 x g. Ten ul of the supernatant was used in the PCR reaction. PCR reaction mixes were prepared with the primer pairs corresponding to the clones that were being screened. The reaction mix was prepared as follows: 5 ul of 10 X buffer (Boehringer Mannheim), 250 uM dNTPs (Pharmacia), one unit of *Taq* (Boehringer Mannheim), 2 mM Mg^{2+}, 10 pmol of primer, made up to a total volume of 40 ul with filtered distilled water. To facilitate screening of large numbers of clones, master mixes for 40 reactions were prepared and stored at -20°C if not used immediately.

3.2.6.2 Alkaline Lysis

3.2.6.2.1 Reagents

LB broth: 10 g of tryptone, 5 g of yeast and 10 g of NaCl were dissolved in 800 ml of distilled water. The pH was adjusted to 7.0 with NaOH (5 M). The volume was adjusted to one litre with water and autoclaved.

AMP: (Appendix I)

Solution I: 5 ml of 1 M glucose (Appendix I), 2.5 ml of 1 M tris HCl pH 8.0 (Appendix I) and 2 ml of 0.5 M EDTA (Appendix I) was diluted to 100 ml with water. (Final concentrations: 50 mM glucose; 25 mM tris HCl (pH 8); 10 mM EDTA.)

Solution II: 20 ul of 1 N NaOH (Appendix I) and 100 ul of 10% SDS (Appendix I) were diluted with 880 ul of water. (Final concentrations: 0.2 N NaOH and 1% SDS.)

Solution III: 60 ml of 5 M potassium acetate (Appendix I) and 11.5 ml of glacial acetic acid were diluted with 28.5 ml of water to give a 3 M solution.

TE buffer: (Appendix I)

RNAse A (10 mg/ml): 10 mg of RNAse A (Boehringer Mannheim) were dissolved in 800 ul of distilled water. The volume was adjusted to 1 ml with distilled water.

RNAse A (0.02 mg/ml): 2 ul of RNAse A (10 mg/ml) was diluted to 1 ml with TE.

3.2.6.2.2 Procedure

Each colony was transferred to a green cap tube (15 ml) containing 3 ml of LB broth and 3 ul of 50 ng/ul AMP. The tubes were incubated overnight at 37°C on a rotary shaker at 250 rpm. One-and-a-half ml of each sample was centrifuged for three minutes at 12500 x g. The supernatant was removed and the sample centrifuged again, until all the supernatant was removed. One-hundred ul of solution I was added to the tubes and vortexed. Two-hundred ul of freshly prepared solution II was added to each tube and rapidly inverted five times. The tubes were placed on ice and cold solution III (150 ul) was added to each tube, vortexed, inverted ten times and stored on ice for five minutes. The tubes were centrifuged for five minutes and the supernatants transferred to fresh tubes. Four-hundred ul of 24:1:1 phenol:chloroform:isoamylalcohol was added, vortexed for 15 seconds and centrifuged for five minutes at 12500 x g. Each aqueous phase was

transferred to a fresh tube. Absolute alcohol (800 ul) was added at room temperature, vortexed and allowed to settle for two minutes. The samples were centrifuged for five minutes and the supernatants removed. The pellets were washed with one ml of 70% ethanol, allowed to air-dry at room temperature (approximately ten minutes), then dissolved in 25 ul tris EDTA (TE) buffer (pH 8) containing RNAse A (0.02 mg/ml).

Restriction digests were carried out on the plasmid DNA products to confirm the presence, and size of the inserts. Two restriction enzymes were used for the digest, *Bam*HI (Boehringer Mannheim) and *Hin*dIII (Boehringer Mannheim). Their recognition sequences are: G/GATCC (*Bam*HI) and A/AGCTT (*Hin*dIII).

The reaction mixture was prepared as follows: 3 ul of DNA product, 1 ul of *Bam*H I, 1 ul of *Hin*d III, 1 ul of buffer B (Boehringer Mannheim) and 1 ul of water.

The digest was incubated for one hour at 37°C. Samples were analysed on a 1% agarose gel. The gel was prepared as described in 2.2.7.1.

3.2.7 DNA Extraction from Clones Using the QIAgen Kit

3.2.7.1 Reagents

All solutions were stored at room temperature unless otherwise stated.

P1 (Resuspension buffer): 50 mM tris HCl, 10 mM EDTA, pH 8.0. Lyophilised RNAse A was added to P1 after opening to give a final concentration of 100 ug/ml. The solution was stored at 4°C.

P2 (Lysis buffer): 200 mM NaOH, 1% SDS.

P3 (Neutralisation buffer): 3 M KAc, pH 5.5. The solution was stored at 4°C.

TE Buffer: (Appendix I)

Buffer QBT (Equilibration buffer): 750 mM NaCl, 50 mM MOPS, 15% ethanol, pH 7.0, 0.15% triton X-100.

Buffer QC (Wash buffer): 1 M NaCl, 50 mM MOPS, 15% ethanol, pH 8.5.

Buffer QF (Elution buffer): 1.25 M NaCl, 50 mM tris HCl, 15% ethanol, pH 8.0. ,

70% ethanol: (Appendix I)

3.2.7.2 Procedure

The extraction protocol was as stated in the manufacturers instructions. Briefly, clones were grown overnight, on an Innova 4000 incubator shaker (New Brunswick Scientific) (250 rpm), at 37°C in 100 ml LB broth containing 100 ng/ul AMP. The cultures were centrifuged for 30 minutes using a Beckman J21 centrifuge and JA14 rotor at 10000 rpm. The supernatants were discarded, and the pellets resuspended in 10 ml of buffer P1 containing RNase A. Ten ml of buffer P2 was added to each tube, gently mixed, and incubated for five minutes at room temperature. Buffer P3 was added and mixed immediately. The contents of each tube was transfered to a QIAfilter Maxi cartridge. These were left for ten minutes at room temperature. The lysates were filtered into a pre-equilibrated QIAgen-tip 500, and allowed to enter the resin by gravity flow. The QIAgen-tips were washed twice with 30 ml of Buffer QC. Fifteen ml of Buffer QF was used to elute the DNA from the resin. The DNA was precipitated at room temperature with 10.5 ml isopropanol, and centrifuged immediately for 30 minutes at 4°C at 10000 rpm. The supernatants were removed and 5 ml of 70% ethanol was added to each tube. The tubes were centrifuged for ten minutes at 10000 rpm. The supernatants were removed and the pelltes were air-dryed for five minutes. Each proviral DNA sample was resuspended in a suitable volume of TE buffer.

3.2.8 Cycle Sequencing

The Thermosequenase Cy5 labelled primer Cycle Sequencing Kit (Pharmacia Biotech) was used to sequence the clones.

3.2.8.1 Kit Components

A Mix

C Mix

G Mix

T Mix

Loading buffer: 100% deionised formamide and DextranBlue 200 (5 mg/ml).

3.2.8.2 Additional Reagents

M13 reverse primer: TTAGCCTA (Cy5 labelled)

T7 promoter primer: TATGGATT (Cy5 labelled)

3.2.8.3 Procedure

A reaction master mix was prepared as follows: 2 ug of the plasmid DNA, 2.1 pmol of cy5' labelled primer (M13 or T7), made up to 22 ul with filtered distilled water. This was vortexed and microfuged. Two ul of nucleotide mix was added to appropriately labelled tubes (A, C, G, T) and kept on ice.

Five ul of master mix was added to each nucleotide, and thouroughly mixed (taking care not to create any bubbles).

Tubes were placed in a Progene (Techne) thermocycler and amplified according to the following cycling conditions: 96°C for five minutes, 30 cylces of 96°C for one minute,

60°C for 30 seconds 70°C for one minute and a final extention of 70°C for five minutes. This was held at 4°C. Four ul of loading buffer was added to each tube. At this point, samples could be temporarily stored at -20°C until sequenced.

3.2.9 Gel Preparation

The gel solution was prepared as follows: 9 ml of 10 X TBE, 6 ml of 50% Longranger (FMC, USA) and 22.8 g of urea (A.L.F. quality), made up to 60 ml with filtered distilled water. This was stirred on a magnetic stirrer for 15 minutes. The solution was filtered, using a 0.45 um filter into a polyethylene bottle suitable for pouring the gel.

3.2.10 Assembling the Sequencing Plate

The sequencing plate consists of two glass plates one of which is attached to the upper buffer chamber and separated by two glass spacers (approximately 300 mm x 10 mm x 1 mm). Before assembling the sequencing plate, both glass plates were thouroughly cleaned, initially using distilled water followed by 70% ethanol, taking care to remove all visible dust. The plate was assembled (with the spacers on either end) and secured with clamps (four on either side). A comb with 40 wells was inserted at the top of the two plates.

3.2.11 Pouring the Sequencing Gel

3.2.11.1 Reagents

10% ammonium persulphate: 10 g of ammonium persulphate was dissolved in 100 ml of distilled water. The solution was aliquoted in 2 ml eppendorfs and stored at -20°C.

3.2.11.2 Procedure

Thirty ul of TEMED (BioRad, USA) and 300 ul of 10% ammonium persulphate (APS) was added to the prepared gel solution and gently mixed. The gel was poured by placing the nozzle of the bottle between the plates, and with gentle squeezing, moving the nozzle back and fourth along the plates. All bubbles were removed using a "bubble catcher" made of thin plastic (approximately 80 cm x 5 cm). The gel was allowed to polymerise for approximately two hours.

3.2.12 Sequencing Plate Preparation

The top of the assembled plate was cleaned briefly. Special care was taken to remove all acrylamide and urea from the spacers where they were exposed.

3.2.13 Preparation of the (A.L.F.) Express DNA Sequencer (Pharmacia, Biotech)

3.2.13.1 Reagents

0.5 X TBE: 200 ml of 5 X TBE (Appendix I) was diluted to two litres with distilled water.

3.2.13.2 Procedure

The assembled sequencing plate was placed in the lower buffer chamber and attached to the sequencer. Both buffer chambers were filled with 0.5 X TBE. The water circulation was connected and laser activated. The A.L.F. Express (Pharmacia, Biotech) is linked to a computer terminal where the data is captured and processed using a specific software package (A.L.F. Manager, Hitachi). A new file was opened in this package, and named

appropriately. The experimental conditions were selected (56°C, 25 W, 1500 V, two second sampling interval, 900 minute running time), and the laser value was monitored. When the temperature reached 50°C, the comb was removed. The wells were washed thoroughly (approximately 10 times each) with buffer (using a 20 ml syringe) to remove any acrylamide or urea present in the wells.

3.2.14 Loading the Sequencing Gel

Samples were denatured at 96°C for three minutes and placed on ice. Four ul of each sample was loaded into the appropriate wells (for example, a tube containing Adenosine for the first plasmid was loaded into the lane marked A of number one). Four ul of loading buffer was added into the smaller marker wells on either end of the gel.

3.2.15 Starting the Run

The electrodes were connected to the sequencer and the lid was closed. If the temperature and the laser values were stable, the run was started.

3.2.16 ABI Prism Dye Terminator Cycle Sequencing

3.2.16.1 Primer Design

A second set of primers was used to amplify an internal portion of the HTLV-I *env* gene. These primers targeted position 5669 (forward): AAT GCG GTT TTC CCT TCT CC 3' and position 6294 (reverse): 5' TTT TGA GTA GAT TTT TGT GG 3'. The PCR primers used to amplify the 1535 kb *env* region were also used for cycle sequencing of the clones obtained for patient 49 (the primer targeting nucleotide 5146) and patient 62 (the primer targeting nucleotide 6681).

Terminator ready reaction mix: A, C, G and T nucleosides (each labelled with a specific dye) and AmpliTaq DNA polymerase, FS.

3.2.16.4 Procedure

For each reaction, the following reagents were mixed in an appropriately labelled tube: 6 ul of the terminator ready reaction mix, 1 ul of template DNA (1 ug), 2 ul of primer (5 pmol/ul) and 6 ul of filtered distilled water.

The reaction mixture was amplified using a Progene (Techne) thermocycler under the following conditions:

95°C for one minute, 30 cycles of 95°C for 30 seconds, 40°C for 30 seconds and 60°C for 3 minutes. The samples were held overnight at 4°C.

3.2.17 Purifying the Extension Products

3.2.17.1 Reagents

70% Ethanol with 0.5 mM Mg^{2+}: 50 ul of 1 M Mg^{2+} (Sigma) was added to 80 ml of absolute ethanol (Merck) and made up to 100 ml with distilled water.

Template suppression buffer (Perkin Elmer)

3.2.17.2 Procedure

The extension products were purified according to the protocol described in the manufacturers instructions. Briefly, the amplicons were transfered to 1.5 ml tubes. Fifty-five ul of 70% ethanol with 0.5 mM Mg^{2+} was added to the sequencing reaction and vortexed briefly. The samples were left at room temperature for ten minutes (to precipitate the extension products), after which they were centrifuged at 12500 x g for 15

minutes, also at room temperature. The supernatant was completely removed using a micro pipet. Each pellet was dried in a vacuum centrifuge and resuspended in 18 ul of template suppression buffer. The extension products were denatured for two minutes in a boiling water bath and quenched immediately on ice for one minute. Samples were centrifuged briefly, and transfered to an autosampler tube with a septum. Samples could be stored at 4°C at this stage for at least a week before analysis.

Loading of samples on the ABI Prism genetic analyser was performed by staff at the S.A. Sugar Research Institute.

3.2.18 Compilation of Sequences

Sequences that were obtained from the A.L.F. Express Sequencer and the ABI Prism sequencer were aligned using DNASIS v7.00 (Hitachi Software Engineering).

3.2.19 Sequence Analysis

Sequences were compared with the cosmopolitan strains from North America (CH and SP: Paine *et.al.*, 1991), Japan (HTVPROP: Seiki *et.al.*, 1983, H5: Tsujimoto *et.al.*, 1988 and HTLV1ENV: Astier-Gin 1991), the Caribbean (HTVPRCAR: Malik *et.al.*, 1988), the Central African strain (EL: Paine *et.al.*, 1991), the Melanesian strain (MEL1: Gessain *et.al.*, 1993), a STLV-I strain (STLENV: Watanabe *et.al.* 1985), and an HTLV-II strain (HL2IIENV: Lee *et.al.*, 1993).

3.3 **Results**

3.3.1 PCR

3.3.1.1 Optimisation of the Single Step Reaction

The external primers (5146 and 6681) were optimised using DNA extracted from patient 61 (Fig. 3.4 and Fig. 3.5). The optimum Mg^{2+} concentration was 2.5 mM (lane three, Fig. 3.4), and the optimum primer concentration was 10 pmol of primers (lane two, Fig. 3.5).

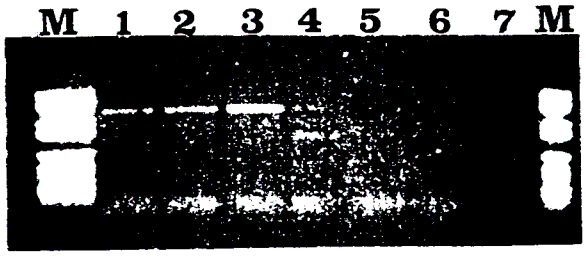


Fig. 3.4: Titration of the Mg^{2+} concentration of the external primer pair (position 5146 - 6681). The Mg^{2+} concentration was titrated as follows: 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM and 4 mM (lanes 1 - 6 respectively). Lane 7 was a negative control. The PCR products were detected on a 1% EtBr stained agarose gel, using marker VI (Boehringer Mannheim). The optimum Mg^{2+} concentration observed was 2.5 mM (lane 3). A further increase in Mg^{2+} concentration resulted in non-specific amplification.

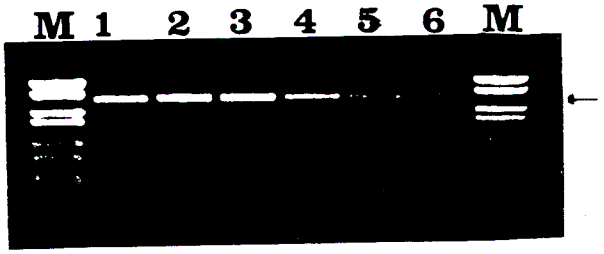


Fig. 3.5: Titration of the primer concentration of the external primer pair (position 5146 - 6681). The primers were titrated as follows: 5 pmol, 10 pmol, 20 pmol, 50 pmol and 100 pmol (lanes 1 - 5 respectively). Lane 6 was a negative control. The PCR products were detected on a 1% EtBr stained agarose gel, using marker VI (Boehringer Mannheim). The optimum primer concentration observed was 10 pmol (lane 2). Non-specific amplification, resulting in smearing was observed for primer concentrations above 50 pmol. ← indicates expected size target 1535 bp.

3.3.1.2 Optimisation of the "Nested" Reaction

Four of the five patients were amplified in a single PCR (49, 61, 62, 68). The fifth patient (48) was amplified using the internal primer pair (5166 and 6658) in a "nested" PCR. The optimum Mg^{2+} and primer concentrations were 2 mM and 10 pmol, respectively (lane two, Fig 3.6 and lane two, Fig 3.7).

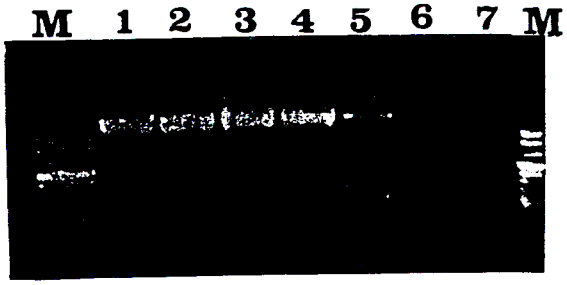


Fig. 3.6: Titration of the Mg^{2+} concentration of the internal primer pair (5166 - 6658). The Mg^{2+} concentration was titrated as follows: 1.5 mM, 2 mM, 2.5 mM, 3 mM, 3.5 mM and 4 mM (lanes 1 - 6 respectively). Lane 7 was a negative control. The PCR products were detected on a 1% EtBr stained agarose gel, using marker VIII (Boehringer Mannheim). The optimum Mg^{2+} concentration observed was 2 mM (lane 2).

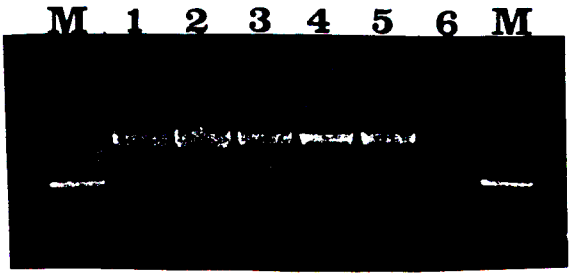


Fig. 3.7: Titration of primer concentration of the internal primer pair (position 5166 - 6658). The primers were titrated as follows: 5 pmol, 10 pmol, 20 pmol, 50 pmol and 100 pmol (lanes 1 - 5 respectively). Lane 6 was a negative control. The PCR products were detected on a 1% EtBr stained agarose gel, using marker VIII (Boehringer Mannheim). The optimum primer concentration was taken as 10 pmol (lane 2).

3.3.2 Screening of Clones

3.3.2.1 Screening by PCR

Clones were screened by PCR using the appropriate primer pair (Fig 3.8).

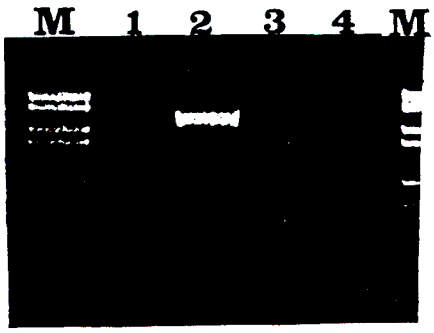


Fig. 3.8: Screening of clones by PCR. On the above gel, white (potentially positive) colonies of clones isolated from patient 61 were screened using primers 5146 and 6681. PCR products were detected on a 1% EtBr stained agarose gel using marker VI (Boehringer Mannheim). Amplification of the 1535 bp product was taken as confirmation of an insert.

3.3.2.2 Screening by Alkaline Lysis

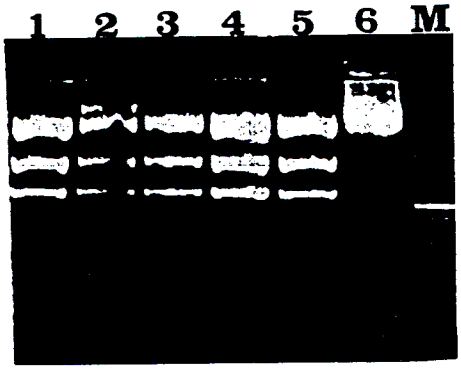


Fig. 3.9: Screening of clones by Alkaline Lysis. The five clones were confirmed by extracting the plasmids using an alkaline lysis method, followed by restriction enzyme digest and analysis on EtBr stained gels.

3.3.3 Sequencing Results

Sequence reads of approximately 600 bp long were obtained from the A.L.F. Express Automated Sequencer (Biotech/Pharmacia). Sequences from both the normal and complementary strands were used. Reads of approximately 500 bp long were obtained with the ABI sequencer for both the forward and complementary strands. The fluorograms of the sequences can be seen in Appendix V. The nucleotide sequences are displayed in Fig. 3.10, using CH as a reference clone.

```
        5203
#CH     ATG GGT AAG TTT CTC GCC ACT TTG ATT TTA TTC TTC CAG TTC TGC CCC CTC
#48     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62     ... ... ... ... ... ... ... ... .C. ... ... ... ... ... ... ... ...
#68     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5     ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ...
#HTVPROP ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP     ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ...
#EL     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1   ... ... ... ... ... A.. ... ... ... ... ... C.. ... ... ... ... .CT
#STLENV1 ... ... ... ... ..T A.. ..C ..A C.. ... ... C.. ..A C.. C.. ... .C.
#HL2IIENV ... ... ..C G.. T.. TT. CTA C.T T.A ..C AGT C.. ACA CA. .T. ..A .CA
        5254
#CH     ATC CTC GGT GAT TAC AGC CCC AGC TGC TGT ACT CTC ACA ATT GGA GTC TCC
#48     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR ... ... ... ... ... ... ... ... ... ... ... ... ... G.. ... ... ...
#H5     ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL     ..T ... A.. ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1   ..T ... A.. TC. ... ..T ... ... ... ... ... ... ..T ... ..G ... ...
#STLENV1 ..T G.. T.. ... C.. ..T ... ... ... ..C ... ... ... G.C ... ... ...
#HL2IIENV G.. .AG CAG AGC CGA T.. A.A CT. ACG GT. GG. A.T T.C TCC TAC CA. ...
        5305
#CH     TCA TAC CAC TCT AAA CCC TGC AAT CCT GCC CAG CCA GTT TGT TCG TGG ACC
#48     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ...
#62     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ...
#H5     ... ... ... ... ... ..T ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ...
#HTLV1ENV ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1   ... ... ... ... ... ... ... ... ... ... ... ... ... ..C ..A ... ...
#STLENV1 ... ..T ..T ... ... ... ... ... ..C ... ..A ... ..C ..C ..A ... ...
#HL2IIENV AGC CC. TGT AGC CC. A.. CAA CCC GTC TG. AC. TGG AAC CTC GAC CTT .AT
        5356
#CH     CTC GAC CTG CTG GCC CTT TCA GCA GAT CAG GCC CTA CAG CCC CCC TGC CCT
#48     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62     ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#68     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ...
#H5     A.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL     ... ... ..A ... ... ... ... ... ... ..A ... ... ... ... ... ... ...
#HTLV1ENV ... ... ... ... ... ... ... ..G ... ... ... ... ... ... ... ... ...
#MEL1   ... ... ... T.. ... ... ... ... ... ... ..T ... ... ... ... ... ...
#STLENV1 ... ... ... ..A ..T ... ... ... ..C ..A ... ... ..A ... ... ... ...
#HL2IIENV TC. CTA AC. AC. .A. .AG CG. CT. C.. .CC C.. TGC .CT AA. .TA ATT A..
        5407
#CH     AAT CTA GTA GCT TAC TCT AGC TAC CAT GCC ACC TAT TCC CTA TAT CTA TTC
#48     ... ... ... .G. ... ... ... ... ... ... ... ... ... ... ... ... ...
#61     ... ... ... .G. ... ... ... ... ... ... ... ... ... ... ... ... ...
#49     ... ... ... .G. ... ... ... ... ... ... ... ... ... ... ... ... ...
#62     ... ... ... .G. ... ... ... ... ... ... ... ... ... ... ... ... ...
#68     ... ... ... .G. ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR ... ... ... AG. ... ..C ... ... ... ... ... ... ... ... ... ... ...
#H5     ..C ... ... AG. ... ..C ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP ..C ... ... AG. ... ..C ... ... ... ... ... ... ... ... ... ... ...
#SP     ... ... ... .G. ... ... ... ... T.. ... ... ... ... ... ... ... ...
#EL     ... ... ... AG. ... ..C ... ... ... ... ... ... ... ... ... ... ..T
#HTLV1ENV ... ... ... AG. ... ..C ... ... ... ... ... ... ... ... ... ... ...
#MEL1   ... ... ... AGC ... ..C .A. ... ... ... ... ... .G. ... ... ... ...
#STLENV1 ... ..G ..G .G. ... ..C .A. ... ... ..T ... ... ... ... ... ... ...
#HL2IIENV T.C TCT .GC TTC C.. AAA .CT ..T TCC TTA TA. .TA .T. .C. C.. TGG A.A
        5458
#CH     CCT CAT TGG ATT AAA AAG CCA AAC CGA AAT GGC GGA GGC TAT TAT TCA GCC
#48     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP ... ... ... .C. ..G ... ... ... ... ... ... ... ... ... ... ... ...
#SP     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1   ... T.. ... ... ... ... ... ... ... ... ... ... ... ..C ... ..G ...
#STLENV1 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..G ...
#HL2IIENV AAG A.G CCA .A. .G. C.. GGC CTA G.. T.C TA. TCG CC. .CC ... AAT .A.
```

Fig. 3.10: Comparison of the nucleotide sequence of the five KZN strains (48, 49, 61, 62 and 68) with selected strains. The North American strain (CH) was used as a reference for alignment (Paine *et.al.*, 1991). Other strains include the Caribbean strain (HTVPRCAR), the Japanese strains (H5, HTVPROP and HTLV1ENV), another North American strain (SP), the Central African strain (EL), the Melanesian strain (MEL1), a STLV-I strain (STLVENV1) and a HTLV-II strain (HL2IIENV).

```
       5509
#CH        TCT TAT TCA GAC CCT TGT TCC TTA AAG TGC CCA TAC CTA GGG TGC CAA TCA
#48        ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#EL        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ...
#MEL1      C.. ... ... ... ... ... ... C.. ... ... ... ..C ... ... ... ...
#STLENV1   ... ... ... ... ... ... ... C.. ... ... ..T ... ... ..A ... ... ...
#HL2IIENV C.. .GC ..G CTA .AA ..C C.. .AC TTA G.. TGC C.A TC. T.. ACA TGC C..
       5560
#CH        TGG ACC TGC CCC TAT ACA GGG GCC GTC TCC AGC CCC TAC TGG AAG TTT CAG
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ..A ...
#H5        ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ... ..A
#SP        ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ..A ...
#HTLV1ENV  ... ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#STLENV1   ... ... ..T ... ..C ... ..A ... ... ... ... ... ..T ... ..A .A. ...
#HL2IIENV .AC ..G G.. ... GTC T.C A.T C.A TC. .GG .AG TTT C.. .CA G.T G.A A.T
       5611
#CH        CAA GAT GTC AAT TTT ACT CAA GAA GTT TCA CGC CTC AAT ATT AAT CTC CAT
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... .A. ... ... ... ... ...
#H5        ..C ... ... ... ... ... ... ... A.. ... ... ... ... ... ... ... ...
#HTVPROP   ..C ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ... ... ... ... ..G ... ..C ... ... ... ... ... ... ...
#STLENV1   ... ... ... ... ..C ... ... ... ..C ... ... ... .GC C.. ... ... ..C
#HL2IIENV TTC ACC CAA G.A G.C .GC ... .TG TCC CTT ..A ..A C.C T.C TC. AAG TGC
       5662
#CH        TTT TCG AAA TGC GGT TTT CCC TTC TCC CTT CTA GTC GAC GCT CCA GGA TAT
#48        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#HTVPRCAR  ... ..A ... ... ... ... T.. ... ... ... ... ... ... ... ... ... ...
#H5        ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ..A ... ... ... ..C ... ... ... ... ... ... ..T ..A ..C ... ...
#STLENV1   ... ..A ... ... ... ..C T.. ... ... ... ..G ... ..T ..A .T. ... ..C
#HL2IIENV GGC ..C TCC ATG ACC C.. .TA G.A GAT GCC .CT .GA T.T .A. ..T TT. .GG
       5713
#CH        GAC CCC ATC TGG TTC CTT AAT ACC GAA CCC AGC CAA CTG CCT CCC ACC ACC
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ..T ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#EL        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G..
#MEL1      ... ... ... ... C.. ... ... ... ... ... ... ... ... ..C ... ..T G..
#STLENV1   ... ... ... ... ... ... ... ..C ... ... ..T ... ... ... ..C ... ..T G..
#HL2IIENV TT. AT. .C. .CA GAA .CC .C. CAG CCT ... CCA ACT GCC T.C ... ..T GGT
       5764
#CH        CCT CCT CTA CTC CCC CAC TCT AAC CTA GAC CAC ATC CTC GAG CCC TCT ATA
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... T.. ... ... ... ... ... ... ..T ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ..T ... ... ..C ... T.. ... ... ... ..T ... ... ... ...
#STLENV1   .T. ... ... ... ... ..T ..C ... T.G ... ... ... ... ..T ..C ...
#HL2IIENV ..A TGA ..C .GA ..T TGA A.A CGT .CT A.. .C. C.. .A. .TC TTG GAC .AC
```

Fig. 3.10: (cont.)

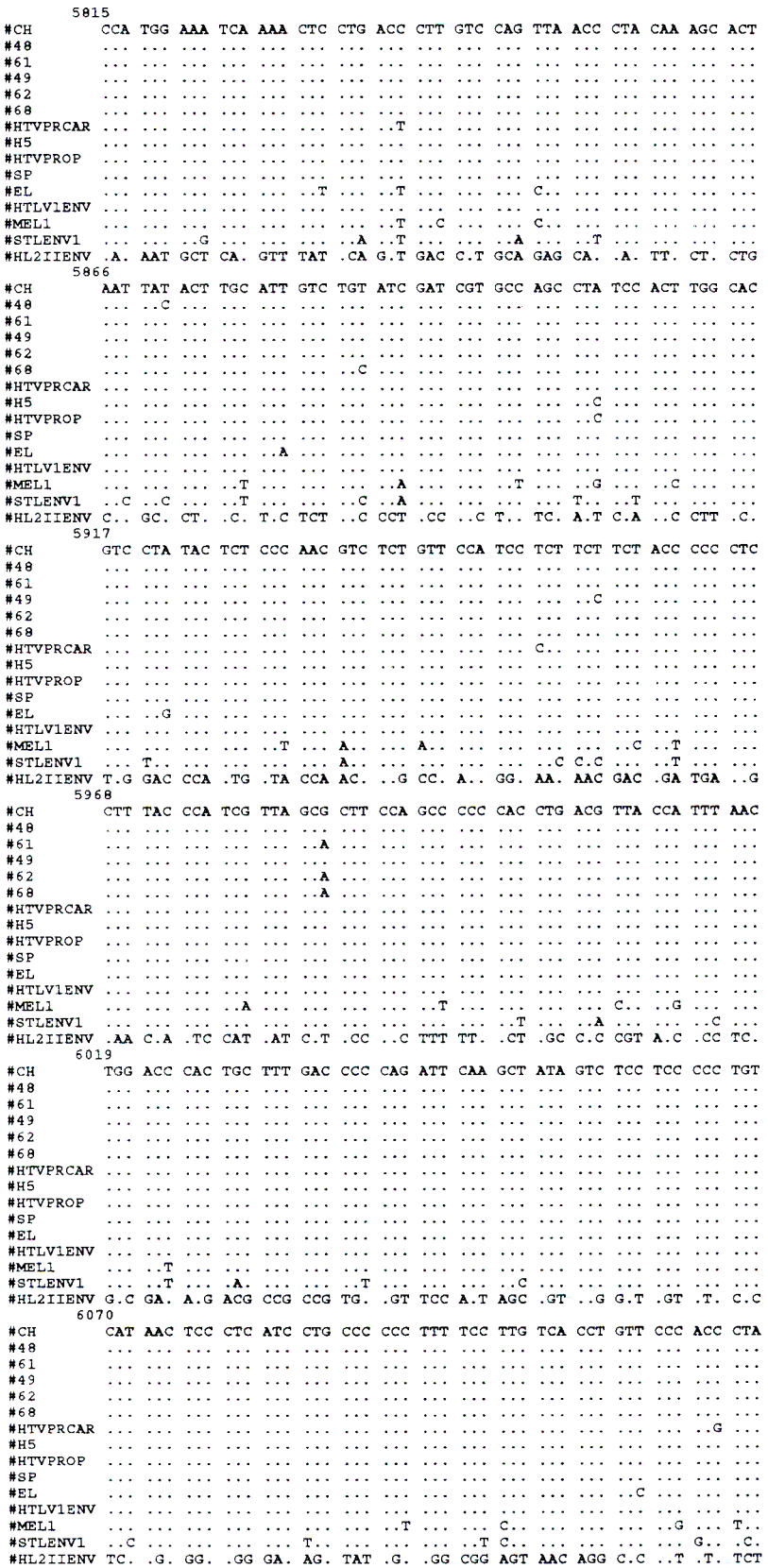

5815
#CH CCA TGG AAA TCA AAA CTC CTG ACC CTT GTC CAG TTA ACC CTA CAA AGC ACT
#48 ... ...
#61 ... ...
#49 ... ...
#62 ... ...
#68 ... ...
#HTVPRCAR ... ...T ... ...
#H5 ... ...
#HTVPROP ... ...
#SP ... ...
#EL ... ...T ... ..T ... ... C.. ... ...
#HTLV1ENV ... ...
#MEL1 ... ...T ..C ... ... C.. ... ...
#STLENV1 ... .G ... ...A ..T ... .A ... ..T ... ...
#HL2IIENV .A. AAT GCT CA. GTT TAT .CA G.T GAC C.T GCA GAG CA. .A. TT. CT. CTG
5866
#CH AAT TAT ACT TGC ATT GTC TGT ATC GAT CGT GCC AGC CTA TCC ACT TGG CAC
#48 ... ..C ... ...
#61 ... ...
#49 ... ...
#62 ... ...
#68 ... ...C ... ...
#HTVPRCAR ... ...
#H5 ... ...C ... ...
#HTVPROP ... ...C ... ...
#SP ... ...
#EL ... ...A ... ...
#HTLV1ENV ... ...
#MEL1 ... ...T ... ...A ... .T ... ..G ... ..C ... ...
#STLENV1 ..C ..C ... ..T ... .C ..A ... ... T.. ..T ... ...
#HL2IIENV C.. GC. CT. .C. T.C TCT ..C CCT .CC ..C T.. TC. A.T C.A ..C CTT .C.
5917
#CH GTC CTA TAC TCT CCC AAC GTC TCT GTT CCA TCC TCT TCT TCT ACC CCC CTC
#48 ... ...
#61 ... ...
#49 ... ...C ... ...
#62 ... ...
#68 ... ...
#HTVPRCAR ... ... C.. ... ...
#H5 ... ...
#HTVPROP ... ...
#SP ... ...
#EL ... ..G ... ...
#HTLV1ENV ... ...
#MEL1 ... ...T ... A.. ... A.. ... ...C ..T ... ...
#STLENV1 ... T.. ... ... A.. ... ...C C.C ... ..T ... ...
#HL2IIENV T.G GAC CCA .TG .TA CCA AC. ..G CC. A.. GG. AA. AAC GAC .GA TGA ..G
5968
#CH CTT TAC CCA TCG TTA GCG CTT CCA GCC CCC CAC CTG ACG TTA CCA TTT AAC
#48 ... ...
#61 ... ...A ... ...
#49 ... ...
#62 ... ...A ... ...
#68 ... ...A ... ...
#HTVPRCAR ... ...
#H5 ... ...
#HTVPROP ... ...
#SP ... ...
#EL ... ...
#HTLV1ENV ... ...
#MEL1 ... ...A ... ...T ... ... C.. ..G ... ...
#STLENV1 ... ...T ... ..A ... .C ...
#HL2IIENV .AA C.A .TC CAT .AT C.T .CC ..C TTT TT. .CT .GC C.C CGT A.C .CC TC.
6019
#CH TGG ACC CAC TGC TTT GAC CCC CAG ATT CAA GCT ATA GTC TCC TCC CCC TGT
#48 ... ...
#61 ... ...
#49 ... ...
#62 ... ...
#68 ... ...
#HTVPRCAR ... ...
#H5 ... ...
#HTVPROP ... ...
#SP ... ...
#EL ... ...
#HTLV1ENV ... ...
#MEL1 ... ..T ... ...
#STLENV1 ... ..T ... .A. ... .T ... ...C ... ...
#HL2IIENV G.C GA. A.G ACG CCG CCG TG. .GT TCC A.T AGC .GT ..G G.T .GT .T. C.C
6070
#CH CAT AAC TCC CTC ATC CTG CCC CCC TTT TCC TTG TCA CCT GTT CCC ACC CTA
#48 ... ...
#61 ... ...
#49 ... ...
#62 ... ...
#68 ... ...
#HTVPRCAR ... ...G ...
#H5 ... ...
#HTVPROP ... ...
#SP ... ...
#EL ... ...C ... ...
#HTLV1ENV ... ...
#MEL1 ... ...T ... ... C.. ... ...G ... T..
#STLENV1 ..C ... ... T.. ... ...T C.. ... ... G.. .C.
#HL2IIENV TC. .G. GG. .GG GA. AG. TAT .G. .GG CGG AGT AAC AGG C.C ..T .T. TCT

Fig. 3.10: (cont.)

```
       6021
#CH        GGA TCC CGC TCC CGC CGA GCG GTA CCG GTG GCG GTC TGG CTT GTC TCC GCC
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ..A ... ... ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ..T ... ... ... ..A ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      CA. ... ..T ... ... ... ... ... ..A ... ... ... ... ..A ... ... ...
#STLENV1   ..G ... ..T ... ... ..G ..A ... ... ... ... ... ..C ... ..T ...
#HL2IIENV A.C .T. .AG .AA AAG .CT T.T C.T .GA .GT TGA CAA A.A TA. C.. C.A C.T
       6172
#CH        CTG GCC ATA GGA GCC GGA GTG GCT GGC GGG ATT ACC GGC TCC ATG TCC CTC
#48        ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ..G ... ... ... ... ... ... A.. ... ... ... ... ... ... ...
#H5        ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ..G ... ... ... A.. ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ..G ... ... ..G ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ... A.. ... A.T ... ... ... ... ..T ... ... ... ... ...
#STLENV1   ... ... ..G ... A.. ... ..T ... ... ... ... ... ... ... ... ... ...
#HL2IIENV TAC C.A GGC CAT AGT CA. AAA T.A TCA AAA CA. C.T CCG GGT TGC A.A A.A
       6223
#CH        GCC TCA GGA AAG AGC CTC CTA CAT GAG GTG GAC AAA GAT ATT TCC CAA TTA
#48        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... .G. ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..G ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..G ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... .G. ... ... ... ... ... ..A ... ... ... ... ... ... ...
#HTLV1ENV  C.. ... ... ... ... ... T.. ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ... ..T ... ... ... ... ..A ... ..G ... ... ... ... ...
#STLENV1   ... ... ... ..A ..T ... ... ... ..A ..A ... ... ... ... ... ... ..G
#HL2IIENV TG. AGC CC. G.A TAG ACG AGG AT. AGA CCT CCT .TT CTG GGA A.A AGG GGG
       6274
#CH        ACT CAA GCA ATA GTC AAA AAC CAC AAA AAT CTA CTC AAA ATT GCG CAG TAT
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ..A ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ... ... ... ... ..T ... ... ... ... ... G.. ..A ..A ...
#STLENV1   ..C ... ... ..T ... ... ... ... ... ... ... ... ..G ... ..A ..A ...
#HL2IIENV TT. GTG CA. .GC CAT .C. GGA GCA .TG TTG ..T .CT C.. TA. CA. T.A C.C
       6325
#CH        GCT GCC CAG AAC AGA CGA GGC CTT GAT CTC CTG TTC TGG GAG CAA GGA GGA
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ..G ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ...
#STLENV1   ... ... ..A ... ... ... ... ... ... ... ... ..T ... ... ... ... ...
#HL2IIENV T.A TGT ATC CGT CCT .C. A.A ACG .CC .C. TCT .GA AAA .C. TGT CAT CAC
       6376
#CH        TTA TGC AAA GCA TTA CAA GAA CAG TGC TGT TTT CTA AAT ATT ACT AAT TCC
#48        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... C.. ... .CG ... ... ..C ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
#EL        ... ... ... ... ... ..G ... ... ... ... ... ... ... ... ..C ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ...
#MEL1      ... ... ..G ..G C.. ... ... ... ... ... ... ... ..C ..C ... ... ...
#STLENV1   C.. ..T ... ... C.G ... ... ..A ..T ... ..C ... ... ... ... ... ...
#HL2IIENV CGG .TG GGG A.T AA. .TG .G. TCT ..G .C. G.C .C. GTG GGC ..G .GA AG.
```

Fig. 3.10: (cont.)

```
       6427
#CH        CAT GTC TCA ATG CTA CAA GAA AGA CCC CCC CTT G-A GAA TCG AGT CCT GAC
#48        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ..A ... ... ..G ... ... ... ... ... A.. ... ... ... ...
#H5        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... C.. ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... C..
#HTLV1ENV  ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ..C ..T ... ..A ... ... ... ... ..A ... ... .G. ... ... ... ... ...
#STLENV1   ... ... ... ..A ..T ... ... ... ... ... ... ... A.. ... ... T.. A..
#HL2IIENV  .C. CCA GAC .G. .AT A.C C.T TCT .A. ..T AC. C.C TCC .T. TCA TA. TGT
       6478
#CH        TGG CTG GGG CCT TAA CTG GGA CCT TGG CCT CTC ACA GTG GGC TCG AGA GGC
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... A..
#H5        A.C ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... G.. ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... T.. ... ... ... ... ... ... ... ... A.. ... C.. ... ...
#STLENV1   ... T.. ... ... C.. ... ... ... ... ... ... G.. A.. ... C.. ... ...
#HL2IIENV  .T. GCC CCT G.A .CC TCC .CC AAA .CC AAG .C. TTC CGC A.. GGT TAC AAA
       6529
#CH        CTT ACA AAC TGG AAT CAC CCT TGT CGC GCT ACT CCT TCT TGT TAT CCT TGC
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... T.. ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        .C. ... ... ... ... ... ... ... ... T.. ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      .C. G.. ... ... C.. ... ... ... ... T.. A.. ... ... ... ... ... ...
#STLENV1   ... ... ... ... ... ... ... ... ... ... AT. ... ... ... G.. C.. ... ..T
#HL2IIENV  ACC GAC .TA GCC .G. ATG ..C .TA TCA A.C .AG AGA C.A ..C ... AA. A.A
       6580
#CH        AGG ACC ATG CAT CCT CCG TCA GCT ACG ACA CCT CCC CTC GCG CGT CAG ATA
#48        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPROP   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... T.. ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      ... ... ... ... --- --- --- --- --- --- --- --- --- --- --- --- ---
#STLENV1   ... G.. ... ... ... T.. C.. A.. ... ... A.. ... ... ... T.. ... ...
#HL2IIENV  CCC G.T .GC TTC TGC AGC AA. T.C C.A TGG TTC AT. .C. CTA .CA TT. .CC
       6631
#CH        CCC CCA TTA CTC TCT TAT AAA CCC TGA GTC ATC CCT GTA AAC CAA GCA CAC
#48        ... ... ... ... C.. ... ... ... ... ... ... ... ... ... ... ... .G.
#61        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#62        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#68        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... .G.
#HTVPROP   ... ... ... ... ... ... ... ... A.. ... ... ... ... ... ... ... .G.
#SP        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#HTLV1ENV  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
#MEL1      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
#STLENV1   ... ... C.. ... ... ... ... ... C.. ... C.. ... A.. ... .T. ... GTT
#HL2IIENV  .A- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
       6682
#CH        AAT TAT TGC AAC CAC ATC GCC TCC
#48        ... ... ... ... ... ... ... ...
#61        ... C.. ... ... ... ... ... ...
#49        ... ... ... ... ... ... ... ...
#62        ... C.. ... ... ... ... ... ...
#68        ... C.. ... ... ... ... ... ...
#HTVPRCAR  ... ... ... ... ... ... ... ...
#H5        ... ... ... ... ... ... ... .G.
#HTVPROP   ... ... ... ... ... ... ... ...
#SP        ... ... ... ... ... ... ... ...
#EL        ... ... ... ... ... ... A.. ...
#HTLV1ENV  --- --- --- --- --- --- --- ---
#MEL1      --- --- --- --- --- --- --- ---
#STLENV1   .C. GCG AC. .CA TTG C.T ..A G..
#HL2IIENV  --- --- --- --- --- --- --- ---
```

Fig. 3.10: (cont.)

```
                          *          **               *      *            60
#CH        MGKFLATLIL FFQFCPLILG DYSPSCCTLT IGVSSYHSKP CNPAQPVCSW TLDLLALSAD
#48        .......... .......... .......... .......... .......... ..........
#61        .......... .......... .......... .......... .......... ..........
#49        .......... .......... .......... .......... .......... ..........
#62        ........T. .......... .......... .......... .......... ..........
#68        .......... .......... .......... .......... .......... ..........
#HTVPRCAR  .......... .......... .......... V......... .......... ..........
#H5        .......... ........F. .......... .......... .......... .I........
#HTVPROP   .......... ........F. .......... .......... .......... ..........
#SP        .......... .......... .......... .......... .......... ..........
#EL        .......... .........S .......... .......... .......... ..........
#HTLV1ENV  .......... .......... .......... .......... .......... ..........
#MEL1      .....T.... .L....P..S S......... .......... .......... ..........
#STLENV1   .....T..L. .L.LR.P.VC .H........ V......... .......... ..........
#HL2IIENV  ..NVFFL.LF SLTHF.PVQQ SRCTLTVGIS SYH..PC.PT QPVCTWNLDL NSLTTDQRLH

           *                                                  *    *    *  120
#CH        QALQPPCPNL VAYSSYHATY SLYLFPHWIK KPNRNGGGYY SASYSDPCSL KCPYLGCQSW
#48        .......... .G........ .......... .......... .......... ..........
#61        .......... .G........ .......... .......... .......... ..........
#49        .......... .G........ .......... .......... .......... ..........
#62        .......... .G........ .......... .......... .......... ..........
#68        .......... .G........ .......... .......... .......... ..........
#HTVPRCAR  .......... .S........ .......... .......... .......... ..........
#H5        .......... .S........ .......... .......... .......... ..........
#HTVPROP   .......... .S........ ........T. .......... .......... ..........
#SP        .......... .G....Y... .......... .......... .......... ..........
#EL        .......... .S........ .......... .......... .......... ..........
#HTLV1ENV  .......... .S........ .......... .......... .......... ..........
#MEL1      .......... .S..N..... C.....Y... .......... ..P....... ..........
#STLENV1   .......... .G..N..... .......... .......... .......... ..........
#HL2IIENV  PPCPNLITYS GFHKT.SLYL FPHWIKKPNR QGLGYYSPS. NDPC.LQ.PY LGCQSWTCPY

            *                                    *                        180
#CH        TCPYTGAVSS PYWKFQQDVN FTQEVSRLNI NLHFSKCGFP FSLLVDAPGY DPIWFLNTEP
#48        .......... .......... .......... .......... .......... ..........
#61        .......... .......... .......... .......... .......... ..........
#49        .......... .......... .......... .......... .......... ..........
#62        .......... .......... .......... .......... .......... ..........
#68        .......... .......... .......... .......... .......... ..........
#HTVPRCAR  .......... .......... ......H... .........S .......... ..........
#H5        .......... ......H... ...K...... .......... .......... ..........
#HTVPROP   .......... ......H... .......... .......... .......... ..........
#SP        .......... .......... .......... .......... .......... ..........
#EL        .......... .......... .......... .......... .......... ..........
#HTLV1ENV  .......... .......... .......... .......... .......... ..........
#MEL1      .......... .......... .......... .......... .......... ....L.....
#STLENV1   .......... ....Y..... .........SL .........S .......L.. ..........
#HL2IIENV  .G.VSSPSWK FHSDVNFTQE VS.VSL..HF SKCG.SMTLL VDAPGYD.LW FITSEPTQP.

                                                      *  *                240
#CH        SQLPPTTPPL LPHSNLDHIL EPSIPWKSKL LTLVQLTLQS TNYTCIVCID RASLSTWHVL
#48        ......A... .......... .......... .......... .......... ..........
#61        .......... .......... .......... .......... .......... ..........
#49        ......A... .......... .......... .......... .......... ..........
#62        .......... .......... .......... .......... .......... ..........
#68        .......... .......... .......... .......... .......... ..........
#HTVPRCAR  ......A... .S........ .......... .......... .......... ..........
#H5        ......A... .......... .......... .......... .......... ..........
#HTVPROP   ......A... .......... .......... .......... .......... ..........
#SP        ......A... .......... .......... .......... .......... ..........
#EL        ......A... .......... .......... .......... .......... ..........
#HTLV1ENV  ......A... .......... .......... .......... .......... ..........
#MEL1      ......A... .......... .......... .......... .......... ..........
#STLENV1   ......AL.. .......... .......... .......... .......... ..........
#HL2IIENV  PTAS..G.*. R.*TRPNPLH VLDNQNAQVY PADLAEHQLL LHALSFS.PA .S.IP.LPLD
```

Fig. 3.11: Comparison of amino acid sequences. The five local strains (48, 49, 61, 62 and 68) and selected strains from around the world were compared to the North American strain (CH). *N*-glycosylation sites are overlined and cysteine residues are indicated by an asterisk. The surface envelope protein (SU) and the Transmembrane protein (TM) is indicated. The likely membrane spanning domain of the TM is underlined and the immunosuppressive peptide region is double underlined.

```
                                                   *            *              300
#CH        YSPNVSVPSS SSTPLLYPSL ALPAPHLTLP FNWTHCFDPQ IQAIVSSPCH NSLILPPFSL
#48        .......... .......... .......... .......... .......... ..........
#61        .......... .......... .......... .......... .......... ..........
#49        .......... .......... .......... .......... .......... ..........
#62        .......... .......... .......... .......... .......... ..........
#68        .......... .......... .......... .......... .......... ..........
#HTVPRCAR  .........P .......... .......... .......... .......... ..........
#H5        .......... .......... .......... .......... .......... ..........
#HTVPROP   .......... .......... .......... .......... .......... ..........
#SP        .......... .......... .......... .......... .......... ..........
#EL        .......... .......... .......... .......... .......... ..........
#HTLV1ENV  .......... .......... .......... .......... .......... ..........
#MEL1      ....I.I... .......... .......... .......... .......... ..........
#STLENV1   ....I..... P......... .......... .....Y.... .......... ..........
#HL2IIENV  PLLPT.PTGN NDR*.QQLHY PP.FFPRPRT SSGDKTPPCR SNSS.ACLRS SGRDRYRWRS
                   SU \/ TM                                                360
#CH        SPVPTLGSRS RRAVPVAVWL VSALAIGAGV AGGITGSMSL ASGKSLLHEV DKDISQLTQA
#48        .......... .......... .....M.... .......... .......... ..........
#61        .......... .......... .....M.... .......... .......... ..........
#49        .......... .......... .....M.... .......... .......... .R........
#62        .......... .......... .....M.... .......... .......... ..........
#68        .......... .......... .....M.... .......... .......... ..........
#HTVPRCAR  .......... .......... .....M.... ..R....... .......... ..........
#H5        .......... .......... .....M.... .......... .......... ..........
#HTVPROP   .......... .......... .....M.... .......... .......... ..........
#SP        .......... .......... .....M.... .......... .......... ..........
#EL        .......... .......... .....M...M .......... ...R...... ..........
#HTLV1ENV  .......... .......... .....M.... .......... P......... ..........
#MEL1      ......Q... .......... .......T.I .......... .......... ..........
#STLENV1   ....AP.... .......... .....M.T.. .......... .......... .......DP.
#HL2IIENV  NRL.ISSFQ* KPSLRG*QRY LPPYPGHSQK SSKHPPGCTI C.PE*TRIRP PILGTRGFVQ
                                              *     * *                    420
#CH        IVKNHKNLLK IAQYAAQNRR GLDLLFWEQG GLCKALQEQC CFLNITNSHV SMLQERPPLE
#48        .......... .......... .......... .......... .......... .I........
#61        .......... .......... .......... .......... .......... .I........
#49        .......... .......... .......... .......... .......... .I........
#62        .......... .......... .......... .......... .......... .I........
#68        .......... .......... .......... .......... .......... .I........
#HTVPRCAR  .......... .......... .......... .......... .......... .I........
#H5        .......... .......... .......... .......... .......... .I........
#HTVPROP   .......... .......... .......... .......... R.P....... PI........
#SP        .......... .......... .......... .......... .......... .I........
#EL        .......... .......... .......... .......... .......... .I........
#HTLV1ENV  .......... .......... .......... .......... .......... .I........
#MEL1      .......... V......... .......... .......... .......... .I.......G
#STLENV1   .......... .......... .......... .......... .......... .I........
#HL2IIENV  SHTGAML.PQ YQ*HSCIRPP RTAPS-KACH HRLGTKLGSW SVPVG.R.PP DRHNHSH.TP
                                                       *                   480
#CH        NRVLTGWGLN WDLGLSQWAR EALQTGITLV ALLLLVILAG PCILRQLRHL PSRVRYPHYS
#48        .......... .......... .......... .......... .......... ..........
#61        .......... .......... .......... .......... .......... ..........
#49        .......... .......... .......... .......... .......... ..........
#62        .......... .......... .......... .......... .......... ..........
#68        .......... .......... .......... .......... .......... ..........
#HTVPRCAR  .......... .......... .......... .......... .......... ..........
#H5        .....A.... .......... .......... .......... .......... ..........
#HTVPROP   .......... .......... .......... .......... .......... ..........
#SP        .......... .......... .......... .......... .......... ..........
#EL        .......... .......... .......... .......... .......... ..........
#HTLV1ENV  .......... .......... .......... .......... .......... ..........
#MEL1      ESSPDWL.S* LGPWPLTMGP RGPANWHHPC CTTPSCYPCR TMH------- ----------
#STLENV1   .......... .......... .......... ........V. ........Q. ..........
#HL2IIENV  PCHIVWPLHP PPNPSPSA.V TKPT*PVCPY QPRDHA.IDP LASAANPHGS SPYH*P----
                   TM/           501
#CH        LINPESSL*T KHTIIATTSP P
#48        .......... ..A....... .
#61        .......... .......... .
#49        .......... .......... .
#62        .......... .......... .
#68        .......... .......... .
#HTVPRCAR  .......... .......... .
#H5        .......... ..A....... A
#HTVPROP   ..K....... ..A....... .
#SP        .......... .......... .
#EL        .......... .......... .
#HTLV1ENV  .........- ---------- -
#MEL1      ---------- ---------- -
#STLENV1   .......... *QLLRPHCLQ .
#HL2IIENV  ---------- ---------- -
```

Fig. 3.11: *(cont.)*

TABLE 3.1 Sequence homology. Comparison of the sequence obtained for patient 61 with the other strains.

Strain	% Similarity
48	99.5
49	99.1
62	99.8
68	99.9
CH	99.2
EL	97.9
H5	98.6
HL2IIENV	54.7
HTLV1ENV	98.9
HTVPRCAR	97.9
HTVPROP	98.3
MEL1	92.5
SP	99.3
STLENV1	89.1

The local sequences showed a 0.1% to 0.9% diversity among themselves. They were closely related to the North American strains, CH and SP (99.3%), followed by the Japanese strains (HTLV1ENV: 98.9%, H5: 98.6% and HTVPROP: 98.3%), the Caribbean strain (HTVPRCAR: 97.9%) and the Central African strain (EL: 97.4%). The Melanesian strain differed from the local strains by 7.5%. The Melanesian strain had the highest number of random base substitution for the HTLV-I strains overall. Isolates from patients 48 and 49 had the highest number of random substitutions from the local isolates (both had three), while the isolate from patient 61 had none. Base substitutions unique to the local subtype were found at position 5694 (A to G) and position 5985 (G to A). Bases that were common only to the local strains and the North American strains were found at positions 5416, 5424 and 5667. Six base substitutions common only to the Japanese strains H5 and HTVPROP were observed (positions 5257, 5409, 5613, 5904, 6270 and 6680). Twenty-five common base substitutions were observed between the Melanesian

strain and the STLV-I strain. The most common substitutions were G to A changes, followed by C to T changes. Five base substitutions were shared by the Melanesian, Central African and STLV-I strain. An insertion was noted in the Melanesian strain at position 6461 (Gessain *et.al.*, 1993). This affected the amino acid sequence for the Melanesian strain, resulting in a stop codon 59 codons before it occurred in the other strains. Amino acid changes observed for the local isolates were: 48 (three), 49 (three), 62 (two), 61 (one) and 68 (one). One of the amino acid changes was common to all the local strains (Alanine to Glycine).

No amino acid changes were observed in the potential *N*-glycosylation sites (Kurata *et.al.*, 1989), the gp61 *env* precursor protein (position 6126-6143) or the membrane-spanning domain. Only one amino acid substitution was observed in the immunosuppressive peptide region (HTVPROP at position 6380).

3.4 **Discussion**

High prevalence of HTLV-I have been found in some parts of sub-Saharan Africa (Goubau *et.al.*, 1993). The virus has been shown to be endemic to KZN by Bhigjee and colleagues (1993 and 1994). However, very little sequence data is available for HTLV-I strains in South Africa, and KZN in particular. To this end, the *env* region of the local HTLV-I subtype was amplified, cloned and sequenced. DNA samples from five patients were sequenced in order to identify a representative "local strain".

Analysis of the local sequences have shown considerable sequence conservation among themselves (0.1% to 0.9%). This is consistent with findings for other HTLV-I strains from the same geographic region (Malik *et.al.*, 1988; Nerurkar *et.al.*, 1993; Gessain

et.al., 1993). However, the local strains did not all share the unique base substitutions found at positions 5694 and 5985. As a result, the isolates from patients 61, 62 and 68 seem to be more closely related to each other than 48 and 49. Clinical presentations of the virus did not correlate with sequence diversity, unlike the neurovirulent strains of murine leukaemic virus (Paquette *et.al.*, 1989). Patient 61 presented with skin disease, while the others presented with HAM/TSP, although the isolate from patient 49 was the most divergent. This is consistent with a report by Bazarbachi *et.al.* (1995), who did not identify sequences specifically assiciated with neurotropic strains. The nucleotide mutations appeared to be more specific for the geographical origin of the HTLV-I strain, rather than the clinical outcome (Bazarbachi *et.al.*, 1995).

The local strains were clearly associated with the cosmopolitan subtype, when compared to the prototype sequences. They were closely related to the North American strains (CH and SP), sharing base substitutions that are not found in the other strains (G bases at position 5416 and 5667, and a T at position 5424). The local strains were also closely related to the Japanese strains, with four of the five isolates sharing a G at position 5547. Position 5547 had previously been identified as a specific site for base substitution for Japanese isolates (De *et.al.*, 1991). The similarity of the American strains with the Japanese strains is consistent with findings by Malik *et.al.* (1988). Malik and colleagues (1988) found the American strains to be more closely related to the Japanese strains than to the Caribbean isolates. The suggested that the evolution of the American strains was not associcated with migration of populations between the Americas and the Caribbean. The Japanese isolates differed among themselves. The Japanese strain HTLV1ENV (Astier-Gin *et.al.*, 1991), isolated from the MT-2 cell line, did not share the six base

substitutions at positions 5257, 5409, 5613, 5904, 6270 and 6680 with the other Japanese strains (H5: Tsujimoto *et.al.*, 1988 and HTVPROP: Seiki *et.al.*, 1983). Vidal and colleagues (1994) demonstrated the existence of two distinct molecular subtypes in Japan, one closely related to the cosmopolitan subtype, the other a separate Japanese subtype.

The Melanesian strain (MEL1) showed the highest number of random base substitutions when compared to the reference strain CH. This can be attributed to the separate evolution of Australo-Melanesian strains from the rest of the HTLV-I strains (Yanagihara *et.al.*, 1994). As HTLV-I is a conserved virus, the sequence diversity between the cosmopolitan and Melanesian strains is consistent with a slow accumulation of base substitutions over a long period of time (Yanagihara *et.al.*, 1994).
Another interesting observation was that the Melanesian and Central African strains had many base substitutions in common with the STLV-I strain. Phylogenetic analysis of the LTR sequences of the Melanesian strain (Saksena *et.al.*, 1992) had suggested that the Melanesian strain diverged from the common ancestor of HTLV-I after Indonesian STLV-I and before African and cosmopolitan HTLV-I diverged from each other. The phylogenetic implications of this will be discussed further in chapter four.

The amino acid sequence analysis showed very few amino acid changes among the local strains. There were no substitutions observed in the possible *N*-glycosylation sites, immunosuppressive peptide region or the membrane spanning domain of the transmembrane envelope protein. There were also no amino acid substitutions noted for the 19 cysteine residues. This suggests strong genetic pressure to conserve the structure and function of the *env* protein (Pique *et.al.*, 1990). Perhaps the conservation is important

for maintenance of a strong tertiary structure which may be essential for envelope processing or receptor binding (Paine *et.al.* 1991). The relationship between the local strains and the North American strains can also been seen at the amino acid level. At codon 72, the KZN strains share an amino acid substitution with the North American strain SP (Alanine to Glycine). This was due to the nucleotide substitution that occurred at position 5417. The other HTLV-I strains also differed from the North American strain CH with an Alanine to Serine substitution. This was due to the added difference at nucleotide position 5416. At codon 187, local strains 61, 62 and 68 showed a Threonine amino acid in common with the North American strain CH. The other local strains (48 and 49) showed a Threonine to Alanine substitution that also occurred in the rest of the strains. This also supports the theory that non-random alteration at specific sites appear to be better tolerated (Paine *et.al.* 1991).

The high level of sequence conservation among the HTLV-I strains is in contrast to the divergent strains of HIV (McNearney *et.al.*, 1990). Possibly, the fidelity of the reverse transcriptases for HTLV-I/STLV-I strains are higher than HIV strains (Preston *et.al.*, 1990). Envelope proteins in HTLV-I strains are also smaller than HIV strains, which may result in fewer regions that will tolerate changes without a loss of function (Paine *et.al.*, 1991). Sequence diversity, assuming a constant rate of mutation, should reflect the evolution and migratory pathways of HTLV-I infection (Vidal *et.al.*, 1994). Sequencing of the entire genome of the local HTLV-I subtype will assist in understanding the dissemination of the virus to South Africa.

Chapter 4

Phylogenetic Analysis

4.1 **Introduction**

Due to the low evolutionary rate and the limited horizontal transmission of HTLV-I, phylogenetic analysis of the virus is revealing in the movements and contacts of ancient populations (Yanagihara *et.al.*, 1994; Vandamme *et.al.*, 1994). There is much evidence that suggests that HTLV-I originated from its Simian counterpart, Simian T-lymphotrophic virus type I (STLV-I), particularly the fact that HTLV-I and STLV-I cannot be separated into distinct phylogenetic lineages according to species of origin (Saksena *et.al.*, 1993; Vandamme *et.al.*, 1994). Due to the large African reservoir (Delaporte *et.al.*, 1989; Goubau *et.al.*, 1993) and because HTLV-I is largely found in Americans (North and South) of African descent, it has been proposed that HTLV-I originated in Africa (Gallo *et.al.*, 1983). Studies using the *env* gene have supported this hypothesis (Gallo *et.al.*, 1983; Gessain *et.al.*, 1992), and have suggested that western African HTLV-I migrated to the Caribbean basin and to Japan during the slave trade of the 16th century. However, the discovery of a very divergent variant in some Melanesian and Australian Aboriginals (Gessain *et.al.*, 1991; Bastian *et.al.*, 1993) and the presence of a divergent Asian STLV-I strain (PtM3, Watanabe *et.al.*, 1985), supports a possible Asian origin of HTLV-I (Saksena *et.al.*, 1992). This second hypothesis suggests that the original Asian HTLV-I spread to Africa, possibly via ancient Asian-African contacts in Madagascar. More recently, the more divergent LTR region has been used to analyse the origin of HTLV-I (Vandamme *et.al.*, 1994, Lui *et.al.*, 1996). Vandamme and colleagues (1994) propose an ancient African origin of STLV-I with three subsequent interspecies transmissions: from western African STLV-I to HTLV-I before spreading to the Caribbean islands and the rest of the world to give the Cosmopolitan subtype; from Central Africa STLV-I (CH) to become Central African (Zairian) HTLV-I; and from

African STLV-I (PHSu1) to Asian STLV-I and subsequently to Melanesian HTLV-I. Lui and colleagues (1996) also suggest that the Melanesian HTLV-I subtype arose from Asian STLV-I, but suggest that HTLV-I/STLV-I originated in Asia, not Africa. However, Lui and colleagues concede that since both Asia and Africa harbour very divergent HTLV-I/STLV-I strains, the origin of HTLV-I/STLV-I will be biased towards Asia when using divergent Asian STLV-I strains (Song *et.al.*, 1994), and towards Africa when using divergent African STLV-I strains (Vandamme *et.al.*, 1994).

In this chapter, the phylogenetic relationship between the *env* sequences of five KZN patients and selected HTLV-I strains from throughout the world, are investigated.

4.2 **Methods**

4.2.1 Virus Strains

Five cases from KZN were studied, four HAM/TSP cases (48, 49, 62, 68) and one ATL case with skin disease (61). Reference strains representative of the different subtypes were selected for the phylogenetic analysis. All strains included in the analysis are given in Table 4.1.

4.2.2 PCR and Sequencing

A 1535 bp region of the *env* gene of each patient was amplified, cloned and sequenced as described in 3.2.4 to 3.2.17.

TABLE 4.1: HTLV-I/STLV-I Strains Used in the Phylogenetic Analysis

Strain	Geographical Origin	Accession no.	Reference
CH	N.America	M67490	Paine *et.al.*, 1991
EL	Zaire	M67514	Paine *et.al.*, 1991
GP68 (MT-2)	Japan	M23823	Takeuchi *et.al.*, 1985
H5	Japan	M37299	Tsujimoto *et.al.*, 1988
HA (HUT-102)	USA	M61048	De *et.al.*, 1991
HB (JA-1)	Japan	M61049	"
HC (JA-2)	Japan	M61050	"
HE (AF-1)	W.Africa	M61052	"
HG (AF-3)	W.Africa	M61054	"
HJ (AF-6)	W.Africa	M61057	"
HK (MT-2)	Japan	M61058	"
HL (ATK)	Japan	M61059	"
HL12091	Martinique	L76069	Baylot *et.al.*, 1996
HL1RHK30	Japan	L03561	Zhao *et.al.*, 1993
HL1RHK34	Japan	L03562	"
HL2IIENV	USA	L20734	Lee *et.al.*, 1993
HM (US-1)	USA	M64263	De *et.al.*, 1991
HTLV1ENV (MT-2)	Japan	X56949	Astier-Gin *et.al.*, 1991
HTVPRCAR	Caribbean	D13784	Malik *et.al.*, 1988
HTVPROP (ATK-1)	Japan	J02029	Seiki *et.al.*, 1983
HL12055	French Guiana	L76056	Baylot *et.al.*, 1996
MEL1	Papua New Guinea	L02533	Gessain *et.al.*, 1993
SP	N. America	M67514	Paine *et.al.*, 1991
ST	Chile	L13774	Dekaban *et.al.*, 1992
STLENV1 (PtM3)	Asia	M11373	Watanabe *et.al.*, 1985

4.2.3 Phylogenetic Analysis

Sequences were aligned using DNASIS v7.00 (Hitachi Software Engineering), with minimal manual editing. Phylogenetic Analysis was performed using the MEGA (Molecular Evolutionary Genetics Analysis, Kumar *et.al.*, 1993) software package. Analysis was performed on the 1535 bp region as well as a 581 bp partial *env* region to incorporate more published cosmopolitan sequences. Phylogenetic trees were constructed for both the nucleotide sequence as well as the amino acid sequence. Multiple distance estimation and tree drawing methods were used to improve the validity of the results. The distance estimation and tree drawing methods are described.

4.2.3.1 Distance estimation

The evolutionary distance between a pair of sequences is usually measured by the number of nucleotide or amino acid substitutions between them. The evolutionary distances that are computed from DNA sequence data are primarily estimates of the number of nucleotide substitutions per site (*d*) between two sequences. Similarly, the distances computed from amino acid data are estimates of the number of amino acid substitutions per site between two sequences.

4.2.3.1.1 *P*-distance

The simplest method of distance estimation is the *p*-distance method. This distance is merely the proportion (*p*) of nucleotide sites or amino acid sites at which the two sequences compared are different. This is obtained by dividing the number of nucleotide/amino acid differences (n_d) by the total number of nucleotides/amino acids compared (n).

Therefore: $p = n_d/n$

The variance of *p* is given by:

$V(p) = p(1 - p)/n$.

The *p*-distance is approximately equal to the number of nucleotide/amino acid substitutions per site (*d*) only when it is small (for example when $p < 0.1$).

4.2.3.1.2 Kimura 2-parameter Model

In actual sequence data, the rate of transitional nucleotide substitution is often higher than that of transversional substitution. A maximum likelihood estimate of *d* for this case is given by Kimura's (1980) 2-parameter method. This estimate and its variance are given

by

$d = -^{1}/2\log_e(1 - 2P - Q) - {}^{1}/4\log_e(1 - 2Q)$,

$V(d) = [c_1^2P + c_3^2Q - (cP + c_3Q)^2]/n$,

where:

$c_1 = 1/(1 - 2P - Q)$, $c_2 = 1/(1 - 2Q)$, and $c_3 = {}^{1}/2(c_1 + c_2)$.

4.2.3.1.3 Jukes Cantor (JC) Distance

This method was developed under the assumption that the rate of nucleotide substitution is the same for all pairs of the four nucleotides (Jukes and Cantor (1969). It gives a maximum likelihood estimate of the number of nucleotide substitutions (d) between two sequences. The estimate and its variance are given by

$d = -^{3}/4\ \log_e(1-^{4}/3p)$,

$V(d) = p(1-p)/[1-^{4}/3p)^2n]$,

where: $p = n_d/n$

The JC distance can only be computed if $p<0.75$.

4.2.3.1.4 Poisson-Correction

This distance method is used for estimating the number of amino acid substitutions per site (d) by assuming that the number of amino acid substitutions at each site follows the Poisson distribution, where:

$d = -\log_e(1 - p)$.

The variance is given by:

$V(d) = p/[(1 - p)]$ (where $p = n_d/n$)

4.2.3.1.5 Gamma distance

This method estimates the number of amino acid substitutions by assuming that the rate of substitutions vary from site to site and follows the gamma distribution with parameter *a*. The distance is computed (Nei *et.al.*, 1976) as follows:

$d = a[(1 - p)^{-1/a} - 1]$

The variance is given by:

$V(d) = p[(1 - p)^{-(1 + 2a)}]/n.$

In the default option of MEGA, $a=2$ is used.

4.2.3.2 Tree building methods

Tree building methods can be classified into distance methods and discrete character methods. In distance methods, a pairwise evolutionary distance is computed for all operational taxonomic units (OTUs), and the tree is constructed by certain algorithms and principles. In discrete-character methods, the tree is constructed by considering the evolutionary relationships of OTUs at each character or nucleotide position.

4.2.3.2.1 NJ method

The NJ method (Saitou and Nei 1987) is an efficient distance method that examines the different possible topologies and produces only one final tree. This method produces an unrooted tree, and requires an outgroup OTU to find the root. In MEGA, the default in the absence of an outgroup, is to give the root at the midpoint of the longest route connecting two OTUs. (This is under the assumption of a constant rate of evolution.)

4.2.3.2.2 UPGMA

This distance method assumes that the rate of nucleotide or amino acid substitution is the same for all evolutionary lineages. This therefore produces a rooted tree, although the root can be removed for bootstrapping.

4.2.3.2.3 MP Method

The MP method is a discrete-character method of tree reconstruction. In MEGA, the MP method that is used assumes that evolutionary change ocurrs between any pair of the four nucleotides (Fitch, 1971). For constructing an MP tree, only parsimony-informative sites (nucleotide sites at which there are at least two different kinds of nucleotides, each represented at least twice) are used. In constructing a MP tree, the nucleotides of ancestral sequences are inferred at each nucleotide site for a given tree topology, and the minimum number of substitutions that are required to explain the nucleotide differences is counted. The sum of this number over all parsimony-informative sites of the sequences for a given topology is called the tree length. The tree length is then computed for all possible topologies, and the topology that shows the smallest tree length is chosen as the final tree (maximum parsimony tree).

In MEGA there are two ways of finding the MP tree: by using the branch-and-bound method (Hendy and Penny, 1982) or the heuristic search method. In the branch-and-bound method, the search for the MP tree starts with a core tree of three OTUs, which has only one unrooted tree. Other OTUs are added to this tree one by one by choosing the OTU with the maximum value out of all the minimum values (minimum number of substitutions), and the tree length (L) is computed at each stage of OTU addition. Tree lengths larger than a predetermined upperbound tree length (L_U), results in that topology

being ignored from further consideration.

The heuristic search method also starts with an initial core tree of three OTUs. The order of OTU addition is determined in a similar fashion to the branch-and-bound method, except that the OTU with the lowest value out of the minimum number of substitutions, is chosen (minimum-of-the-minimum-values-algorithm).

4.2.3.3 Alignment gaps

Alignment gaps were inserted during the alignment of homologous regions of the sequences. These sites were ignored in distance estimation by choosing the Complete-Deletion option in MEGA.

4.2.3.4 Statistical Tests of a Tree Obtained

NJ and UPGMA trees were statistically evaluated using 1000 bootstrap replications (Felsenstein, 1985; Zharkikh and Li, 1992). The bootstrapping procedure was the same for NJ and UPGMA trees, once MEGA had eliminated the root of the UPGMA tree. Briefly, in bootstrap tests, the same number of nucleotides as the actual number used for constructing the tree are randomly sampled with replacement from the original sequence data, and a tree is produced from this set of resampled nucleotide data. The topology of the tree is then compared with the original tree. Any interior branch of the tree that gives the same partition of sequences as that of the bootstrap tree is given a value 1 (identity value), whereas other interior branches are given 0. This process is repeated several hundred times, and the percentage of times each interior branch of the tree receives identity value 1 is computed. This is called the bootstrap confidence level (BCL). Branches with bootstrap values above 95% are usually considered to be robust.

In MEGA, MP trees are only bootstrapped if the number of OTUs is small (<20) and the number of equally parsimonius trees is one or few. If there is only one global MP tree, the bootstrap test is performed on this tree. However, if there are several global MP trees, a strict consensus tree (a composite tree of many equally parsimonious trees by forming a multifurcating branching pattern) is constructed. The bootstrap test is then performed on the consensus tree.

4.3 **Results**

4.3.1 Nucleotide Sequence Analysis

Phylogenetic analysis of the 1535 bp segment using the *p*-distance and NJ methods, showed the local strains to cluster within the cosmopolitan subtype (Fig. 4.1), particularly with the North American strains CH (Paine *et.al.*, 1991) and SP (Paine *et.al.*, 1991). Statistical analysis (1000 bootstrap replications) showed the branching pattern among the local strains and the North American strains to be unstable (BCL of 14% separating 49 and SP). The stability of the branching pattern increased for the cluster of Japanese strains (Seiki *et.al.*, 1983, Tsujimoto *et.al.*, 1988, Zhoa *et.al.*, 1993) (BCL: 100%). The continental American and Japanese strains were separate from the Caribbean strain (HTVPRCAR, Malik *et.al.*, 1988) within the cosmopolitan subtype (BCL of 98%). The Central African strain (EL, Paine *et.al.*, 1991) was separated from the cosmopolitan subtypes, with a BCL of 96%. The Melanesian strain (Gessain *et.al.*, 1993) was also separate, with a BCL of 100%. Similar topologies were obtained when using the Kimura-2-parameter and JC distance methods with the NJ tree drawing method.

Analysis using the UPGMA tree drawing method (Fig. 4.2) gave identical results for the p-distance, Kimura-2-parameter and JC distance methods. The topologies were similar to the NJ method, with low BCL for the branching pattern separating the local and North American strains (30%-63%). The BCL's increased for the branch separating the Japanese strains H5 (Tsujimoto *et.al.*, 1988) and HTVPROP (Seiki *et.al.*, 1983) from the other Japanese strains (BCL of 82%). The distinction between the Central African strain and the cosmopolitan strains was still evident (BCL of 72%).

MP analysis using both the Heuristic (Fig. 4.3) and Branched-and-bound search methods gave similar topologies for the p-distance, Kimura-2-parameter and JC distance methods. The local strains once again clustered with the North American strains (SP and CH: Paine *et.al.*, 1991). The Japanese strains HL1RHK30, HL1RHK34 (Baylot *et.al.*, 1996) and HTLV1ENV (Astier-Gin *et. al.*, 1991) clustered together, but were separate from the other Japanese strains H5 (Tsujimoto *et.al.*, 1988) and HTVPROP (Seiki *et.al.*, 1983). The BCL for this separation was 66%.

With the inclusion of more cosmopolitan strains (partial *env* analysis), the topology of the NJ tree using the *p*-distance method (Fig. 4.4) was essentially the same as the 1535 bp *env* analysis (Fig. 4.1). The local strains still clustered within the cosmopolitan subtype, particularly with the North American strains (CH and SP). The Martinique strain (HL12091, Baylot *et.al.*, 1996) and Chile strain (ST, Dekaban *et.al.*, 1992) were also included in this cluster. The Japanese strains once again were separated with the HL1RHK30, HL1RHK34, HTLV1ENV and HTVGP68 (Zhao *et.al.*, 1993; Astier-Gin *et.al.*, 1991; Takeuchi *et.al.*, 1985) strains clustering with the North American and KZN

strains. The other Japanese strains, HTVPROP (Seiki *et.al.*, 1983), H5 (Tsujimoto *et.al.*, 1988), HL, HB, HC and HK (De *et.al.*, 1991) formed a separate cluster (BCL of 60%). The branching pattern within the cluster was more reliable (BCL 67%-83%). A separate cluster was formed containing the USA (HA and HM, De *et.al.*, 1991), West African (HE, HG and HJ, De *et.al.*, 1991) and Caribbean (HTVPRCAR, [Malik *et.al.*, 1988] and HL12055 [Baylot *et.al.*, 1991]) strains. The Central African strain once again was separated from the rest of the cosmopolitan subtypes (BCL 53%).

A slightly different topology was obtained for the Kimura-2-parameter and NJ tree (Fig. 4.5). Four of the five strains clustered with the North American strain CH. Patient 49, fell outside of this cluster.

The trees obtained using the UPGMA method (Fig. 4.6) were the same for both the *p*-distance and the Kimura-2-parameter methods. A significant difference from the other trees was that the Central African strain (EL) fell within the cosmopolitan subtype, although it did not not cluster with any strains in particular. The West African, USA (HA and HM) and Caribbean strains did not form a cluster as they did for the NJ trees, although the BCL separating them were low (16% - 31%).

4.3.2 Amino Acid Sequence Analysis

Amino acid sequences were analysed using the *p*-distance (Fig. 4.7), Poisson correction (Fig 4.8), and Gamma distance (Fig 4.9) methods and the NJ tree drawing method. Three of the local strains (61, 62 and 68) clustered with the North American strain (CH), while the other two (48 and 49) clustered with the North American strain (SP). For the Poisson-correction and Gamma distance methods, the Caribbean strain (HTVPRCAR) clustered

with the Japanese MT-2 strains. The Caribbean strain did not cluster with the Japanese strains using the *p*-distance method. For all distance methods, the Central African strain (EL) fell within the cosmopolitan cluster.

For completion, the amino acid sequence of the KZN isolates were compared with the four South African strains reported by Mahieux *et.al.*, (1997) (Fig 4.10). The strains predictably clustered together, confirming the results that KZN trains are of the cosmopolitan subtype.

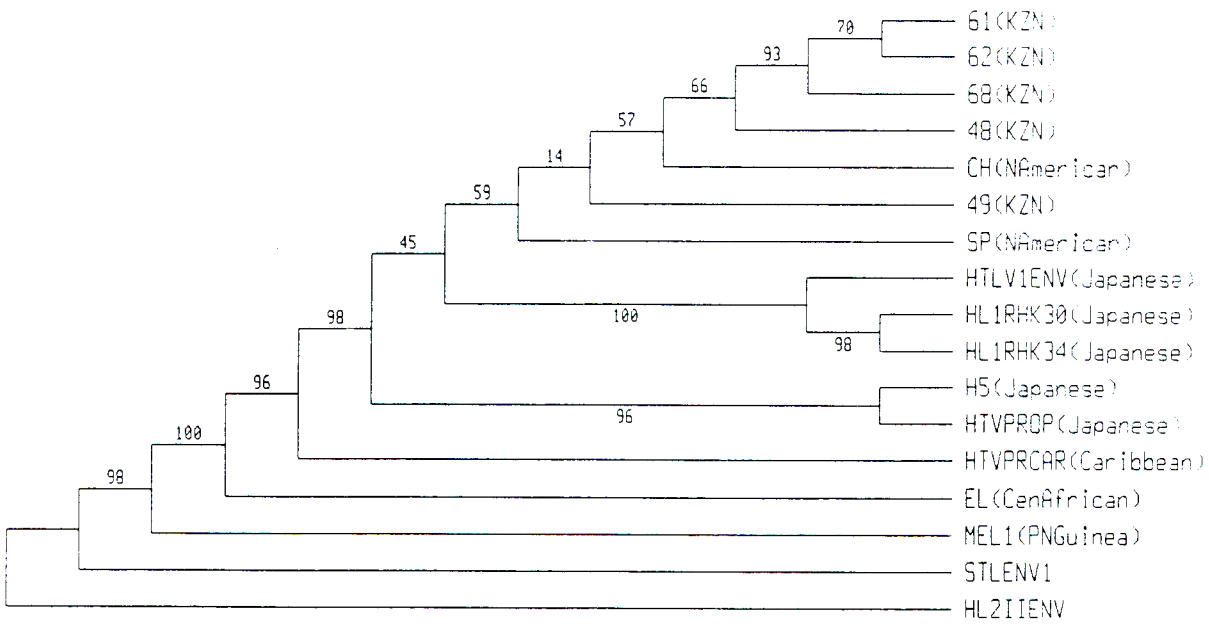


Fig. 4.1 Phylogeny of five HTLV-I KZN isolates (*env* nucleotide analysis) with other prototype isolates. The tree was constructed using the Neighbour-Joining (NJ) method and the *p*-distance estimation method. Bootstrap confidence levels are given along the branches.

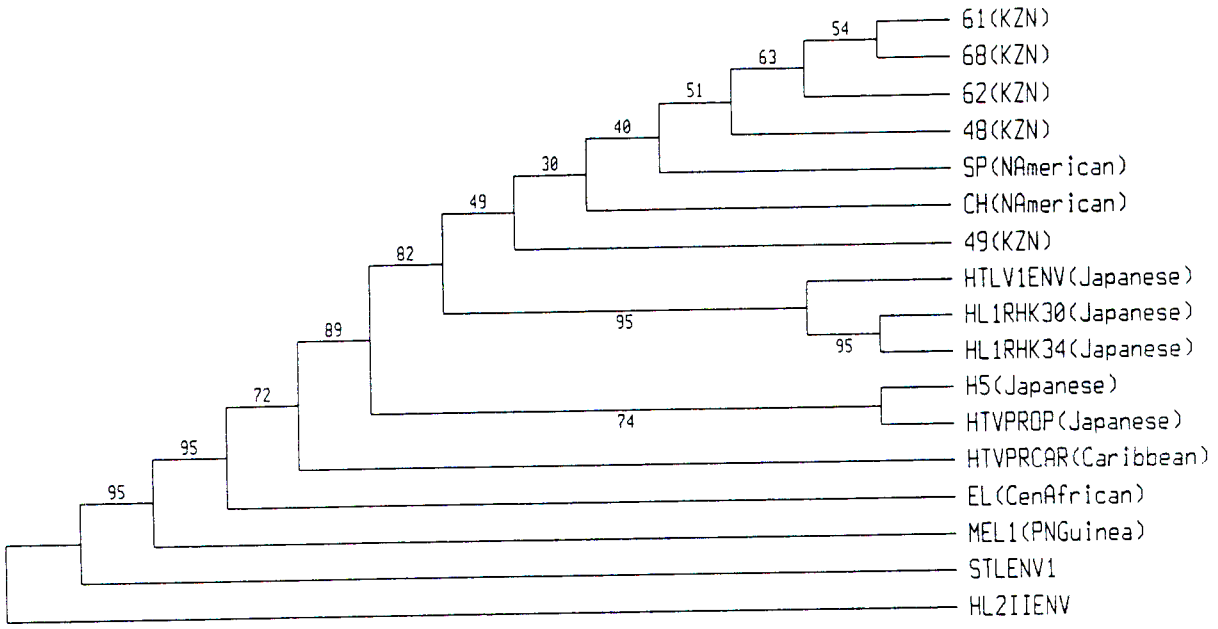


Fig. 4.2: Phylogeny of five HTLV-I KZN isolates (*env* nucleotide analysis) with other prototype isolates. The tree was constructed using the UPGMA and the Kimura-2-parameter distance estimation method. Bootstrap confidence levels are given along the branches.

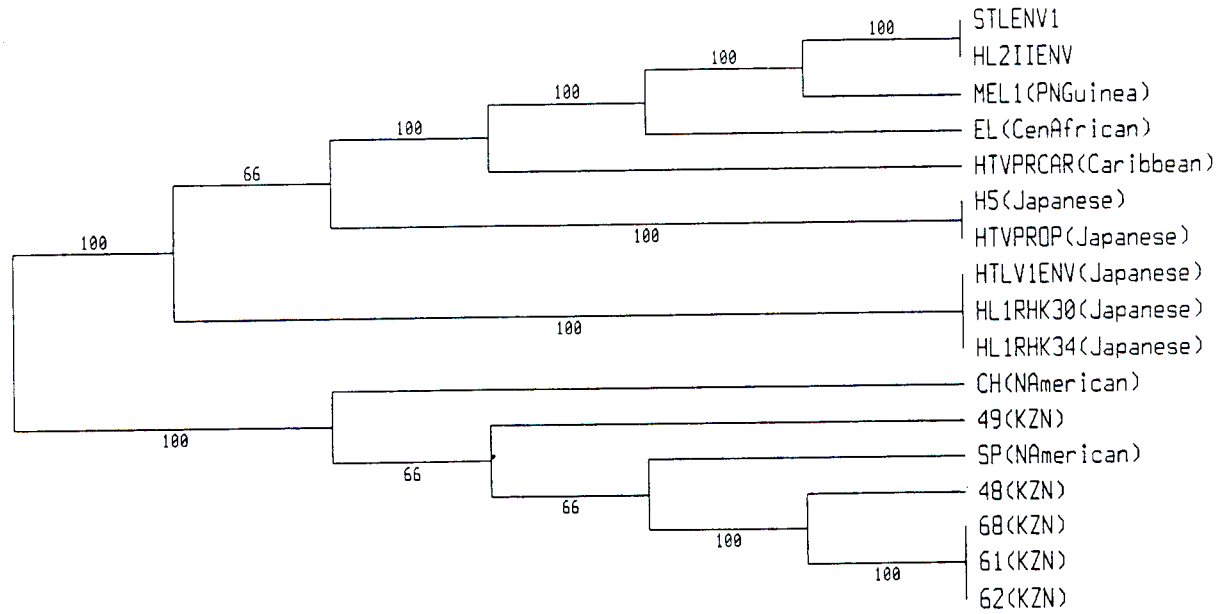


Fig. 4.3: Phylogeny of five HTLV-I KZN isolates (*env* nucleotide analysis) with other prototype isolates. The tree was constructed using the Maximum Parsimony (MP) method (Heuristic search) and the *p*-distance estimation method. Bootstrap confidence levels are given along the branches.

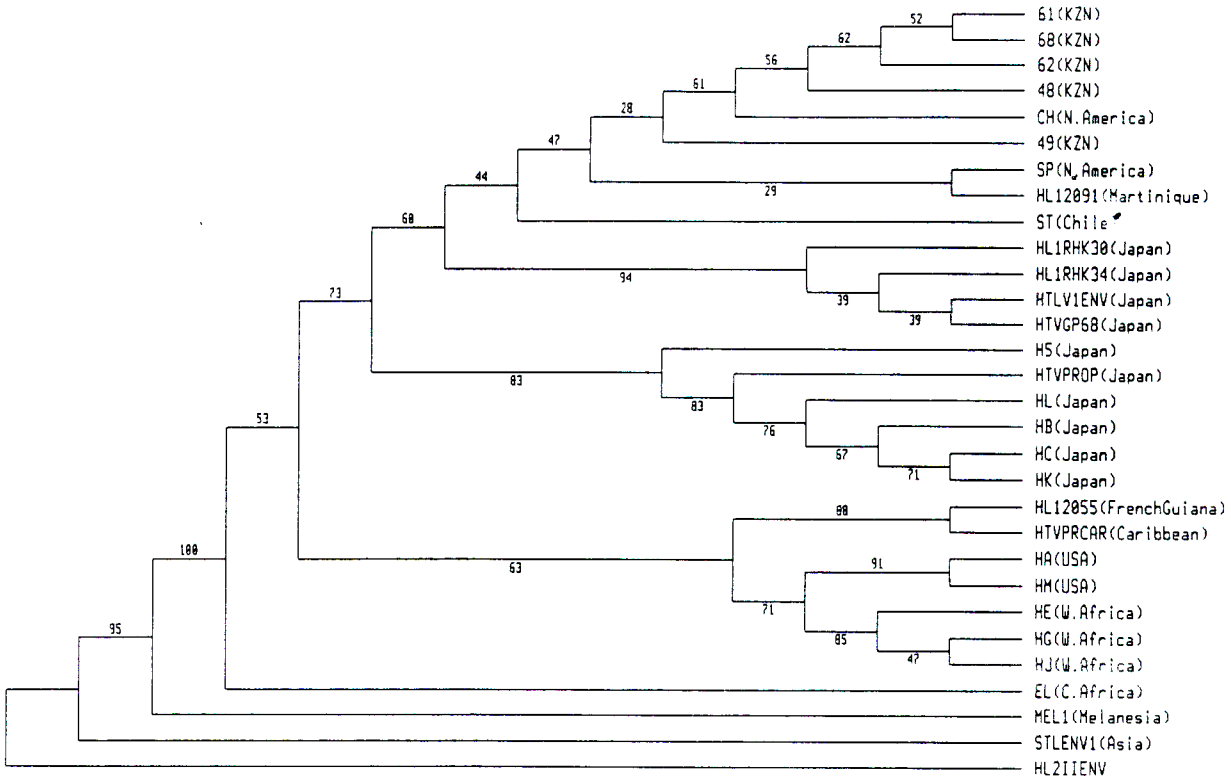


Fig. 4.4: Phylogeny of five HTLV-I KZN isolates (581 bp region of the *env* gene) with other prototype isolates. The tree was constructed using the NJ method and the *p*-distance estimation method. Bootstrap confidence levels are given along the branches.

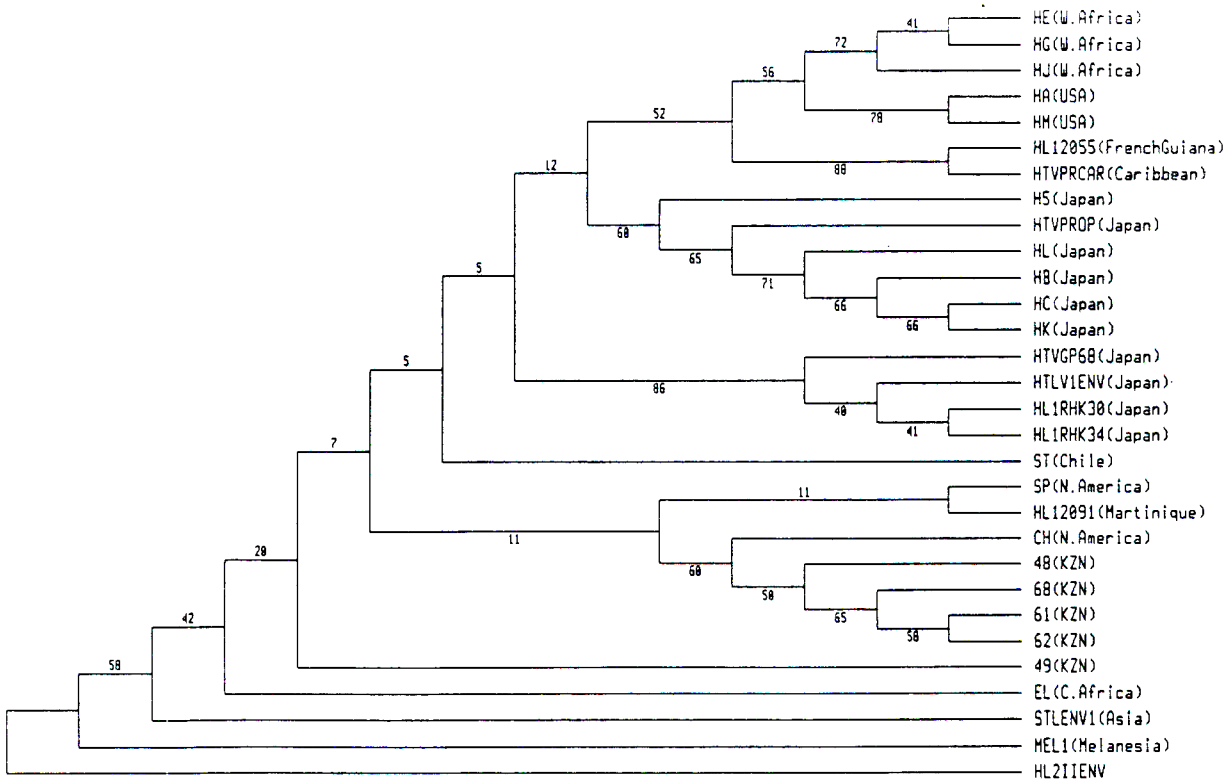


Fig. 4.5: Phylogeny of five HTLV-I KZN isolates (581 bp region of the *env* gene) with other prototype isolates. The tree was constructed using the NJ method and the Kimura-2-parameter distance estimation method. Bootstrap confidence levels are given along the branches.

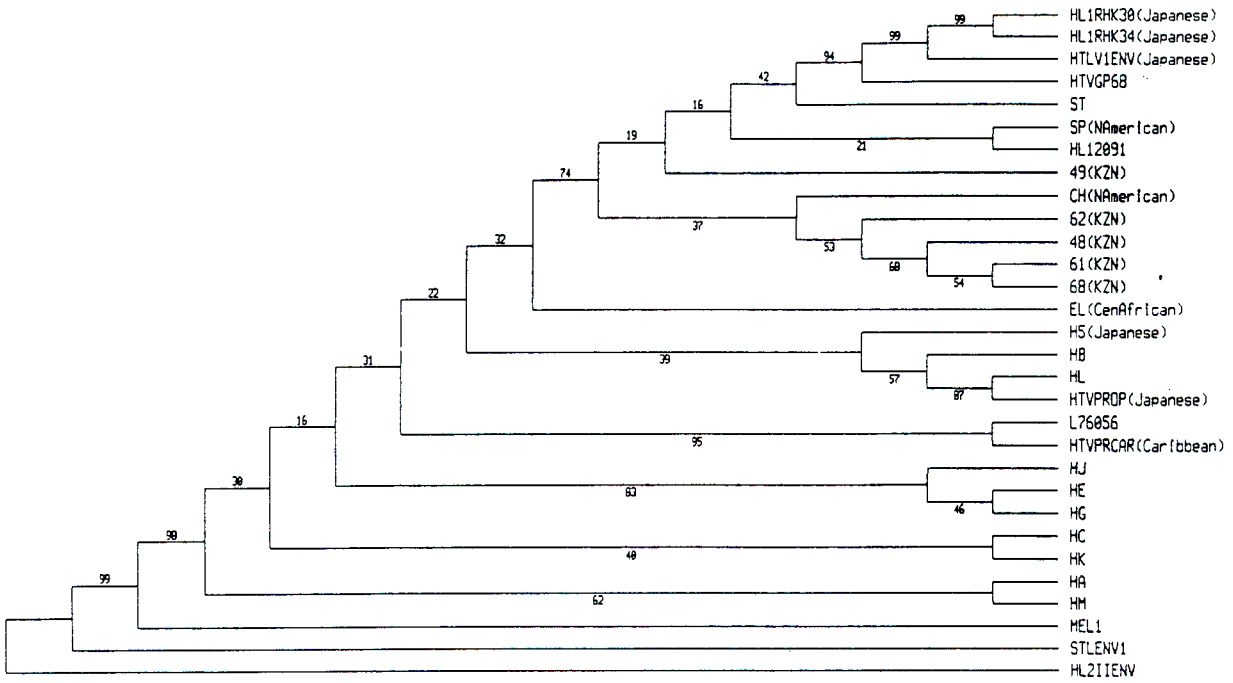


Fig. 4.6 Phylogeny of five HTLV-I KZN isolates (581 bp region of the *env* gene) with other prototype isolates. The tree was constructed using the UPGMA method and the *p*-distance estimation method. Bootstrap confidence levels are given along the branches.

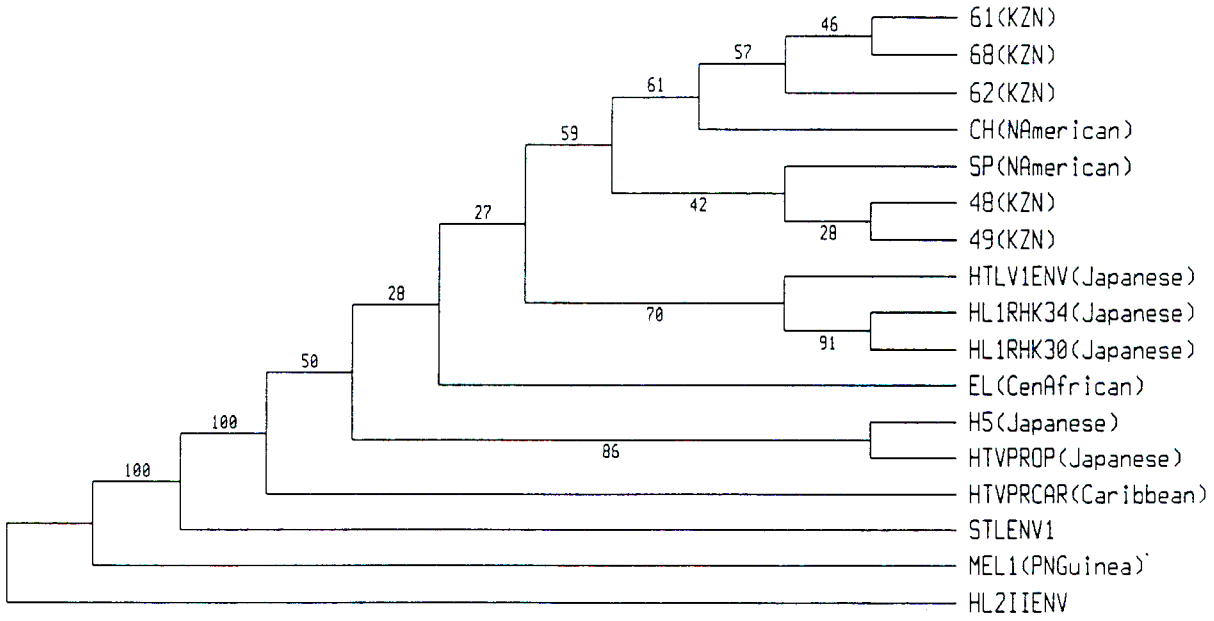


Fig. 4.7 Phylogeny of five HTLV-I KZN isolates (amino acid sequence analysis) with other prototype isolates. The tree was constructed using the NJ method and the *p*-distance estimation method. Bootstrap confidence levels are given along the branches.

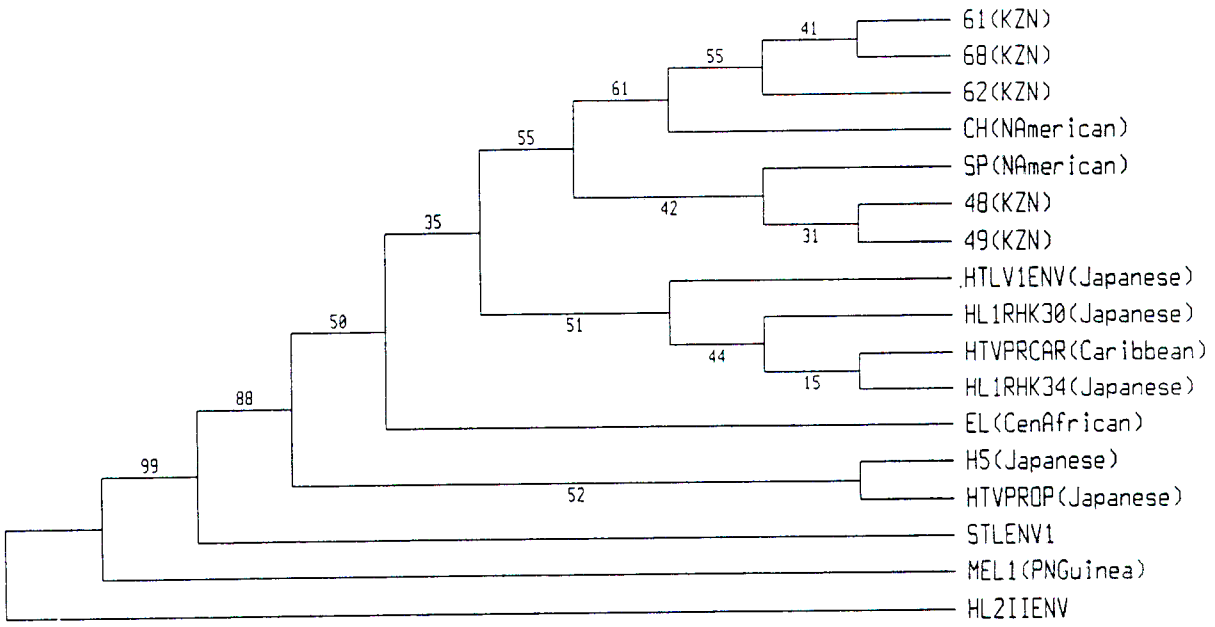


Fig. 4.8: Phylogeny of five HTLV-I KZN isolates (amino acid sequence analysis) with other prototype isolates. The tree was constructed using the NJ method and the Poisson-correction distance estimation method. Bootstrap confidence levels are given along the branches.

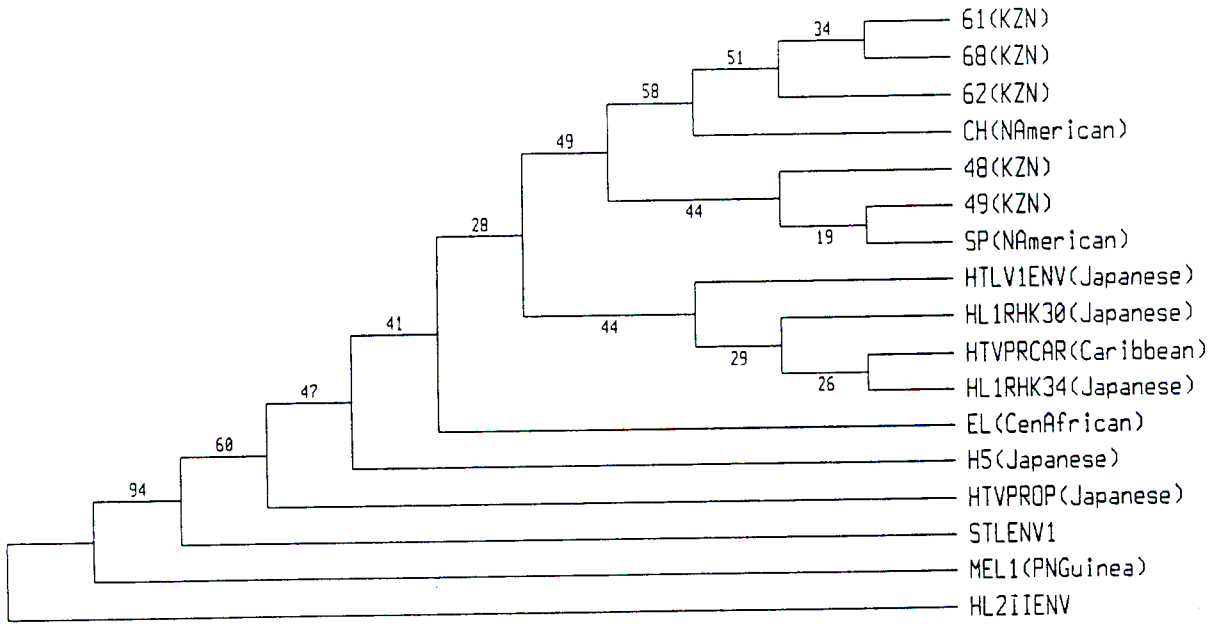


Fig. 4.9: Phylogeny of five HTLV-I KZN isolates (amino acid sequence analysis) with other prototype isolates. The tree was constructed using the NJ method and the Gamma distance estimation method, where a=2. Bootstrap confidence levels are given along the branches.

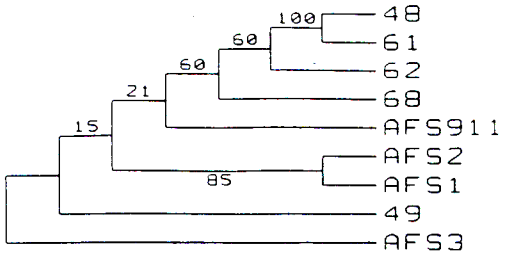


Fig. 4.10 Phylogeny of five HTLV-I KZN isolates (amino acid sequence analysis) with other South African isolates (ASF911, ASF1, ASF2 and ASF3; Mahieux *et.al.*, 1997). The tree was constructed using the NJ method and the *p*- distance estimation method. Bootstrap confidence levels are given along the branches.

4.4 **Discussion**

In chapter 3, new sequence data was provided on the HTLV-I *env* gene isolated from five KZN patients (48, 49, 61, 62, 68). In this chapter, the phylogenetic relationship of these sequences and published sequence data were investigated using the MEGA software package (Kumar *et.al.*, 1993). The branching pattern of the phylogenetic trees obtained were also compared with published tree data.

Phylogenetic analysis of the 1535 bp nucleotide sequences clearly associated the local strains with the cosmopolitan subtype. This is consistent with evidence by Engelbrecht *et.al.* (1996) and Mahieux *et.al.*, (1997), who have also reported South African strains to fall within the cosmopolitan subtype. The KZN strains were compared with the four South African strains isolated by Mahieux *et.al.*, (1997), and this further supported the classification of the KZN strains as the cosmopolitan subtype. Unfortunately, the strains isolated by Engelbrecht *et.al.*, (1996) could not be included in the analysis, as sequences from the LTR region were used in their analyses. The presence of a cosmopolitan subtype further emphasises the genetic diversity of HTLV-I isolates in Africa. Previous reports have noted a western and Central African variant (Vidal *et.al.*, 1994) as well as a Northern African variant (Gasmi *et.al.*, 1994).

The local strains clustered with the North American strains within the cosmopolitan

subtype. The stability of the branching pattern separating the local strains from the North American strains was low when tested with 1000 bootstrap replications (66% and 14%). Bootstrap values below 75% are generally believed not to be conclusive to support a topology, while values above 95% are usually significant (Zharkikh and Li, 1992). The low stability of the branches can be explained by the low genetic diversity between the local strains and the North American strains (Table 3.1). As base substitutions occur in HTLV-I over a long period in the highly conserved genome (Saksena *et.al.*, 1993), the low genetic diversity suggests a recent transmission from the North American strains. The bootstrap value of the branch separating isolates 61, 62 and 68 from isolates 48 and 49 was 93% (Fig. 4.1). This is consistent with the sequence data described in 3.3, where isolates 61, 62 and 68 shared base substitutions not common to the other two. Partial *env* (581 bp) and amino acid sequence analysis also clustered the local strains with the North American strains. However, one difference was the placement of the local isolate (49) outside the North American cluster for most of the trees. Isolate 49 only remained within this cluster for the *p*-distance and NJ analysis.

Partial *env* analysis was more informative, with the inclusion of more cosmopolitan strains. Three clusters could be seen within the cosmopolitan subtype: 1) the West African, American and Caribbean strains; 2) Japanese strains only; 3) the Japanese (MT-2 cell line), North American, South American and South African strains. However, the bootstrap values separating the different clusters were not very reliable. The most reliable separation was between the West African/Caribbean strains and the rest of the cosmopolitan strains (Japanese and American). The clustering of the West African strains with the Caribbean strain is consistent with Vandamme and colleagues (1994). It is widely accepted (Vandamme *et.al.*; 1994, Lui *et.al.* ,1996) that cosmopolitan HTLV-I spread

from West African HTLV-I to the Caribbean and then to the rest of the world. The separation of the Japanese strains into a cluster of Japanese strains only, with the others clustering with the cosmopolitan strains is consistent with Vidal *et.al.* (1994). Vidal and colleagues demonstrated the existence of two distinct molecular genotypes in Japan, one occurring throughout Japan, and the other mainly associated with port cities (cosmopolitan strains). Sequence data described in 3.3 showed the Japanese strains isolated from the MT-2 cell line to be more closely related to the North American strains. This is also supported by Vandamme *et.al.*, (1994), who reported a mixed Japanese/American group and a separate Japanese group. Lui *et.al.* (1996) have suggested that HTLV-I was introduction to the Americas in a stepwise process. It is possible that this stepwise process was via the Japanese strains.

In the amino acid sequence analysis, the Central African strain did not separate from the cosmopolitan strains, as it did in the nucleotide sequence analysis. However, the branching patterns were not very stable (BCL's of less than 35%). Strain EL appeared to associate with the Japanese strains. This contradicts reported data where Vandamme *et.al.*(1994) and Lui *et.al.* (1996), focusing on the nucleotide sequence for their phylogenetic analysis, found the Central African strain to separate from the cosmopolitan strains. Possibly, the nucleotide divergence of EL did not result in amino acid changes. This was probably due to the genomic constraints on the *env* region which inhibit amino acid changes that will result in a loss of function or infectivity (Pique *et.al.*, 1990). In this way, the difference between the Central African subtype and the cosmopolitan subtype did not emerge. For all trees, the Melanesian subtype was closely associated with the STLV-I Asian strain. This is consistent with reports suggesting an origin of the Australo-Melanesian subtype from an Asian STLV-I strain (Saksena *et.al.*, 1992;

Yanagihara *et.al.*, 1994; Vandamme *et.al.*, 1994; Lui *et.al.*, 1996). However, researchers disagree on the ancient origin of HTLV-I. Vandamme *et.al.* (1994) support an ancient African origin of HTLV-I, while Lui *et.al.* (1996) support an Asian origin of HTLV-I. The sequence data described in chapter three showed the Melanesian strain to have the most random substitutions of the HTLV-I strains analysed. This is consistent with a slow accumulation of bases over a prolonged period (Yanagihara *et.al.*, 1994), with the Australo-Melanesian strain evolving separately from the rest of the HTLV-I strains. However, Lui *et.al.* (1994) found the Melanesian strains to have a higher evolutionary rate than the rest of the HTLV-I strains, which could also explain the genetic diversity of the Melanesian strain. There is no conclusive evidence, therefore, for either an African or Asian origin of HTLV-I. Lui *et.al.* (1996) suggest that more gene regions and divergent strains from as many geographical locations as possible be included in a phylogenetic analysis if the origin of HTLV-I is to be solved.

It is important when comparing phylogenetic relationships, that distance and tree drawing methods are carefully chosen. In this study, the NJ method gave the most reliable topologies. This is consistent with previous reports that suggest that the UPGMA and MP methods are not as reliable as the NJ method (Saitou and Imanishi, 1989; Vandamme *et.al.*, 1994). The *p*-distance estimation method seemed to be the most reliable, although the Kimura-2-parameter method is also widely used (Engelbrecht *et.al.*, 1996; Vandamme *et.al.*, 1994). Ultimately, the comparison of as many topologies as possible is the best. The next step in the phylogenetic analysis of the local strains would be to amplify and sequence the more variable LTR region, and ultimately the entire genome. Phylogenetic trees obtained for the different genes will then be compared with the phylogenies obtained for the *env* region.

General Discussion and Conclusion

The initial objectives of the study were to establish an in-house diagnostic assay for HTLV-I and to clone and sequence amplified products from the *env* region of the genome. Five isolates were sequenced to identify the local strain.

Two areas of the genome were targeted for molecular diagnosis, the *pol* and *env* regions. The *pol* primers proved to be sensitive and specific. Amplification using the *env* primer pair was not reproducible, and was not pursued further. A novel technique, the AmpliSensor assay (Acugen Systems, Lowell, MA), was also tested. The AmpliSensor assay was very specific, but not as sensitive as PCR. However, there were many positive aspects of the AmpliSensor assay. Post-amplification manipulations were unnecessary, reducing the possibility of cross-contamination, detection occurred simultaneously with amplification and the progress of the amplification reaction could be monitored, making the assay quantitative.

In the second aspect of the study, a 1.535 kb region of the HTLV-I genome was amplified from peripheral blood obtained from five local seropositive patients. Four of the patients presented with HAM/TSP, and the fifth presented with a skin disease. Sequencing of these isolates showed no difference at the genomic level between the two clinical manifestations. The local isolates differed by 0.1% to 0.9% among themselves. They diverged by 1.1 % to 1.7% from the cosmopolitan strains, by 2.1% from the central African strain and by 7.5% from the Melanesian strain. Phylogenetic analysis clearly clustered the local strains within the cosmopolitan subtype. The local strains were

most closely related to the North American strains suggesting a possible transmission link from North America to South Africa.

For the future, the entire HTLV-I genome from one of the isolates (possibly 68), should be amplified and sequenced. Sequence data does not exist for the entire HTLV-I genome from this region.

References

Akizuki S, Setojuchi M, Nakazato O, Yoshida S, Higuchi Y, Yamamoto S, Okajima T (1988). **An autopsy case of human T-lymphotrophic virus type 1-associated myelopathy.** *Human Pathology*, **19:** 988-990.

Ando Y, Nakano S, Saito K, Shimamoto I, Ichijo M, Toyamata T and Hinuma Y (1987). **Transmission of adult T-cell leukaemia retrovirus (HTLV-I) from mother to child: Comparison of bottle- with breast-fed babies.** *Japanese Journal of Cancer Reasearch*, **78:** 322-324.

Asher D, Goudsmit J, Pomeroy KL, Garruto RM, Bakker M, Ono SG, Elliot N, Harris K, Askins, Eldadah Z, Goldstein AD and Gajdusek DC (1988). **Antibodies to HTLV-I in populations of the southwestern pacific.** *Journal of Medical Virology*, **26:** 339-351.

Astier-Gin T and Guillermain B (1991). **Nucleotide sequence of the envelope (env) gene from HTLV-I isolate MT-2.** *Genebank Accession number X56949.*

Bastian I, Gardner J, Webb D and Gardner I (1993). **Isolation of a strain of human T-lymphotropic virus type I strain from Australian Aboriginals.** *Journal of Virology*, **67:** 843-851.

Baylot F, Edouard E, Moynet DJ and Guillemain B (1996). **HTLV-I sequences.** *Genebank Accession number L76056.*

Bazabachi A and Hermine O (1996). **Treatment with a combination of zidovudine and alpha-interferon in naive and pretreated adult T-cell leukemia/lymphoma patients.** *Journal of Acquired Immune Deficiency Syndrome and Human Retrovirology,* **13:** s186-s190.

Bazarbachi A, Huang M, Gessain A, Saal F, Saib A, Peries J, de The H and Galibert F (1995). **Human T-cell-leukemia virus type I in post-transfusional spastic paraparesis: complete proviral sequence from uncultured blood cells.** *International Journal of Cancer,* **63:** 494-499.

Bhigjee AI, Moodley AA, Madurai S, York D and Bill PLA (1995). **Infantile onset of HTLV-I-associated myelopathy (HAM/TSP).** *South African Medical Journal,* **85:** 687-688.

Bhigjee AI, Thaler D, Madurai S, Gouws E and Bill PLA (1994). **Seroprevalence of HTLV-I in Natal/KwaZulu.** *South African Medical Journal,* **84**: 368.

Bhigjee AI, Vinsen C, Windsor IM, Gouws E, Bill PLA and Tait D (1993). **Prevalence and transmission of HTLV-I infection in Natal/KwaZulu.** *South African Medical Journal,* **83:** 665-667.

Blattner WA (1990). **Epidemiology of HTLV-1 and associated diseases.** In: (WA Blattner [ed]) *Human Retrovirology: HTLV,* Raven, New York, **pp** 251-265.

Blattner WA, Nomura A, Clark JW, Ho GYF, Nakao Y, Gallo R and Robert-Guroff M (1986). **Modes of transmission and evidence for viral latency from studies of human T-cell lymphotropic virus type 1 in Japanese migrant populations in Hawaii.** *Proceedings on the National Academy of Sciences, USA,* **83:** 4895-4898.

Blattner WA, Kalyanaraman VS, Robert-Guroff M, Lister TA, Galton DAG, Sarin PS, Crawford MH, Catovsky D, Greaves M and Gallo RC (1982). **The human type-C retrovirus, HTLV, in blacks from the Caribbean region and its relationship to adult T-cell leukemia/lymphoma.** *International Journal of Cancer,* **30:** 257-264.

Borg A, Yin JA, Johnson PR, Tosswill J, Saunders M and Morris D (1996). **Successful treatment of HTLV-I-associated acute adult T-cell leukaemia lymphoma by allogeneic bone marrow transplantation.** *British Journal of Haematology,* **94:** 713-715.

Bos JL, Fearon ER, Hamilton SR, Verlaan-de Vries M, van Boom JH, van der Eb AJ, and Vogelstein B (1987). **Prevalence of *ras* gene mutations in human colorectal cancers.** *Nature,* **327:** 293.

Burnette WN (1981). **Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A.** *Analytical Biochemistry,* **112:** 195-203.

Cann AJ and Chen ISY (1990). **Human T-cell leukemia virus types I and II.** In: (Fields BN, Knipe DR [eds]) *Virology,* Raven Press, New York, **pp** 1501-1527.

Cann AJ, Rosenblatt JD, Wachsman W, Shah NP and Chen ISY (1985). **Identification of the gene responsible for human T-cell leukaemia virus transciptional regulation.** *Nature,* **318:** 571-574.

Cartier L, Mora C, Araya F, Castillo J, Verdugo R, Miller MA, Gadjusek DC and Gibbs CJ (1989). **HTLV-I positive spastic paraparesis in a temperate zone.** *Lancet*, **1:** 556-557.

Cavalli-Sforza LL, Piazza A, Menozzi P and Mountain J (1988). **Reconstruction of human evolution: bringing together genetic, archaeological and linguistic data.** *Proceedings of the National Academy of Sciences, USA,* **35:** 6002-6006.

Cavalli-Sforza LL and Edwards AWF (1967). **Phylogenetic analysis: Models and estimation procedures.** *American Journal of Human Genetics,* **19:** 233-257.

Centres for Disease Control (1988). **Licencure of screening tests for antobody to human T-lymphotropic virus type 1.** *Morbidity and Mortality Weekly Report,* **37:** 736-747.

Chamberlain JS, Gibbs RA, Ranier JE, Nguyen PN and Caskey CT (1988). **Deletion scanning of the Duchenne muscular dystrophy locus via mulitiplex DNA**

amplification. *Nucleic Acids Research,* **16:** 11141-11156.

Clark JM (1988). **Novel non-templated nucleotide addition reactions catalysed by procaryotic and eucaryotic DNA polymerases.** *Nucleic Acids Research,* **16:** 9677-9686.

Coates SR, Harris AJ, Parkes DL, Smith CM, Lui H-L, Akita RW, Ferrer MM, Sampson EK, Brandis JW and Sliwkowski MX (1990). **Serological evaluation of *Escherichia coli*-expressed human T-cell leukaemia virus type I env, gag p24, and tax proteins.** *Journal of Clinical Microbiology,* **28:** 1139-1142.

Dalgleish A, Richardson J and Matutes E (1988). **HTLV-I infection in tropical spastic paraparesis: lymphocyte culture and serologic response.** *AIDS Research and Human Retroviruses,* **4:** 475-485.

De BK, Lairmore MD, Griffis K, Williams LJ, Villinger F, Quinn TC, Brown C, Nzilambi, Sugimoto M, Araki S and Folks T (1991)). **Comparative analysis of nucleotide sequence of the partial envelope gene (5' domain) among human T lymphotropic virus type I (HTLV-I) isolates.** *Virology,* **182:** 413-419.

Dekaban GA, King EE, Waters D and Rice GPA (1992). **Nucleotide sequence analysis of an HTLV-I isolate from a Chilean patient with HAM/TSP.** *AIDS Research and Human Retroviruses,* **8:** 1201-1207.

Delaporte E, Monplaisir N, Louwagie J, Peeters M, Martin-Prevel Y, Louis J-P, Trebuco A, Bedjabaga L, Ossari S, Honore C, Larouze B, D'Auriol L, Van der Groen G and Piot P (1991). **Prevalence of HTLV-I and HTLV-II infection in Gabon, Africa: Comparison of the serological and PCR results.** *International Journal of Cancer*, **49:** 373-376.

Delaporte E, Peeters M, Simoni M and Piot P (1989). **HTLV-I infection in western equatorial Africa.** *Lancet*, **2:** 1226.

DiCesare J, Grossman B, Katz E, Picozza E, Ragusa R and Woudenberg T (1993). **A high-sensitivity electrochemiluminescence-based detection system for automated PCR product quantitation.** *Biotechniques,* **15:** 152-157.

Dodon M, Bernard A and Gazzolo L (1989). **Peripheral T-lymphocyte activation by human T-cell leukaemia virus type 1 interferes with the CD2 but not with the CD3/TCR pathway.** *Journal of Virology,* **63:** 5413-5419.

Duncan J and Rudge P (1990). **Methylprednisolone therapy in tropical spastic paraparesis.** *Journal of Neurology Neurosurgery and Psychiatry,* **53:** 173-174.

Ehrlich GD, Greenberg S and Abbott MA (1990). **Detection of human T-cell lymphoma/leukemia virus.** In: (Innis M, Gelfand D, Sninsky J and White T [eds]) *PCR Protocols: A Guide to Methods and Applications*, Academic Press, CA, **pp** 325-335.

Engelbrecht S, van Rensburg EJ and Robson BA (1996). **Sequence variation and subtyping of human and simian T-cell lymphotropic virus type I strains from South Africa.** *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology,* **12:** 298-302.

Engelke DR, Hoener PA and Collins FS (1988). **Direct sequencing of enzymatically amplified human genomic DNA.** *Proceedings of the National Academy of Sciences,* **85:** 544.

Erlich HA, Gelfand DH and Saiki RK (1988). **Specific DNA amplification.** *Nature,* **331:** 461.

Essex M, Matsuda Z, Yu X and Lee TH (1995). **Gene therapy against retroviral diseases.** *Leukemia,* **9:** s71-s74.

Evans BK, Gore I, Harrell LE, Arnold T and OH SJ (1989). **HTLV-I associated myelopathy and polymyositis in a US native.** *Neurology,* **39:** 1572-1575.

Felsenstein J (1989). **PHYLIP - Phylogeny inference package (version 3.2).** *Cladistics,* **5:** 164-166.

Felsenstein J (1985). **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution,* **39:** 783-791.

Fitch WM (1971). **Towards defining the course of evolution: Minimum change for a specific tree topology.** *Systematic Zoology,* **20:** 406-416.

Fleming AF, Maharajan R, Abraham M, Kulkarni A,G, Bhusnurmath SR, Okpara RA, Williams E, Akinsete I, Schneider J, Bayer H and Hunsmann G (1986). **Antibodies to HTLV-I in Nigerian blood donors, their relatives and patients with leukaemia, lymphomas and other diseases.** *International Journal of Cancer,* **38:** 809-813.

Gallo RC, Wong-Staal F, Montagnier L, Haseltine WA and Yoshida M (1988). **HIV/HTLV gene nomenclature.** *Nature,* **333:** 504.

Gallo RC, Sliski A and Wong-Staal F (1983). **Origin of human T-cell leukaemia-lymphoma virus.** *Lancet,* **2:** 962-963.

Gasmi M, Farouqi B, d'Incan M and Desgranges C (1994). **Long terminal repeat sequence analysis of HTLV type I molecular variants identified in four North African patients.** *AIDS Research and Human Retroviruses,* **10:** 1313-1315.

Gessain A, Boeri E, Yanagihara R, Gallo RC and Franchini G (1993). **Complete nucleotide sequence of a highly divergent human T-cell leukaemia (lympotropic) virus type I (HTLV-I) variant from Melanesia: genetic and phylogenetic relationship to HTLV-I strains from other geographical regions.** *Journal of Virology,* **67:** 1015-1023.

Gessain A, Gallo RC and Franchini G (1992). **Low degree of human T-cell leukemia/lymphoma virus type I genetic drift in vivo as a means of monitoring viral transmission and movement of ancient human populations.** *Journal of Virology,* **66:** 2288-2295.

Gessain A, Yanagihara R, Franchini G, Garruto RM, Jenkins CL, Ajdukiewicz AB, Gallo RC and Gajdusek DC (1991). **Highly divergent molecular variants of human T-lymphotropic virus type I from isolated populations in Papua New Guinea and the Solomon Islands.** *Proceedings of the National Academy of Sciences, USA,* **88:** 7694-7698.

Gessain A, Francis H, Sonan T, Fiordano C, Akani F, Piquemal M, Caudie C, Malone G, Essex M and de The' G (1986). **HTLV-I and tropical spastic paraparesis in Africa.** *Lancet,* **2:** 698.

Gessain A, Barin F, Vernant JC, Gout O, Maurs L, Calender A and de The G (1985). **Antibodies to human T lymphotropic virus type I in patients with tropical spastic paraparesis.** *Lancet,* **2**: 407-410.

Gilliand G, Perrin S, Blanchard K and Bunn HF (1990). **Analysis of cytokine mRNA and DNA: Detection and quantitation by competitive polymerase chain reaction.** *Proceedings of the National Academy of Sciences, USA,* **87:** 2725-2729.

Goubau P, Desmyter J, Swanson P, Reynders M, Shih J, Surmont I, Kazadi K and Lee H (1993). **Detection of HTLV-I and HTLV-II infection in Africans using type-specific envelope peptides.** *Journal of Medical Virology.* **39:** 28-32.

Goudreau G, Karpati G and Carpenter S (1988). **Inflammatory myopathy in association with chronic myelopathy in HTLV-I seropositive patients.** *Neurology,* **38(1):** 205.

Green P and Chen I (1990). **Regulation of Human T-cell leukaemia virus expression.** *Federation of American Societies for Experimental Biology Journal,* **4:** 169-175.

Haff LA (1994). **Improved quantitative PCR using nested primers.** *PCR Methods and Applications,* **3:** 332-337.

Hartley T, Khabbaz R, Cannon R, Kaplan J and Lairmore M (1990). **Characterization of antibody reactivity to human T-cell lymphotropic virus types1 and 2 using immunoblot and radioimmunoprecipitation assays.** *Journal of Clinical Microbiology,* **28:** 646-650.

Haseltine WA, Sodroski JG and Patarca R (1985). **Structure and function of the genome of HTLV.** In: (Vogt PK [ed]) *Current Topics in Microbiology and Immunology,* Springer, Berlin, **pp** 177-209.

Hendy MD and Penny D (1982). **Branch and bound algorithms to determine minimal evolutionary trees.** *Mathematical Biosciences,* **59:** 277-290.

Heneine W, Khabbaz RF, Lal R and Kaplan JE (1992). **Sensitive and specific polymerase chain reactions assays for diagnosis of human T-cell lymphotropic virus type I (HTLV-I) and HTLV-II infections in HTLV-I/II-seropositive individuals.** *Journal of Clinical Microbiology,* **30:** 1605-1607.

Higuchi R, Dollinger G, Walsh PS and Griffith R (1992). **Simulataneous amplification and detection of specific DNA sequences.** *Bio/Technology,* **10:** 413-417.

Higuchi R, von Beroldingen CH, Sensabaugh GF and Erlich HA (1988). **DNA typing from single hairs.** *Nature,* **332:** 543.

Hino S (1989). **Milk-borne transmission of HTLV-I as a major route in the endemic cycle.** *Acta Paediatrica Japan*, **31:** 428-435.

Hinuma Y, Komoda H, Chosa T, Kondo T, Kohakura M, Takenaka T, Kikuchi M, Ichimaru M, Yunoki K, Sato I, Matsuo R, Takiuchi Y, Uchino H and Hanaoka M (1982). **Antibodies to adult T-cell leukaemia virus-associated antigen (ATLV) in sera from patients with ATLL and controls in Japan: A nationwide sero-epidemiolgical study.** *International Journal of Cancer*, **29:** 631-635.

Hinuma YK, Nagata K, Hanaoka M, Nakai M, Matsumoto T, Kinoshita KI, Shirakawa S and Miyoshi I (1981). **Adult T-cell leukemia: antigen in a ATL cell line and detection of antibodies to the antigen in human sera.** *Proceedings of the National Academy of Sciences, USA,* **78:** 6476.

Hjelle B, Cyrus S and Swenson SG (1992). **Evidence for sexual transmission of Human T-lymphotropic virus type II.** *Annals of Internal Medicine*, **116:** 90-91.

Hollsberg P and Hafler D (1993). **Pathogenesis of diseases induced by human T-lymphotropic virus type I infection.** *New England Journal of Medicine,* **328:** 1173-1182.

Jacobson S, McFarlin DE, Robinson S, Voskuhl R, Martin R, Brewah A, Newell AJ and Koenig S (1992). **HTLV-I specific cytotoxic T lymphocytes in the cerebrospinal fluid of patients with HTLV-I-associated neurological disease.** *Annals of Neurology,* **32:** 651-657.

Jacobson S, Shida H, McFarlin DE, Fauci AS and Koening S (1990). **Circulating CD8+ cytotoxic T lymphocytes specific for HTLV-I pX in patients with HTLV-I-associated neurological disease.** *Nature,* **348:** 245-248.

Jukes TH and Cantor CR (1969). **Evolution of protein molecules.** In: (Munro HN [ed]) *Mammalian Protein Metabolism*, Academic press, New York, **pp** 21-132.

Katamine S, Moriuchi R, Yamamoto T, Terada K, Eguchi K, Tsuji Y, Yamabe T, Miyamoto T and Hino S (1994). **HTLV-I proviral DNA in umbilical cord blood of babies born to carrier mothers.** *Lancet,* **343:** 1326-1327.

Kawano F, Yamaguchi K, Nishimura H, Tsuda H and Takatsuki K (1985). **Variation in the clinical courses of adult T-cell leukemia.** *Cancer,* **55:** 851-856.

Khabbaz RF, Hartley TM, Lairmore MD and Kaplan JE (1990). **Epidemiologic assessment of scrreening tests for antibody to human T-lymphotropic virus type 1 (HTLV-I).** *American Journal of Public Health,* **80:** 190-192.

Kimura M (1980). **A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences.** *Journal of Molecular Evolution,* **16:** 111-120.

Kira J-I, Koyanagi Y, Yamada T, Itoyama Y, Tateishi J, Akizuki S-I, Kishikawa M, Baba E, Nakamura M, Suzuki J, Nakamura T, Nakamura N, Yamamoto N and Goto I (1994). **Sequence heterogeneity of HTLV-I proviral DNA in the central nervous system of patients with HTLV-I-associated myelopathy.** *Annals of Neurology,* **36:** 149-156.

Kira J-I, Koyanagi Y, Yamada T, Itoyama Y, Goto I, Yamamoto N, Sasaki H and Sasaki Y (1991). **Increased HTLV-I proviral DNA in HTLV-I-associated myelopathy: a quantitative polymerase chain reaction study.** *Annals of Neurology,* **29:** 194-201.

Komurian F, Pelloquin F and de The' G (1991). **In vivo genomic variability of human T-cell leukemia virus type I depends more upon geography than upon pathologies.** *Journal of Virology*, **65:** 3770-3778.

Koralnik IJ, LoMonico A, Fullen J, Saxinger C, Gessain A, Gallo RC, Markham P, Kalyanaraman V, Boeri E, Hirsch V, Allan J, Murthy P, Guo H-G, Alford P and

Franchini G (1993). **A wide spectrum of simian T-lymphotropic/leukemia type I virus variants in nature: Evidence for interspecies transmission in Equatorial Africa.** *AIDS Journal,* **6:** 745.

Korber B, Okayama A, Donnelly R, Tachibana N and Essex M (1991). **Polymerase chain reaction analysis of defective human T-cell leukemia virus type I proviral genomes in leukemic cells of patients with adult T cell leukemia.** *Journal of Virology,* **65:** 5471-5476.

Kumar S, Tamura K and Nei M (1993). **MEGA: molecular evolutionary genetics analysis, version 1.01.** University Park, Pennsylvania: The Pennsylvania State University.

Kwok S, Kellog D, Ehrlich GD, Poiesz BJ, Bhagavati S and Sninsky JJ (1988). **Characterisation of a sequence of human T-cell leukemia virus type I from a patient with chronic progressive myelopathy.** *Journal of Infectious Diseases,* **158:** 1193-1197.

Kwok S, Mack DH, Mullis KB, Poiesz BJ, Ehrlich GD, Blair D, Friedman-Kein A, Sninsky JJ (1987). **Identification of human immunodeficiency virus sequences by using in vitro enzymatic amplification and oligomer cleavage detection.** *Journal of Virology,* **61:** 1690-1694.

Lairmore M, Jacobson S, Gracia F, De B, Castillo L, Larreategui M, Roberts B, Levine P, Blattner W and Kaplan J (1990). **Isolation of human T-lymphotropic virus type 2**

from Guaymi indians in Panama. *Proceedings of the National Academy of Sciences, USA,* **87:** 8840-8844.

Lee H, Idler KB, Swanson P, Aparicio JJ, Chin KK, Lax JP, Nguyen M, Mann T, Leckie G, Zanetti A, Marinucci G, Chen ISY and Rosenblatt JD (1993). **Complete nucleotide sequence of HTLV-II isolate NRA: comparison of env sequence variation of HTLV-II isolates from U.S blood donors and US and Italian IV drug users.** *Virology,* **196:** 57-69.

Lee H, Swanson P, Shorty VS, Zack JA, Rosenblatt JD and Chen IS (1989). **High rate of HTLV-II infection in seropositive IV drug abusers in New Orleans.** *Science,* **244:** 471-475.

Lion T (1994). **Clinical implications of qualitative and quantitative polymerase chain reaction analysis in the monitoring of patients with chronic myelogenous leukemia. The European Investigators on Chronic Myeloid Leukemia Group.** *Bone Marrow Transplant,* **14:** 505-509.

Loeffelholz MJ, Lewinski CA, Silver SR, Purohit AP, Herman SA, Buonagurio DA and Dragon EA (1992). **Detection of Chlamydia trachomatis in endocervical specimens by polymerase chain reaction.** *Journal of Clinical Microbiology,* **30:** 2847-2851.

Lui H-F, Goubau P, Van Brussel M, Van Laethem K, Chen Y-C, Desmyter J and Vandamme A-M (1996). **The three human T-lymphotropic virus type I subtypes arose**

from three geographically ditinct simian reservoirs. *Journal of General Virology,* **77:** 359-368.

Major ME, Nightingale S and Desselberger U (1993). **Complete sequence conservation of the human T cell leukaemia type 1 *tax* gene within a family cluster showing different topologies.** *Journal of General Virology,* **74:** 2531-2537.

Malik KTA, Even J and Karpas A (1988). **Molecular cloning and complete nucleotide sequencing of an adult T-cell leukemia virus type I isolate of Carribean origin: relationship to other members of the HTLV-1 subgroup.** *Journal of General Virology,* **69:** 1695-1700.

Marchuk D, Drumm M, Saulino A and Collins FS (1991). **Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products.** *Nucleic Acids Research,* **19:** 1154.

Matsumoto C, Mitsunaga S, Oguchi T, Mitomi Y, Shimada T, Ichikawa A, Watanabe J and Nishicka K (1990). **Detection of human T-cell leukemia virus type I (HTLV-I) provirus in an infected cell line and in peripheral mononuclear cells of blood donors by the nested double polymerase chain reaction method: comparison with HTLV-I antibody tests.** *Journal of Virology,* **64:** 5290-5294.

McNearney T, Westervelt P, Thielan BJ, Trowbridge DB, Garcia J, Whittier R and Ratner L (1990). **Limited sequence heterogeneity among biologically distinct HIV-I**

isolates from individuals involved in a clustered infectious outbreak. *Proceedings of the National Academy of Sciences, USA,* **87:** 1917-1921.

Miura T, Fukunaga T, Igarishi T, Yamashita M, Ido E, Funahashi S-I, Ishida T, Washio K, Ueda S, Hashimoto K-I, Yoshida M, Osame M, Singhal BS, Zaninovic V, Cartier L, Sonoda S, Tajima K, Ina Y, Gojobori T and Hayami M (1994). **Phylogenetic subtypes of human T-lymphotropic virus type I and their relations to the anthropological background.** *Proceedings of the National Academy of Sciences, USA,* **91:** 1124-1127.

Miyamoto MM and Cracraft J (1991). **Phylogenetic Analysis of DNA Sequences**, Oxford University press, New York.

Miyoshi I, Kubonoshi, Sumida M, Hiraki S, Kimura I, Miyamoto K and Soto J (1980). **A novel T-cell line derived from adult T-cell leukemia.** *Gann,* **71:** 155-156.

Moore W, Traugott U, Scheinberg L and Raine C (1989). **Tropical spastic paraparesis: a model of virus-induced, cytotoxic t-celll demyelination?** *Annals of Neurology,* **26:** 523-530.

Mullis KB and Faloona FA (1987). **Specific synthesis of DNA *in vitro* via a polymerase catalysed chain reaction.** *Methods in Enzymology,* **155:** 335-350.

Nei M (1991). **Relative efficiencies of different tree making methods for molecular data.** In: (Miyamoto MM and Cracraft JL [eds]) *Recent Advances in Phylogenetic Studies*

of DNA Sequences, Oxford University Press, Oxford, **pp** 90-128.

Nei M, Chakraborty R and Fuerst PA (1976). **Infinite allele model with varying mutation rate.** *Proceedings of the National Academy of Sciences, USA,* **73:** 4164-4168.

Nerurkar VR, Song K-J, Bastian IB, Garin B, Franchini G and Yanagihara R (1994). **Genotyping of human T cell lymphotropic virus type I using Australo-Melanesian topotype-specific oligonucleotide primer-based polymerase chain reaction: insights into viral evolution and dissemination.** *Journal of Infectious Diseases,* **170:** 1353-1360.

Nerurkar VR, Babu PG, Song K-J, Melland RR, Gnanamuthu C, Saraswathi NK, Chandy M, Godec MS, John TJ and Yanagihara R (1993). **Sequence analysis of human T lymphotropic virus type I strains from southern India: gene amplification and direct sequencing from whole blood blotted onto filter paper.** *Journal of General Virology,* **74:** 2799-2805.

Nicholson SR, Efandis T, Dimitrakakis M, Karopoulos A, Lee H and Gust ID (1992). **HTLV-I infection in selected populations in Australia and the western Pacific region.** *Medical Journal of Australia*, **156:** 878-880.

Okochi K, Sato H and Hinuma Y (1984). **A retrospective study on transmission of adult T-cell leukemia virus by blood transfusion: seroconversion in recipients.** *Vox Sang*, **46:** 245-253.

Osame M, Igata A, Matsumoto M, Usuku K, Izumo S and Kosaka K (1987). **HTLV-I-associated myelopathy: a report of 85 cases.** *Annals of Neurology,* **22:** 116.

Osame M, Usuku M, Izumo S, Ijichi N, Amitani H, Igata A, Matsumoto M and Tara M (1986). **HTLV-I-assicoated myelopathy, a new clinical entity.** *Lancet,* **1:** 1031-1032.

Paine E, Garcia J, Philpott TC, Shaw G and Ratner L (1991). **Limited sequence variation in human T-lymphotropic virus type I isolates from North American and African patients.** *Virology*, **182:** 111-123.

Papsidero L, Swartzwelder F, Sheu M, Montagna R, Ehrlich G, Bhagavati S, Dosik H, Sninsky J and Poiesz B (1990). **Immunodetection of human T-cell lymphotropic virus type I core protein in biological samples by using a monoclonal antibody assay.** *Journal of Clinical Microbiology,* **28:** 949-955.

Paquette Y, Hanna Z, Savard P, Broussean R, Robitaille Y and Jolicoeur P (1989). **Retrovirus-induced murine motor-neurone disease: mapping the determinant of spongiform degeneration within the envelope gene.** *Proceedings of the National Academy of Sciences, USA,* **86:** 3896-3900.

Pardue ML and Gall JG (1970). **Chromosomal localisation of mouse satellite DNA.** *Science,* **168:** 1356-1358.

Pique C, Torsz T and Dokelar M-C (1990). **Mutations induced along the HTLV-I envelope gene results in a non-functional protein: a basis for envelope conservation?** *EMBO Journal,* **7:** 4243-4248.

Poiesz B, Russeti F, Gazdar A, Bunn P, Mina J, Gallo R (1980). **Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of patients with cutaneous T-cell lymphoma.** *Proceedings of the National Academy of Sciences, USA*, **77:** 7415-7419.

Preston BD, Poiesz BJ and Loeb LA (1988). **Fidelity of reverse transcriptase.** *Science,* **242:** 1168-1172.

Ratner L and Poiesz BJ (1988). **Leukemias associated with human T-cell lymphotropic virus type I in a non-endemic region.** *Medicine*, **67:** 401-417.

Rodgers-Johnson P, Gajdusek DC, Morgan O, zaninovic V, Sarin PS and Graham DS (1985). **HTLV-I and HTLV-III antibodies and tropical spastic paraparesis.** *Lancet*, **2:** 1247-1248.

Roman GC and Osame M (1988). **Identity of HTLV-I associated tropical paraparesis and HTLV-I associated myelopathy.** *Lancet,* **1:** 651.

Roman GC and Roman LN (1988). **Tropical spastic paraparesis. A clinical study of 50 patients from Tumaco (Colombia) and review of the worldwide features of the**

syndrome. *Journal of Neurological Science,* **87:** 121-138.

Roman GC, Spencer PS, Schoenberg BS, Hugon J, Ludolph A, Rodgers-Johnson P, Osuntokun BO and Shamlaye CF (1987). **Tropical spastic paraparesis in the Seychelles islands. A clinical and case-control neuroepidemiologic study.** *Neurology,* **37:** 1323-1327.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB and Erlich HA (1988). **Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase.** *Science,* **239:** 487-491.

Saiki RK, Scharf S, Faloona FA, Mullis KB, Horn GT, Erlich HA and Arnheim N (1985). **Enzymatic amplification of B-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia.** *Science,* **230:** 1350-1354.

Saitou N and Imanishi T (1989). **Relative efficiencies of the Fitch-Margolish, maximum parsimony, maximum likelihood, minimum- evolution, and neighbour-joining methods of phylogenetic tree construction in obtaining the correct tree.** *Molecular Biology and Evolution,* **6:** 514-525.

Saitou N and Nei M (1987). **The neighbor-joining method: A new method for reconstructing phylogenetic trees.** *Molecular Biology and Evolution,* **4:** 406-425.

Saksena NK, Herve V, Sherman MP, Duran JP, Mathiot C, Muller M, Love JL, LeGuenno B, Barre-Sinoussi F, Dube DK and Poeisz BJ (1993). **Sequence and phylogenetic analyses of a new STLV-I from a naturally infected tantalus monkey from Central Africa.** *Virology,* **192:** 312-320.

Saksena NK, Sherman MP, Yanagihara R, Dube DK and Poiesz BJ (1992). **LTR sequence and phylogenetic analysis of a newly discovered variant of HTLV-I from the Hagahai of Papua New Guinea.** *Virology,* **189:** 1-9.

Sandler SG, Fang CT and Williams AE (1989). **Blood transfusion and HTLV-I.** In: (GC Roman, JC Vernant and M Osame eds) *HTLV-I and the Nervous System*, pp. 541-546, Liss, New York.

Saxton E, Lee H, Swanson P, Chem IS, Ruland C, Chin E, Aboulafia D, Delamarter R, Rosenblatt J (1989). **Detection of Human T-cell leukemia and lymphoma virus type I in a transfusion recipient with chronic myelopathy.** *Neurology*, **39:** 841-844.

Schooley RT (1995). **Correlation between viral load measurements and outcome in clinical trails of antiviral drugs.** *AIDS Journal,* **9:** s15-s19.

Schupbach J, Kalyanaraman VS, Sarngadharan MG, Blattner WA and Gallo RC (1983). **Antibodies against three purified proteins of the human type C retrovirus, human T-cell leukemia-lymphoma virus, in adult T-cell leukemia-lymphoma patients and healthy blacks from the Caribbean.** *Cancer Research,* **43:** 886-891.

Seiki M, Hattori S, Hirayama Y and Yoshida M (1983). **Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA.** *Proceedings of the National Academy of Sciences, USA,* **80:** 3618-3622.

Seiki M, Hattori S and Yoshida M (1982). **Human adult T-cell leukemia virus: molecular cloning of the provirus and the unique terminal structure.** *Proceedings of the National Academy of Sciences,* **79:** 6899-6902.

Sherman MP, Saksena NK, Dube DK, Yanagihara R and Poiesz BJ (1992). **Evolutionary insights on the origin of HTLV-I derived from sequence analysis of a new HTLV-I variant from Papua New Guinea.** *Journal of Virology,* **66:** 2556-2563.

Shibasaki H, Endo C, Kuroda Y, Kakigi R, Oda K and Komine S (1988). **Clinical picture of HTLV-I associated myelopathy.** *Journal of Neurological Sciences,* **87:** 15-24.

Song K-J, Nerurkar VR, Saitou N, Lazo A, Blakeslee JR, Miyoshi I and Yanagihara R (1994). **Genetic analysis and molecular phylogeny of simian T-cell lymphotropic virus type I: evidence for independent virus evolution in Asia and Africa.** *Virology,* **199:** 56-66.

Southern EM (1975). **Detection of specific sequences among DNA fragments separated by gel electrophoresis.** *Journal of Molecular Biology,* **98:** 503-517.

Sugiyama H, Boi H, Yamaguchi K, Tsuji Y, Miyamoto T and Hinas S (1986). **Significance of postnatal mother to child transmission of human T-cell leukemia/lymphoma virus type 1 on the development of adult T-cell leukemia/lymphoma.** *Journal of Medical Virology*, **20:** 253-258.

Tajima K, Ito S-I and Tsushima ATL Study Group (1990). **Prospective studies of HTLV-I and associated diseases in Japan.** In: (Blattner WA [ed]) *Human Retrovirology: HTLV,* New York, Raven Press, **pp** 267-279.

Takatsuki K, Uchiayama T, Sagawa K and Yodoi J (1977). **Adult T-cell leukemia in Japan.** In: (Seno S, Takaku F and Irino S [eds]) *Topics in Hematology,* Excerpta Medica, Amsterdam, **pp** 73-77.

Takeuchi K, Kobayashi N, Nam SH, Yamamoto N and Hatanaka M (1985). **Molecular cloning of cDNA encoding gp68 of adult T-cell leukaemia-associated antigen: Evidence for expression of the pX IV region of human T-cell leukaemia virus.** *Journal of General Virology,* **66:** 1825-1829.

Tara M, Tokunaga M, Osame M (1989). **Neurological complications of adult T-cell leukemia/lymphoma.** In: (Roman GC, Vernant J-C and Osame M [eds]) *HTLV-I and the Nervous System*, New York, Liss, **pp** 73-82.

Tsujimoto A, Teruuchi T, Imamura J, Shimotohno K, Miyoshi I and Miwa M (1988). **Nucleotide sequence analysis of a provirus derived from HTLV-I-associated**

myelopathy (HAM). *Molecular Biology and Medicine,* **5:** 29-42.

Vallejo A and Garcia-Saiz A (1995). **Typing human T-cell lymphotropic virus (HTLV-I and HTLV-II) by nested polymerase chain reaction: application to clinical specimens.** *Journal of Virological Methods,* **51:** 9-18.

Vandamme A-M, Lui H-F, Goubau P and Desmyter J (1994). **Primate T-lymphotropic virus type I sequence variation and its phylogenetic analysis: compatibility with an African origin of PTLV-I.** *Virology,* **202:** 212-223.

Varmus H (1988). **Retroviruses.** *Science,* **240:** 1427-1435.

Verdier M, Denis F, Léonard G, Sangare A, Patillaud S, Prince-David M and Essex M (1990). **Comparison of immunofluorescence, particle agglutination, and enzyme immunoassays for detection of human T-cell leukemia virus type I antibody in African sera.** *Journal of Clinical Microbiology,* **28:** 1988-1993.

Van der Peer Y and De Wachter R (1993). **TREECON: a software package for the construction and drawing of evolutionary trees.** *Computer and Applied Biosciences,* **9:** 177-182.

Vidal AU, Gessain A, Yoshida M, Tekaia F, Garin B, Guillemain B, Schulz T, Farid R and de The G (1994). **Phylogenetic classification of human Tcell leukaemia/lymphoma virus type I genotypes in five major molecular and geographical subtypes.** *Journal of*

General Virology, **74:** 3655-3666.

Wang C-NJ, Wu KY and Wang H-T (1995). **Quantitative PCR using the AmpliSensor assay.** In: (Dieffenbach CW and Dveksler GS [eds]) *PCR Primer: A Laboratory Manual*, Cold Spring Harbour Laboratory Press, New York, **pp** 193-202.

Watanabe T, Seiki M, Tsujimoto H, Miyoshi I, Hayami M and Yoshida M (1985). **Sequence homology of the simian retrovirus genome with human T-cell leukemia virus type I.** *Virology*, **144:** 59-65.

Wiktor SZ, Cannon RO and Atkinson WL (1992). **Infection with HTLV types I and II in sexually transmitted disease clinics in Baltimore and New Orleans.** *Journal of Infectious Diseases*, **165:** 920-924.

Williams AE, Fang CT, Slamon DJ, Poiesz BJ, Sandler SG, Darr WF, Shulman G, MoGowan EI, Douglas DK, Bowman RJ, Peetom F, Kleirman SH, Lenes B and Dodd RY (1988). **Serolprevalence and epidemiological correlates of HTLV-I infection in U.S. blood donors.** *Science*, **240:** 643-646.

World Health Organisation (1989). **Report. Scientific Group on HTLV-I Infections and Associated diseases, Kagoshima, Japan, 10-15 December 1988.** WHO Regional Office for the Western Pacific, Manila, Philippines.

Wucherpfennig KW, Hollsberg P, Richardson JH, Benjamin D and Hafler DA (1992). **T-cell activation by autologous human T-cell leukemia virus type I-infected T-cell clones.** *Proceedings of the National Academy of Sciences, USA,* **89:** 2110-2114.

Yanagihara R (1994). **Geographic-specific genotypes of Human T-cell lymphotropic virus type I as markers for early and recent migrations of human populations.** *Advances in Viral Research,* **43:** 147-186.

Yanagihara R, Hefner C and Ajdukiewicz AB (1992). **Human T lymphotropic virus type I infection among blood donors in the Solomon Islands.** *Transfusion,* **32:** 89.

Yanagihara R, Nerurkar VR, Garruto RM, Miller MA, Leon-Monzon ME, Jenkins CL, Sanders RC, Liberski PP, Alpers MP, and Gajdusek DC (1991). **Characterisation of a variant of human T lymphotropic virus type I isolated from a member of a remote, recently contacted group in Papua New Guinea.** *Proceedings of the National Academy of Sciences, USA,* **88:** 1446-1450.

Yanagihara R, Garruto RM, Miller MA, Leon-Monzon ME, Liberski PP, Gajdusek DC, Jenkins CL, Sanders RC and Alpers MP (1990). **Isolation of HTLV-I from members of a remote tribe in New Guinea.** *New England Journal of Medicine,* **323:** 993-994.

Yoshida M (1989). **Molecular biology of HTLV-I from adult T-cell leukemia (ATL) and chronic myelopathy (TSP/HAM).** In: (Roman GC, Vernant and Osame M [eds]) *HTLV-I and the Nervous System,* New York, Liss, **pp** 19-29.

Yoshida M, Seiki M, Yamaguchi K and Takasuki K (1984). **Monoclonal intergration of human T-cell leukemia provirus in all primary tumors of adult T-cell leukemia suggests a causative role of human T-cell leukemia virus in the disease.** *Proceedings of the National Academy of Sciences, USA,* **81:** 2534-2537.

Yoshida M, Miyoshi I and Hinuma Y (1982). **Isolation and characterisation of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease.** *Proceedings of the National Academy of Sciences, USA,* **79:** 2031-2035.

Zhao T-M, Robinson MA, Sawasdikosol S and Kindt TJ (1993). **Variation in HTLV-I sequences from rabbit cell lines with diverse in vivo effects.** *Virology,* **195:** 271-274.

Zharkikh A and Li W-H (1992). **Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock.** *Molecular Biology and Evolution,* **9:** 1119-1147.

Additional References

Botha MC, Jones M, De Klerk WA and Yamamoto N (1985). **Distribution and possible spread of HTLV-I in human communities in the northern and eastern Transvaal.** *South African Medical Journal,* **67:** 668-671.

Mahieux R, Ibrahim F, Mauclere P, Herve V, Michel P, Tekaja F, Chappey C, Garin B, van der Ryst E, Guillemain B, Ledru E, Delaporte E, De The G and Gessain A (1997). **Molecular epidemiology of 58 new African human T-cell leukemia virus type I (HTLV-I) strains: identification of a new and distinct HTLV-I molecular subtype in central Africa and in pygmies.** *Journal of Virology,* **71:** 1317-1333.

van der Ryst E, Smith MS and Visagie HMM (1996 [a]). **Comparison of the polymerase chain reaction and serology for the diagnosis of HTLV-I infection.** *Journal of Infection,* **32:** 109-112.

van der Ryst E, Joubert G, Smith MS, Terblanche M, Mollentze F and Pretorius AM (1996 [b]). **HTLV-I infection in the Free State region of South Africa: a sero-epidemiologic study.** *Central African Journal of Medicine,* **42:** 65-68.

Appendix I

Reagents

AMP: 5 g of ampicillin was reconstituted with 5 ml of water to give a stock solution of 100 mg/ul. The solution was aliquoted and stored at -20°C.

Boehringer 10 X *Taq* polymerase buffer (with Mg^{2+}): 15 mM Mg^{2+}, 500 mM KCl, Tris-HCl, pH 8.3.

dNTPs: 20 ul of each nucleotide (Pharmacia, 100 mM concentration) was diluted with 720 ul of water to give a final concentration of 2.5 mM.

0.5 M EDTA: 186.1 g of disodium ethylenediaminetetra-acetate H_2O) was dissolved in 800 ml of water by stirring on a magnetic stirrer. The pH was adjusted to 8.0 with NaOH pellets. The volume was adjusted to 100 ml with water.

70% ethanol: 70 ml of absolute ethanol was diluted with 30 ml of distilled water.

95% ethanol: 95 ml of absolute ethanol was diluted with 5 ml of distilled water.

1 M Mg^{2+}: 203.3 g of $MgCl_3 6H_2O$ was disssolved in 800 ml of water. The volume was adjusted to one litre with water. The solution was sterilised by autoclaving.

5 M K acetate: 49.1 g of potassium acetate was dissolved in 90 ml of distilled water. The pH was adjusted to 7.5 with 2 M acetic acid. The volume was adjusted to 100 ml with distilled water. The solution was stored at 4°C.

5 M NaCl: 292.2 g of NaCl was dissolved in 800 ml of water. The volume was adjusted to one litre with water. The solution was sterilised by autoclaving.

1 M tris: 121.1 g of tris base was dissolved in 800 ml of water. The pH was adjusted to pH 8.0 by adding concentrated HCl. The solution was allowed to cool to room temperature before making final adjustments to the pH with concentrated HCl. The volume was adjusted to one litre with water and sterilised by autoclaving.

TE buffer: 1 ml of 1 M tris HCl (pH 8.0), and 2 ml of 0.5 M EDTA was diluted to 100 ml with water (1 mM tris HCl and 1 mM EDTA final concentration).

Appendix II

Sequence Homology of the 410 bp *env* Region with the Japanese Prototype Sequence ATK-1 (J02029)

```
                                                  Match%  Over.  INIT   OPT
                                                    98.5   412   1120   1596

                10         20         30         40         50         60
CJ9.SEQ   GGATGACTCAGGGTTTATAAGAGAGTAATGGGGGTATCTGACGCGCGAGGGGAGGTGTCG
          X:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
CJ8.SEQ   GGATGACTCAGGGTTTATAAGAGAGTAATGGGGGTATCTGACGCGCGAGGGGAGGTGTCG
                70         80         90        100        110        120
                  70         80         90        100        110        120
CJ9.SEQ   TAGCTGACGGAGGATGCATGGTCCTGCAAGGATAACAAGAAGGAGTAGCGCGACAAGGGT
          :::::::::::::::::::::::::::::: ::::::::::::::::::::::::::::::
CJ8.SEQ   TAGCTGACGGAGGATGCATGGTCCTGCAA-GATAACAAGAAGGAGTAGCGCGACAAGGGT
               130        140        150        160        170        180
                 130        140        150        160        170        180
CJ9.SEQ   GATTCCAGTTTGTAAGGCCTCTCGAGCCCACTGTGAGAGGCCAAGGTCCCAGTTAAGGCC
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
CJ8.SEQ   GATTCCAGTTTGTAAGGCCTCTCGAGCCCACTGTGAGAGGCCAAGGTCCCAGTTAAGGCC
                190        200        210        220        230        240
                 190        200        210        220        230
CJ9.SEQ   CCAGCCAGTCAGGACTCGATTCTCAA-GGGGGGGTCTTTCTTGTAGTATTGAGACATGGG
          :::::::::::::::::::::::::: ::::::::::::::::::::::::::::::::
CJ8.SEQ   CCAGCCAGTCAGGACTCGATTCTCAAGGGGGGGGTCTTTCTTGTAGTATTGAGACATGGG
                250        260        270        280        290        300
        240        250        260        270        280        290
CJ9.SEQ   AATTAGTAATATTCAGAAAACAGCACTGTTCTTGTAATGCTTTGCATAATCCTCCTTGTT
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
CJ8.SEQ   AATTAGTAATATTCAGAAAACAGCACTGTTCTTGTAATGCTTTGCATAATCCTCCTTGTT
                310        320        330        340        350        360
        300        310        320        330        340        350
CJ9.SEQ   CCCAGAACAGGAGATCAAGGCCTCGTCTGTTCTGGGCAGCACACTGCGCAATTTTGAGTA
          ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
CJ8.SEQ   CCCAGAACAGGAGATCAAGGCCTCGTCTGTTCTGGGCAGCACACTGCGCAATTTTGAGTA
                370        380        390        400        410        420
        360        370        380        390        400        410
CJ9.SEQ   GATTTTTGTGGTTTTTGACTATTGCTTGAGAATCGGATCCCCGGTACCGAGC
          :::::::::::::::::.:::::::::::::::::::::::::::::    ::::X
CJ8.SEQ   GATTTTTGTGGTTTTTGMCTATTGCTTGAGAATCGGATCCCCGGGTACGAGC
                430        440        450        460        470
```

Appendix III

Raw Data from the Quantitative AmpliSensor Assay.

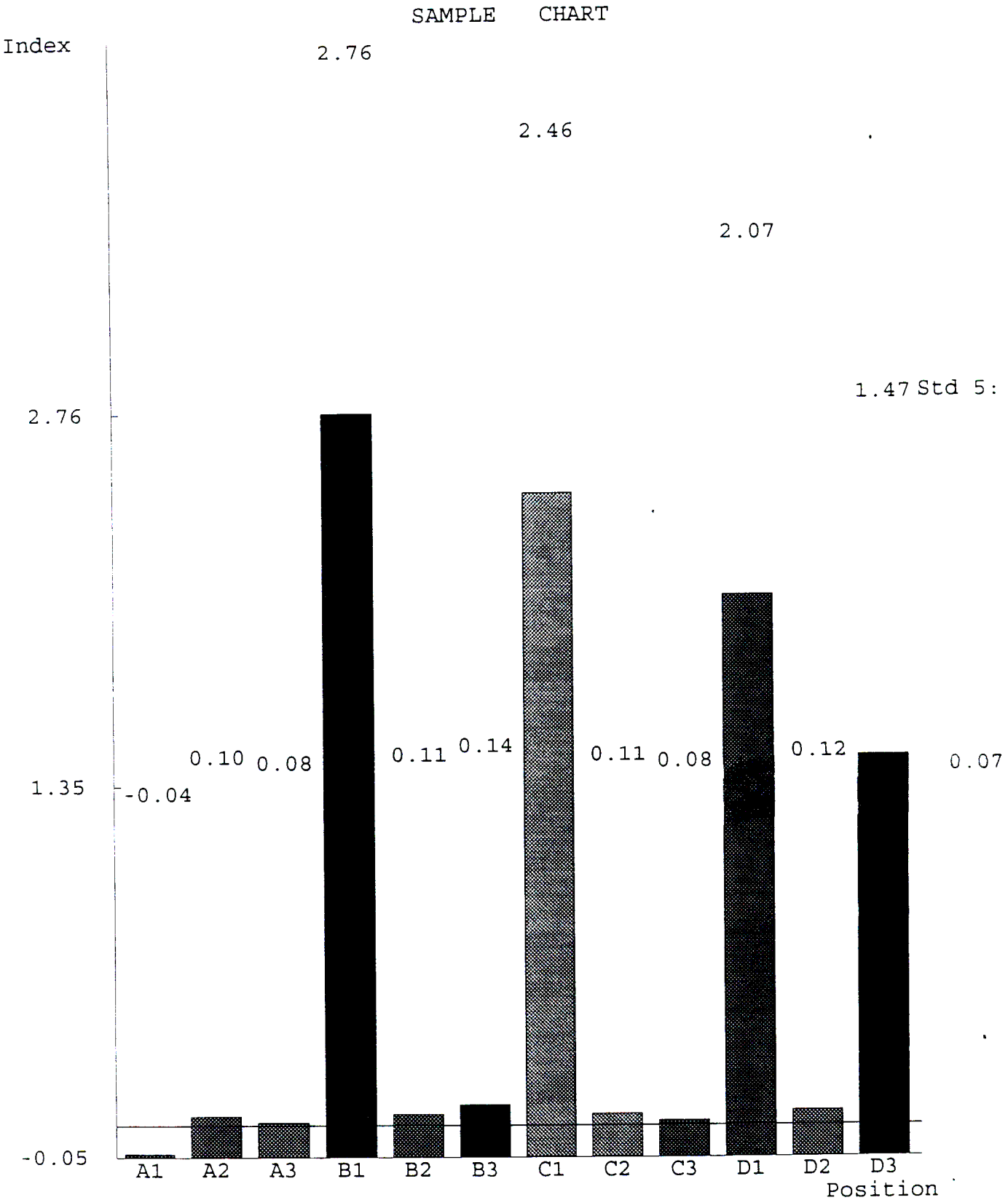
SUMMARY REPORT (Template: 0 Dilution Factor: 1.666667)

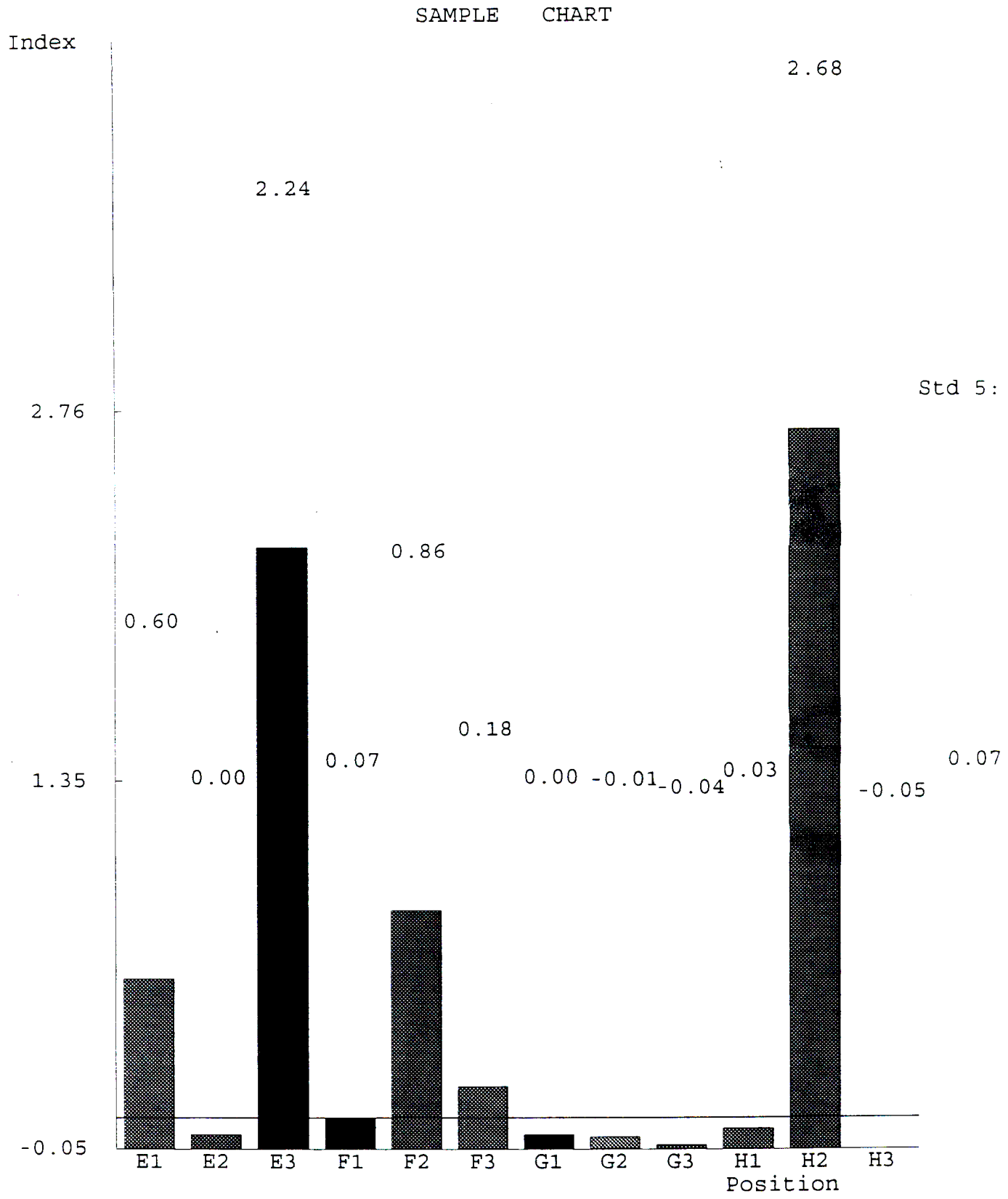
Well Position	Quantity AVG.	Quantity STD.	Titers	Label	Status
B1	+5.74e+003	+2.26e+003		Std 1	+
C1	+1.35e+003	+5.76e+002		Std 2	+
D1	+8.73e+002	-5.00e-003		Std 3	+
E1	+1.44e+002	+9.89e+001		Std 4	+
F1	+9.89e+000	-5.00e-003		Std 5	?
A2	+1.63e+001	+5.91e-001	+2.71e+001		+
A3	+7.04e+000	-5.00e-003	+1.17e+001		+
B2	+1.80e+001	+9.36e-001	+3.00e+001		+
B3	+2.30e+001	+4.72e+000	+3.83e+001		+
C2	+2.02e+001	+3.11e+000	+3.37e+001		+
C3	+1.84e+001	-5.00e-003	+3.06e+001		+
D2	+1.61e+001	+1.35e+000	+2.68e+001		+
D3	+3.08e+002	+4.72e+001	+5.13e+002		+
E3	+1.10e+003	+1.23e+002	+1.83e+003		+
F2	+1.84e+002	+1.07e+002	+3.07e+002		+
F3	+2.76e+001	+7.76e+000	+4.60e+001		+
H2	+3.01e+003	+1.21e+001	+5.02e+003		+

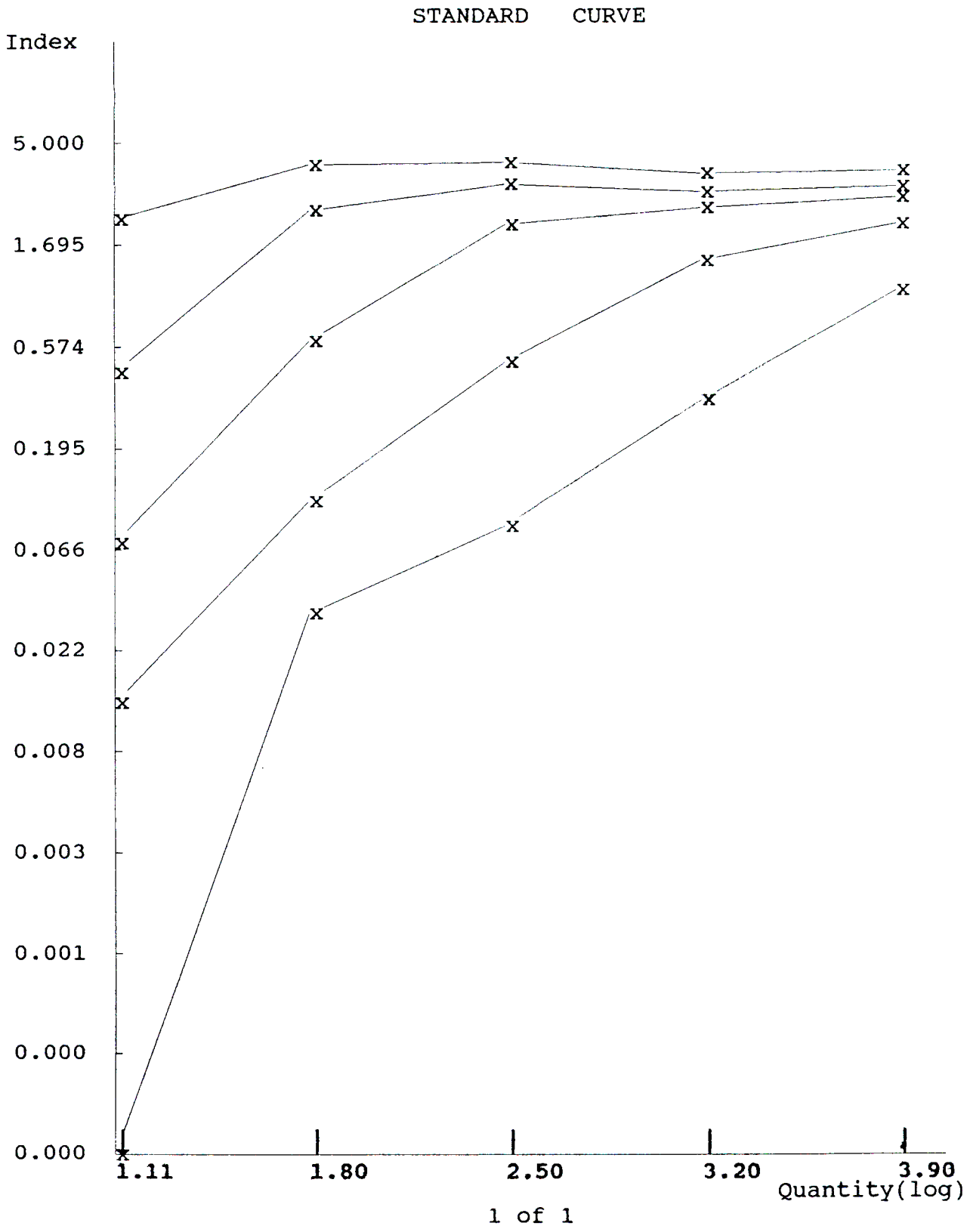
Samples out of the quantitation range :

G1	Apex
G2	Blank
A1	Neg 1
H1	Neg 2
E2	
G3	
H3	

Note: * denotes sample repeat;
( ) denotes ICP aberrant.


SAMPLE CHART
Index
2.76
2.46
2.07
1.47 Std 5:
2.76
1.35
-0.05
-0.04
0.10
0.08
0.11
0.14
0.11
0.08
0.12
0.07
A1
A2
A3
B1
B2
B3
C1
C2
C3
D1
D2
D3
Position


SAMPLE CHART
Index
2.68
2.24
Std 5:
2.76
0.86
0.60
0.18
0.07
0.07
1.35
0.00
0.00
-0.01
-0.04
0.03
-0.05
-0.05
E1
E2
E3
F1
F2
F3
G1
G2
G3
H1
H2
H3
Position


STANDARD CURVE
Index
5.000
1.695
0.574
0.195
0.066
0.022
0.008
0.003
0.001
0.000
0.000
1.11
1.80
2.50
3.20
3.90
Quantity(log)
1 of 1

Appendix IV

Primer Sequences and Dilutions

Primer were designed from the Japanese prototype sequence (Genebank accession no: J02029).

9 (forward): 5' TTG GAG TGT AGT TCT GAC 3'

OD: 25.3; length: 18; dilution: 1 ul primer + 5.39 ul water

2016 (reverse): 5' CTG GTT CTG GGA TAG TGG 3'

OD: 20.9; length: 18; dilution: 1 ul primer + 4.27 ul water

1985 (forward): 5' CAC TGG AGG CGA GAC TGC 3'

OD: 17.2; length: 18; dilution: 1 ul primer + 3.34 ul water

4118 (reverse): 5' CGA AGG TAA TGA TAA AGA 3'

OD: 18.45; length: 18; dilution: 1 ul primer + 3.66 ul water

3921 (forward): 5' GCC CCC TGC CTG TTT TCA 3'

OD: 12.95; length: 18; dilution: 1 ul primer + 2.27 ul water

6666 (reverse): 5' ATT GCG TGC TTG GTT TAC 3'

OD: 18.7; length: 18; dilution: 1 ul primer + 3.72 ul water

4239 (forward): 5' CGC AGC CAT ACC AAT CTA 3'

OD: 13; length: 18; dilution: 1 ul primer + 2.28 ul water

6188 (reverse): 5' TAA TCC CGC CAG CCA CTC 3'

OD: 16.15; length: 18; dilution: 1 ul primer + 3.07 ul water

6275 (forward): 5' CTC AAG CAA TAG TCA AAA 3'

OD: 22.3; length: 18; dilution: 1 ul primer + 4.63 ul water

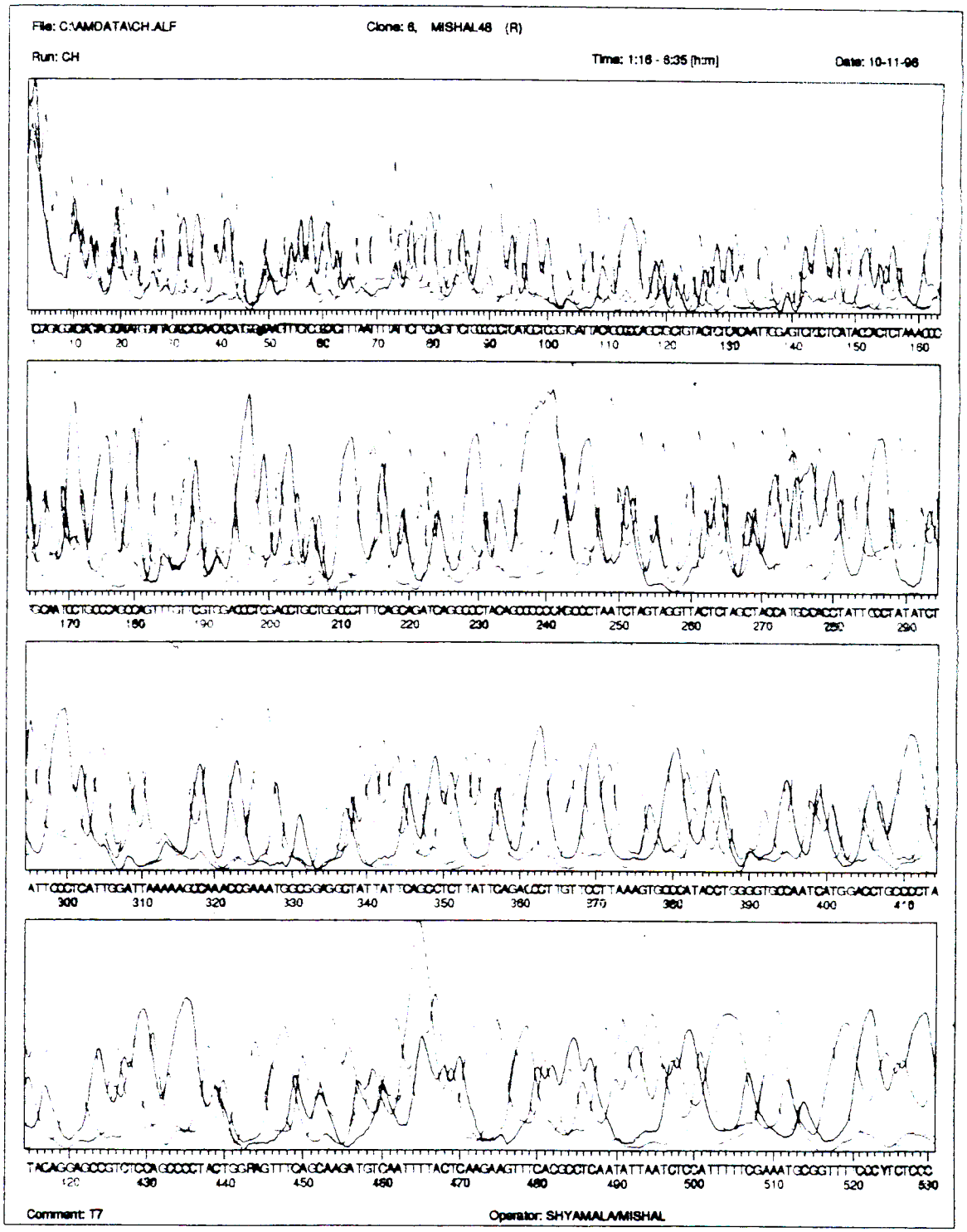
8055 (reverse): 5' GTA AAT GTC CAA ATA AGG 3'

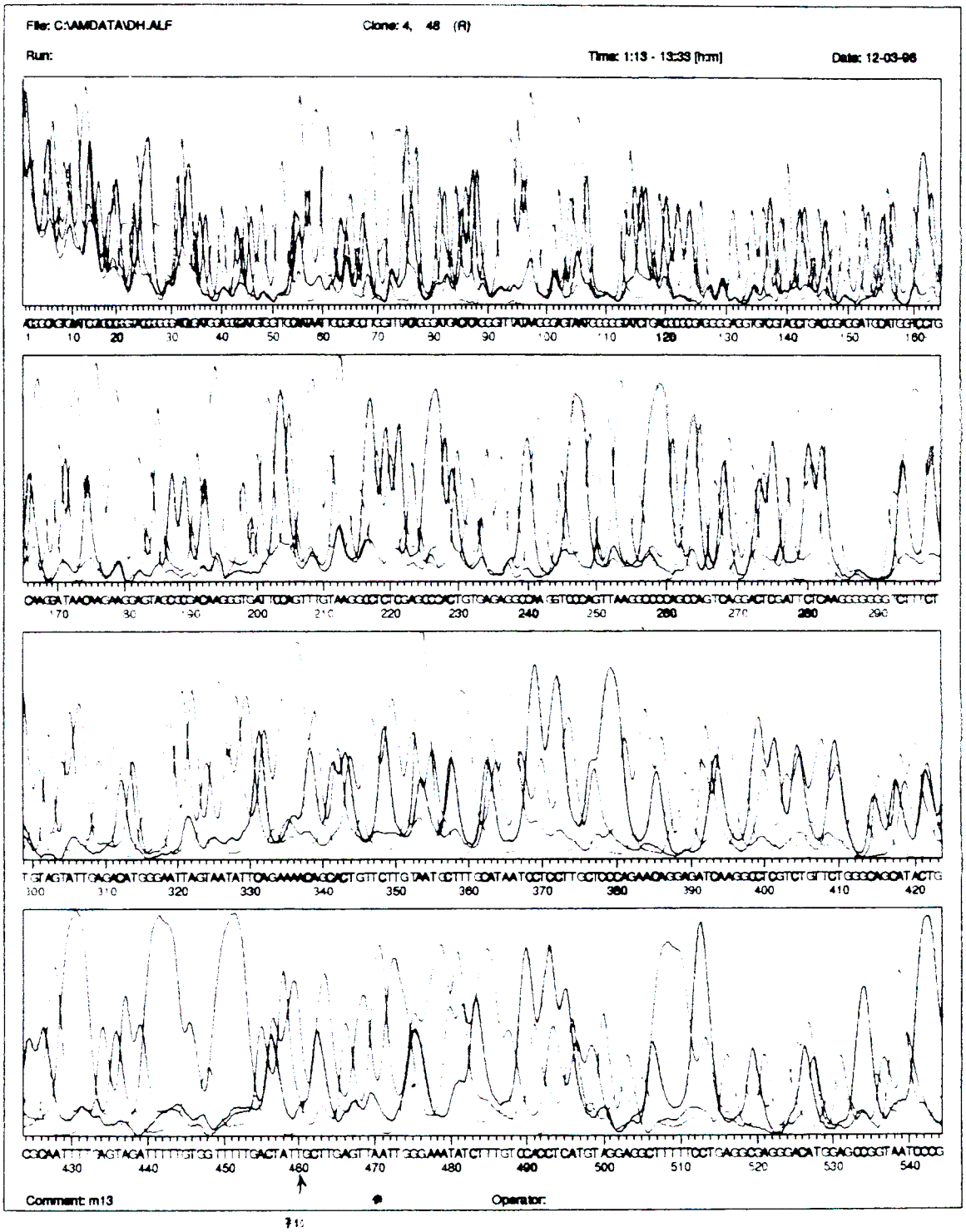
OD: 22.8; length: 18; dilution: 1 ul primer + 4.76 ul water

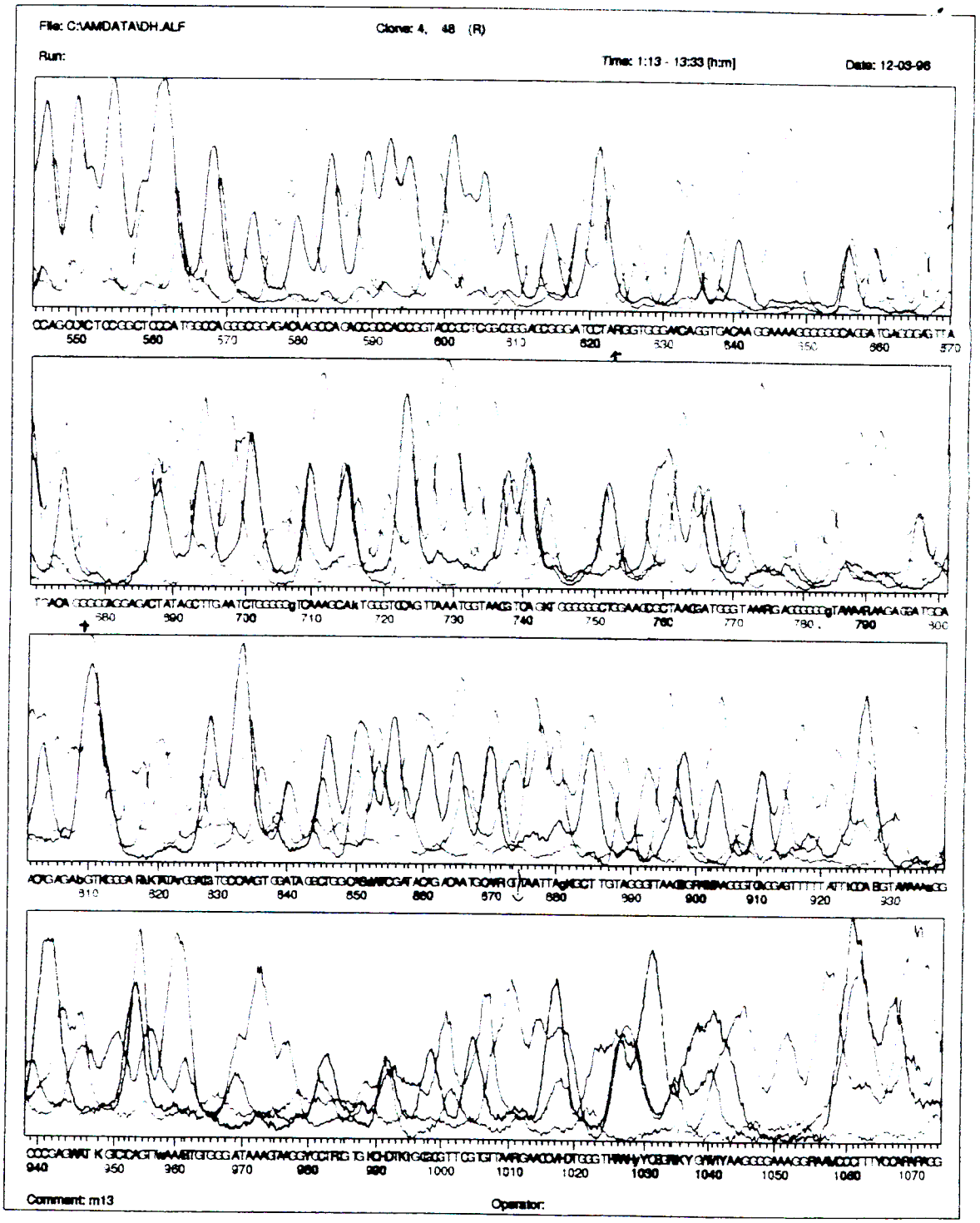
Appendix V

Fluorographs Obtained from the A.L.F Express and ABI Automated Sequencers for the Five Isolates from KZN.

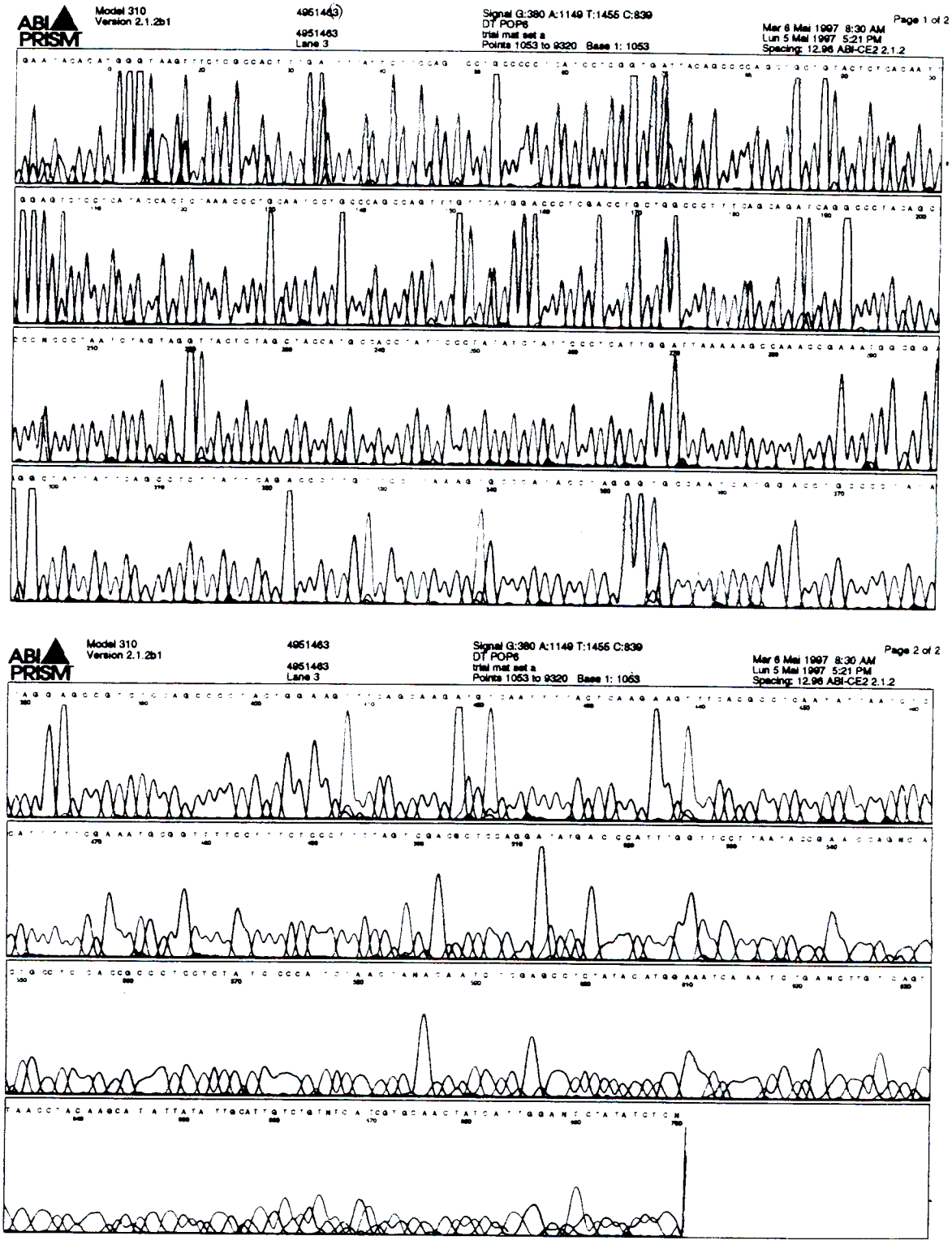
i) 48

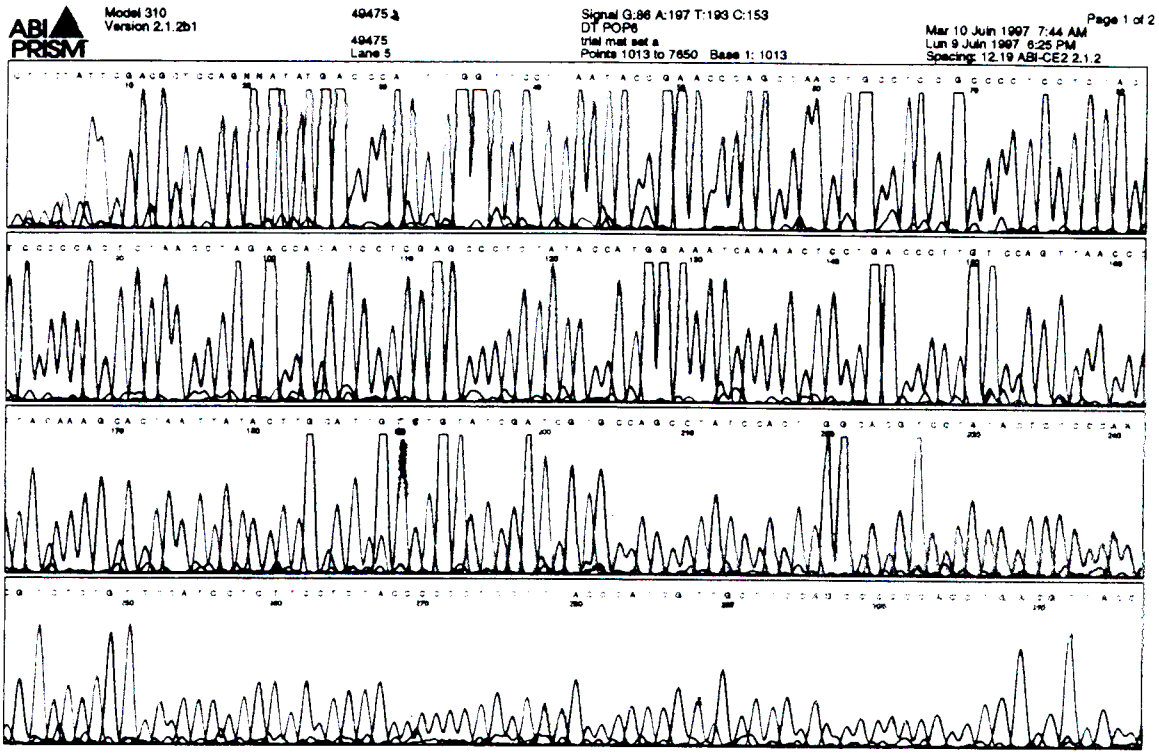


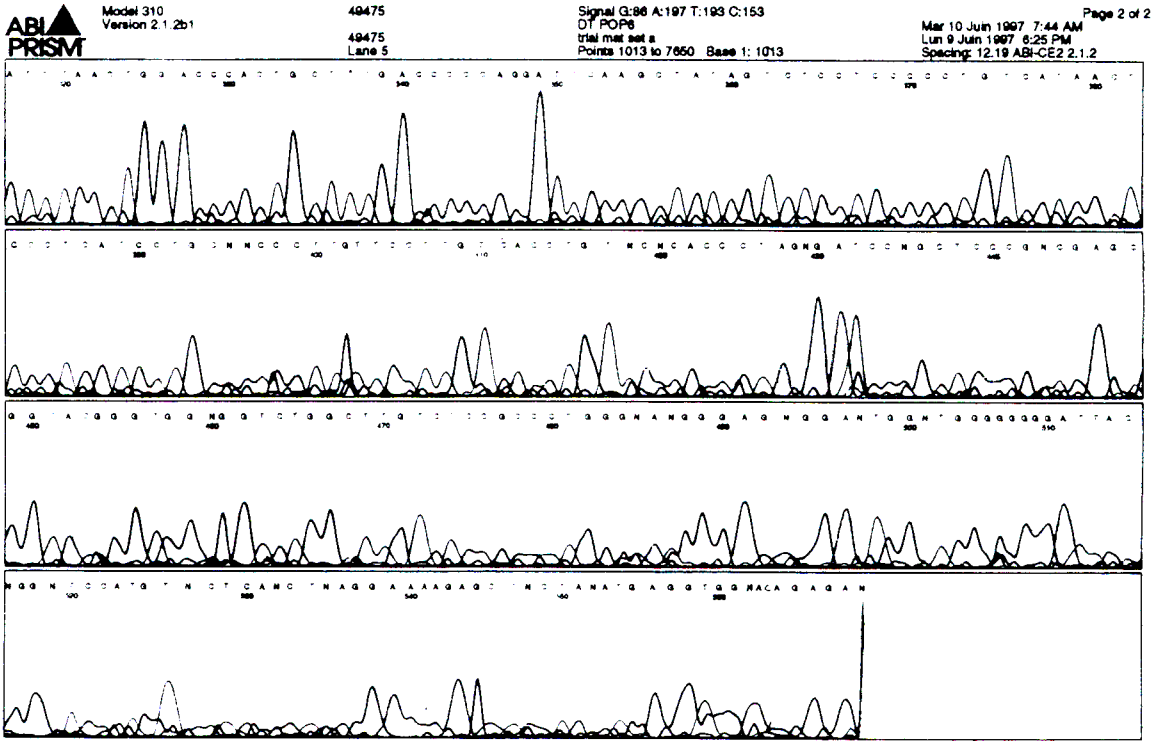

File: C:\AMDATA\DH.ALF
Clone: 4, 48 (R)
Run:
Time: 1:13 - 13:33 [h:m]
Date: 12-03-96
Comment: m13
Operator:

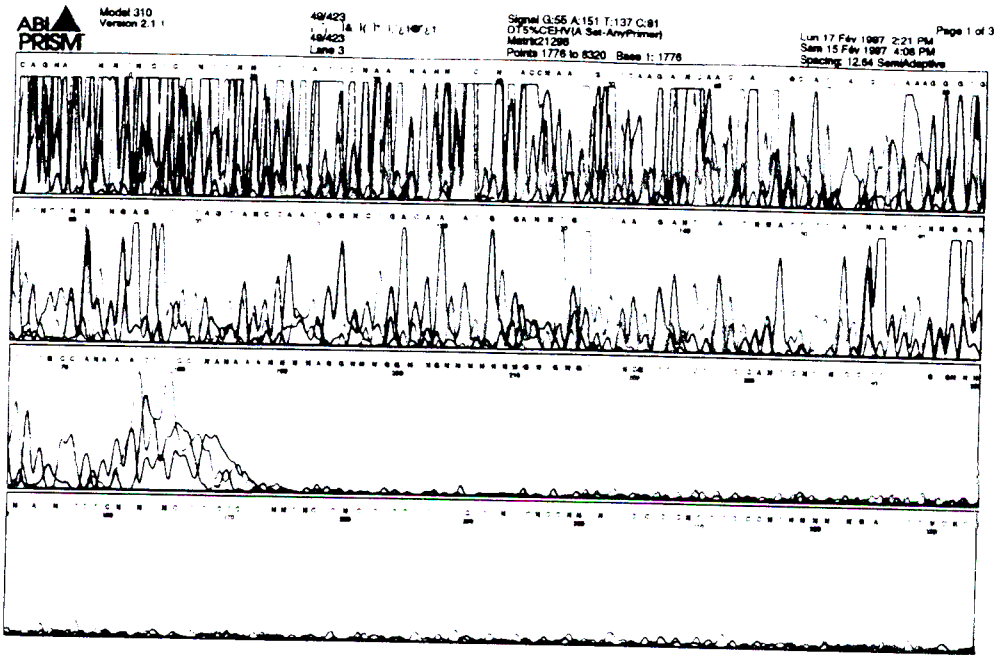

File: C:\AM\DATA\DH.ALF
Clone: 4, 48 (R)
Run:
Time: 1:13 - 13:33 [h:m]
Date: 12-03-96
Comment: m13
Operator:

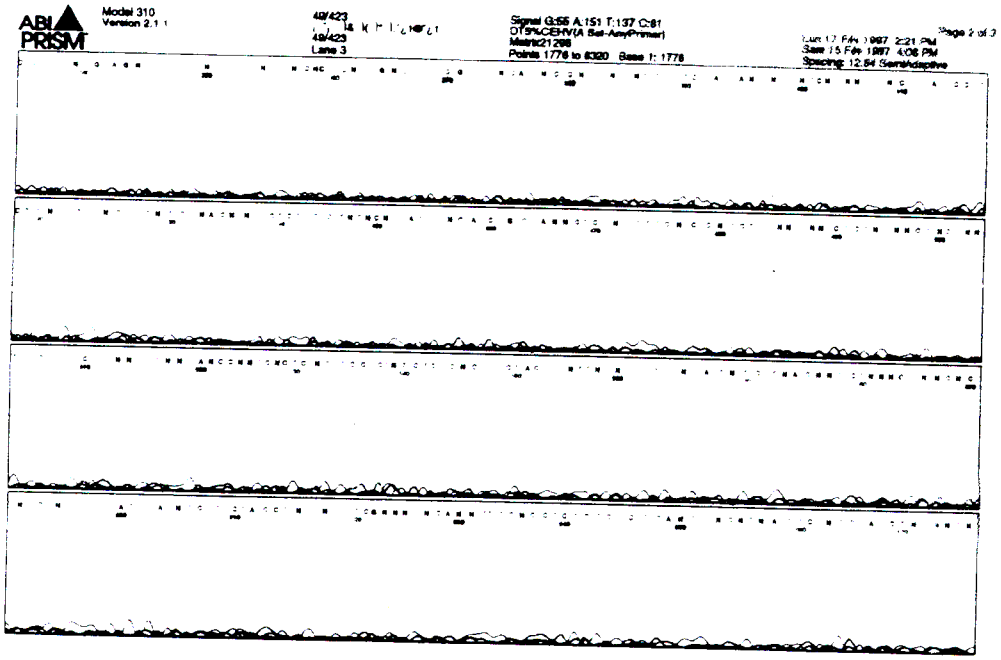
ii) 49

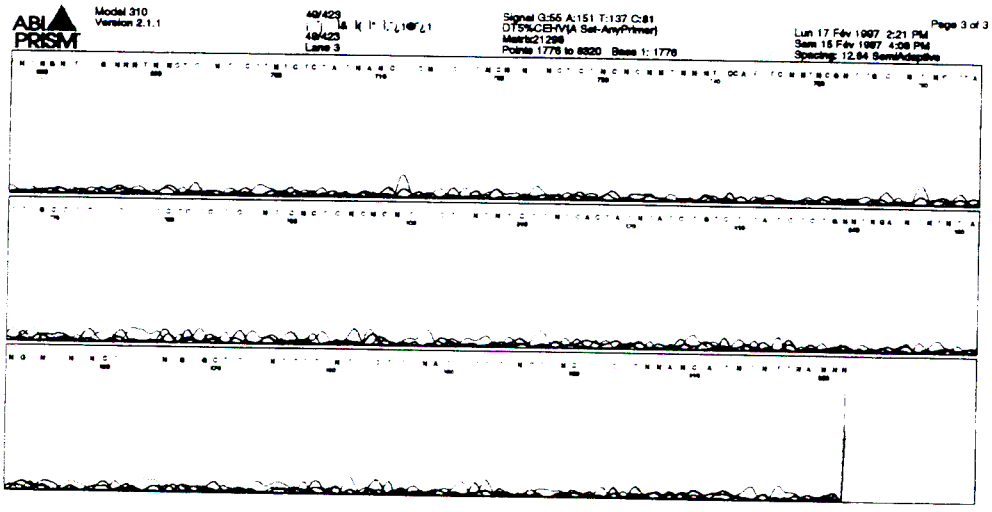

ABI PRISM
Model 310
Version 2.1.2b1
4951463
4951463
Lane 3
Signal G:380 A:1149 T:1455 C:839
DT POP6
trial mat set a
Points 1053 to 9320 Base 1: 1053
Mar 6 Mai 1997 8:30 AM
Lun 5 Mai 1997 5:21 PM
Spacing: 12.96 ABI-CE2 2.1.2
Page 1 of 2
ABI PRISM
Model 310
Version 2.1.2b1
4951463
4951463
Lane 3
Signal G:380 A:1149 T:1455 C:839
DT POP6
trial mat set a
Points 1053 to 9320 Base 1: 1053
Mar 6 Mai 1997 8:30 AM
Lun 5 Mai 1997 5:21 PM
Spacing: 12.96 ABI-CE2 2.1.2
Page 2 of 2


ABI PRISM
Model 310
Version 2.1.2b1
49475
49475
Lane 5
Signal G:86 A:197 T:193 C:153
DT POP6
trial mat set a
Points 1013 to 7650 Base 1: 1013
Page 1 of 2
Mar 10 Juin 1997 7:44 AM
Lun 9 Juin 1997 6:25 PM
Spacing: 12.19 ABI-CE2 2.1.2


ABI PRISM
Model 310
Version 2.1.2b1
49475
49475
Lane 5
Signal G:86 A:197 T:193 C:153
DT POP6
trial mat set a
Points 1013 to 7650 Base 1: 1013
Page 2 of 2
Mar 10 Juin 1997 7:44 AM
Lun 9 Juin 1997 6:25 PM
Spacing: 12.19 ABI-CE2 2.1.2

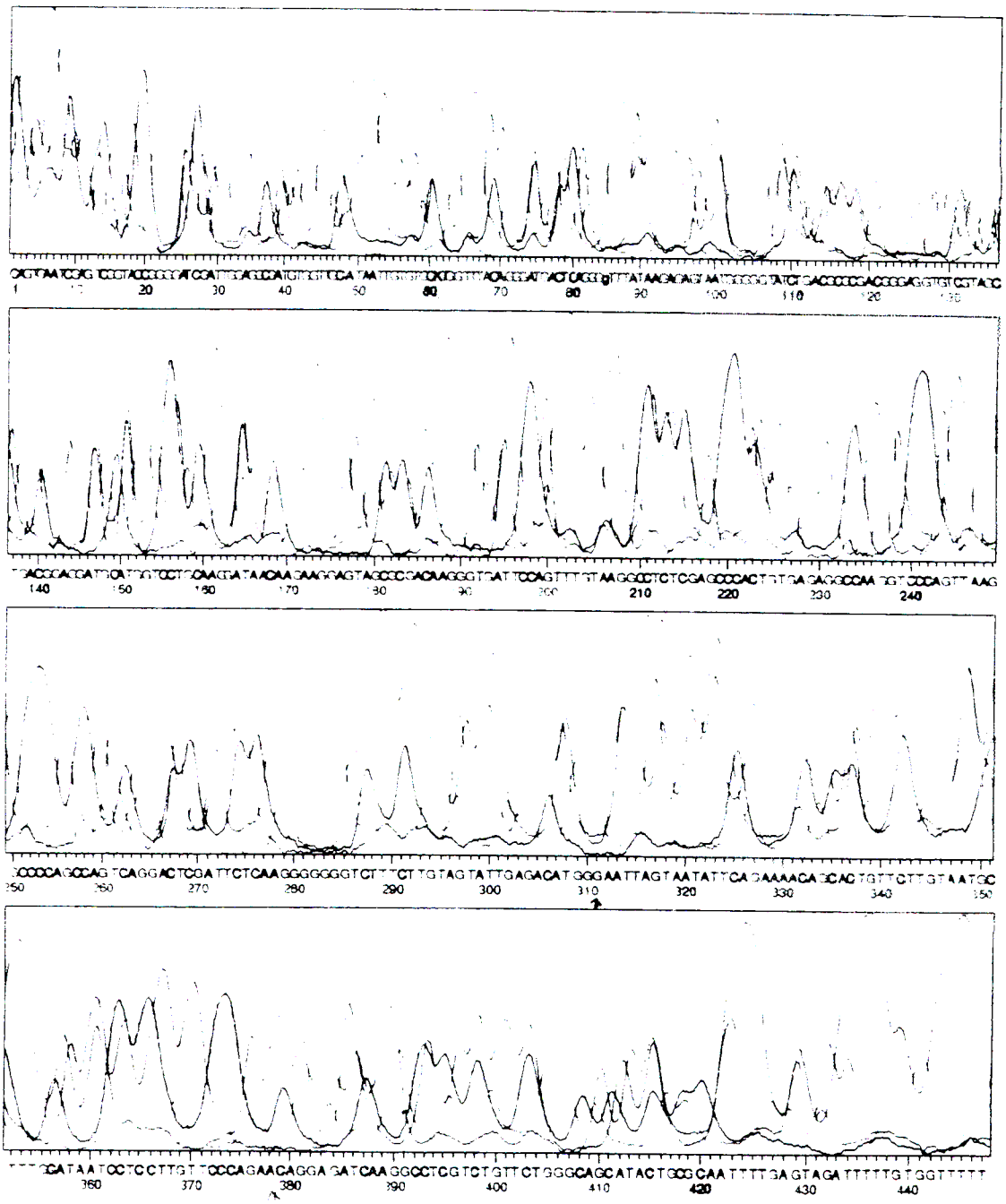

ABI PRISM
Model 310
Version 2.1.1
Signal G:55 A:151 T:137 C:81
Page 1 of 3
Lun 17 Fév 1997 2:21 PM
Sam 15 Fév 1997 4:08 PM
Spacing: 12.84 Semi/Adaptive
Lane 3


ABI PRISM
Model 310
Version 2.1.1
Signal G:55 A:151 T:137 C:81
Lane 3

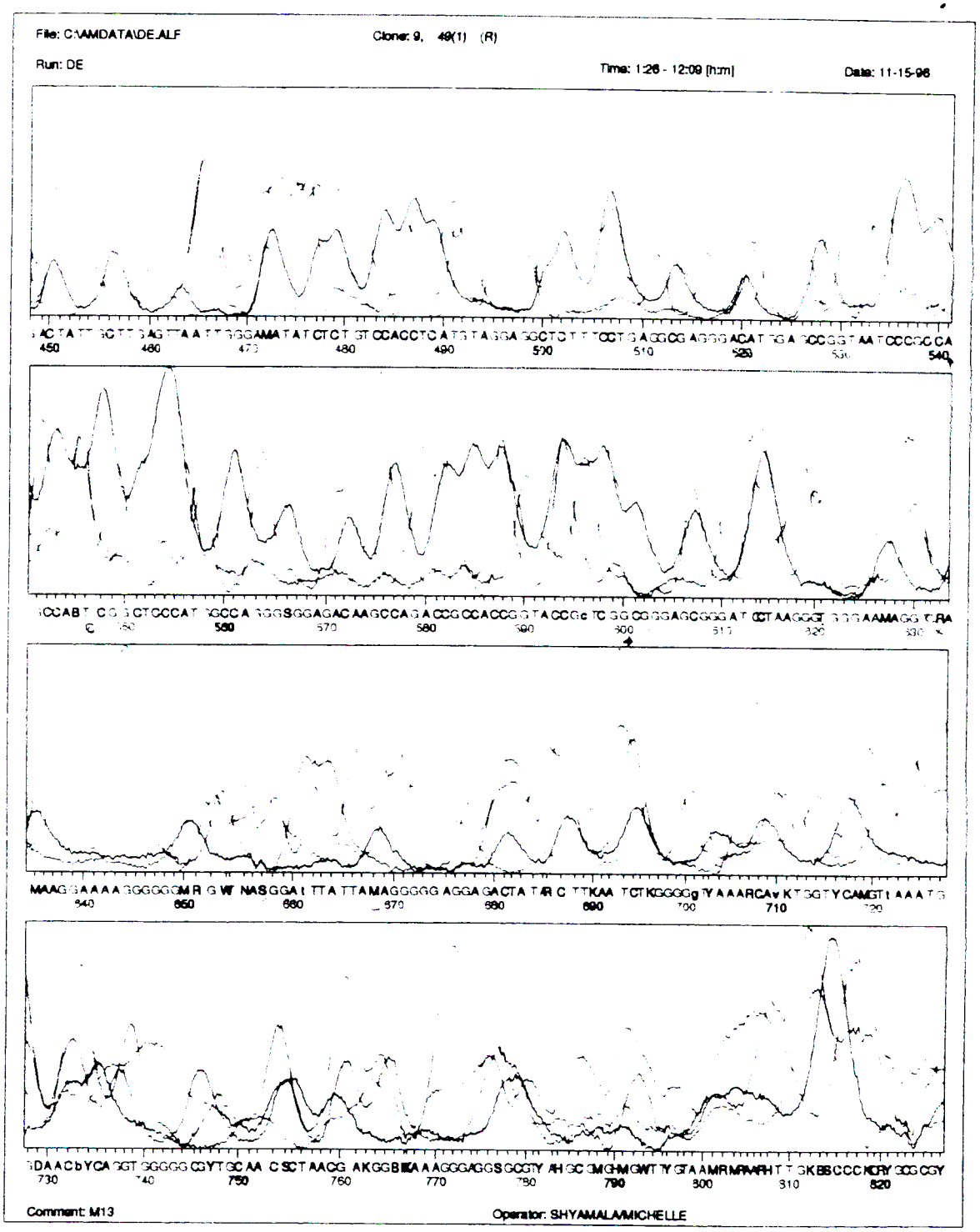

ABI PRISM
Model 310
Version 2.1.1
Signal G:55 A:151 T:137 C:81
Page 3 of 3
Lun 17 Fév 1997 2:21 PM
Sam 15 Fév 1997 4:08 PM
Spacing: 12.84 Semi/Adaptive
Lane 3

File: C:\AM\DATA\DE.ALF Clone: 9, 49(1) (R)

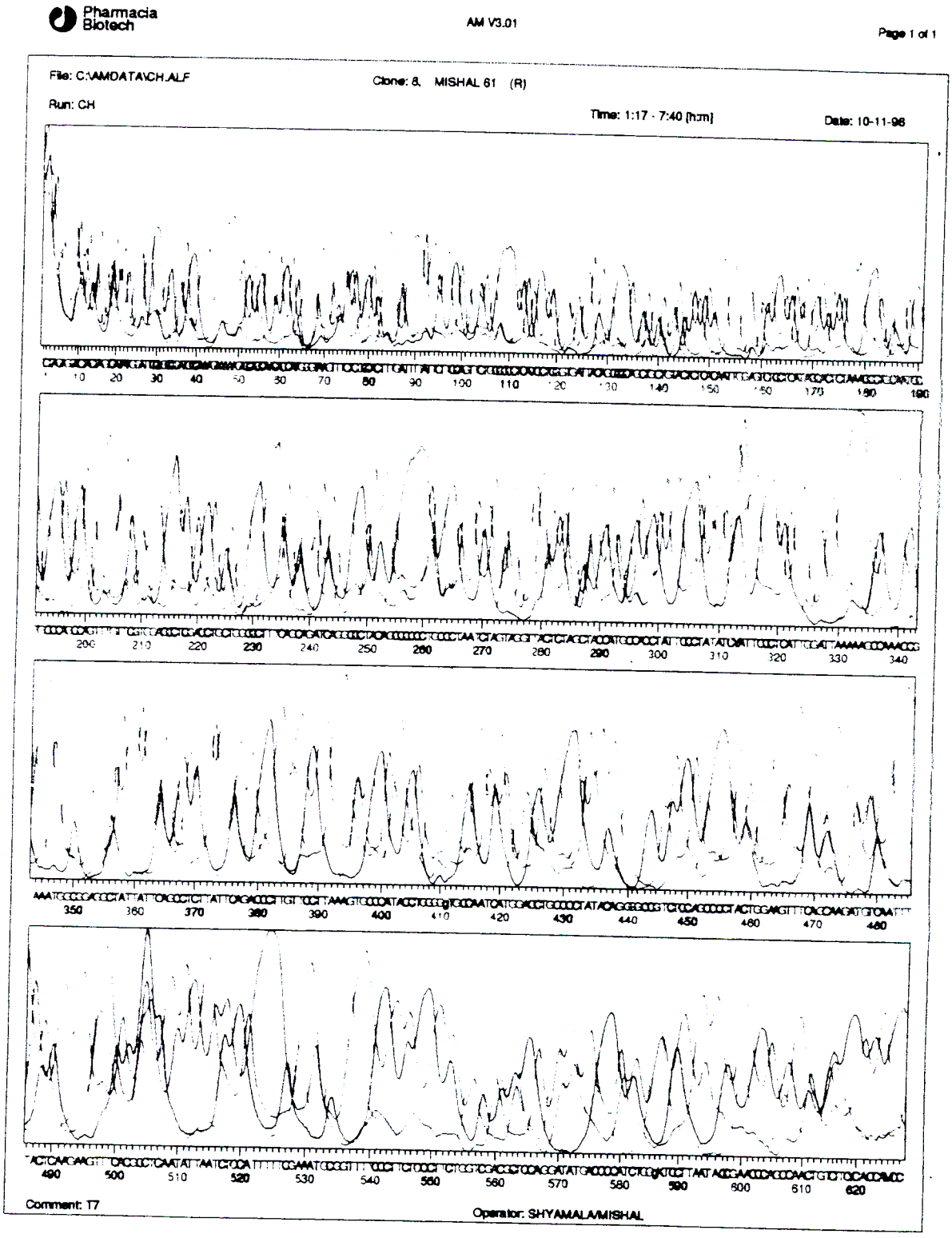
Run: DE Time: 1:26 - 12:09 [h:m] Date: 11-15-96

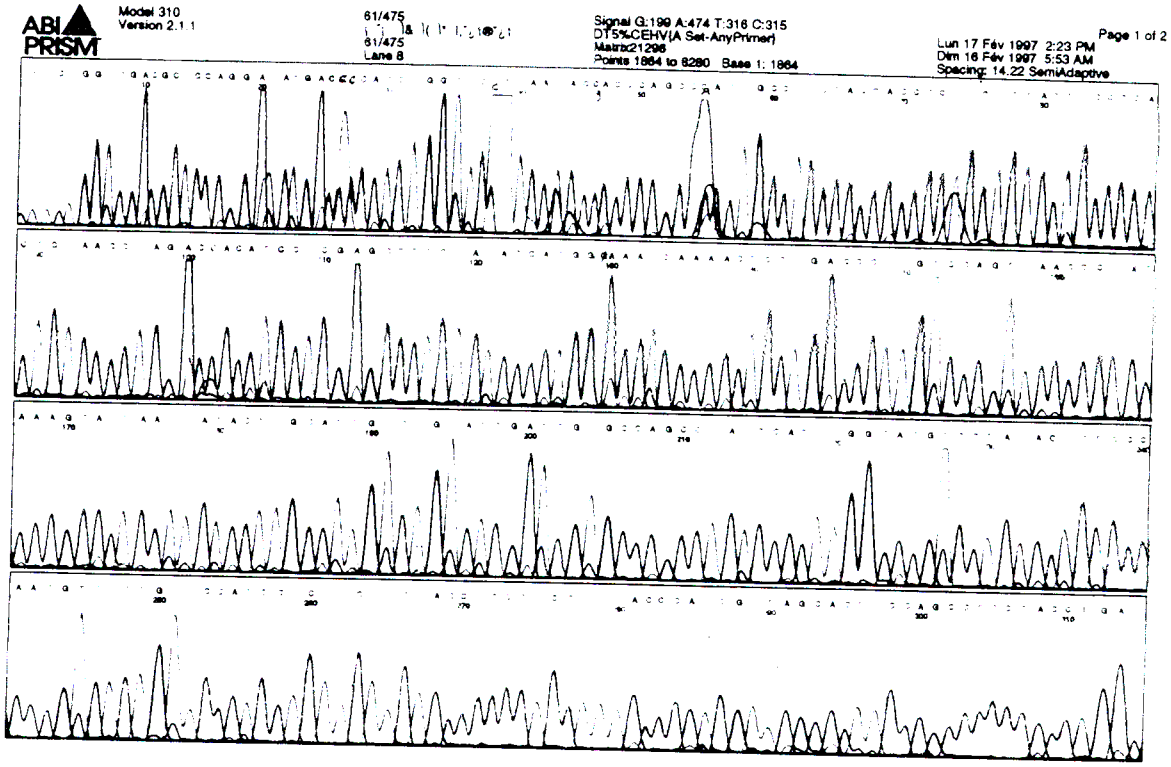


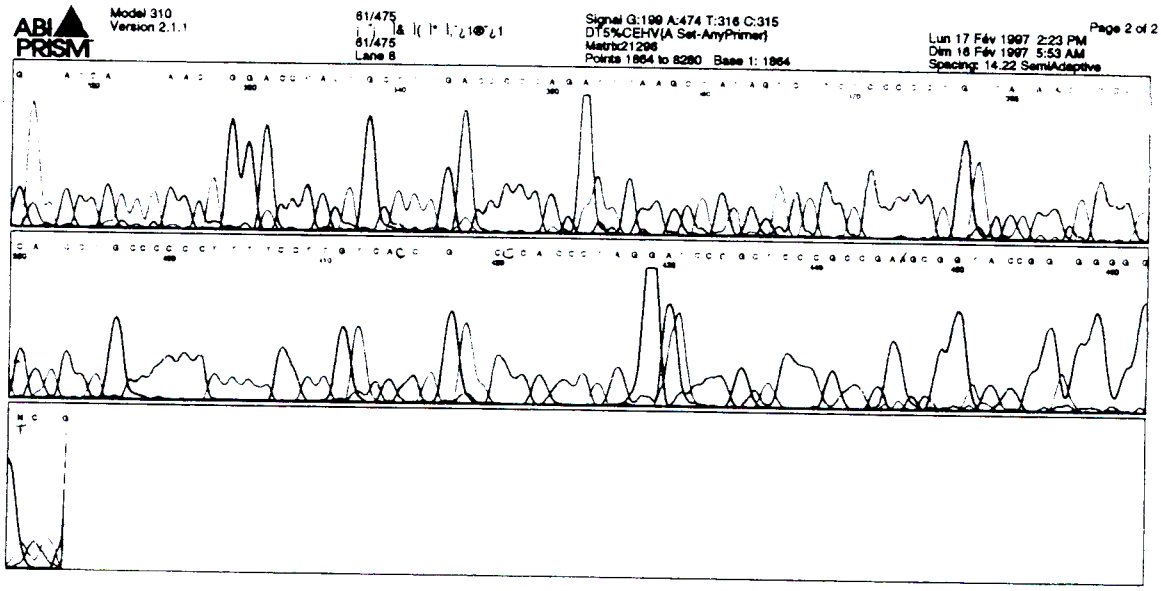
Comment: M13 Operator: SHYAMALA/MICHELLE

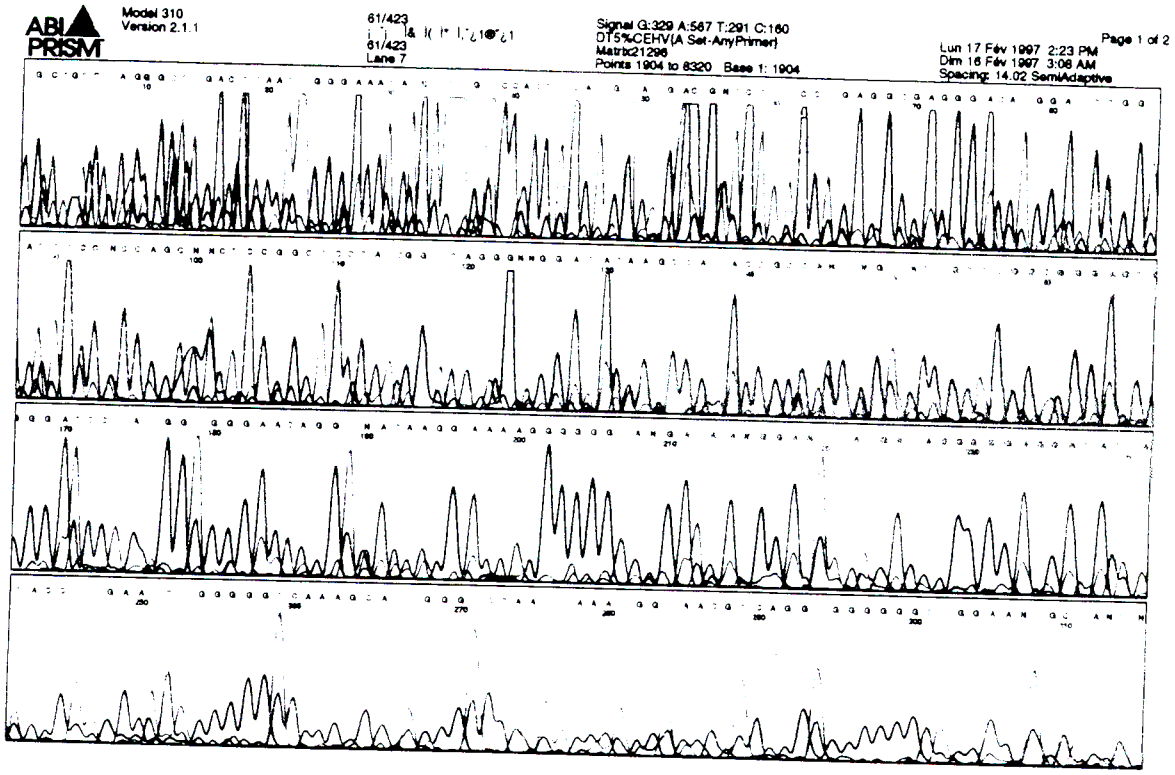

File: C:\AMDATA\DE.ALF
Clone: 9, 49(1) (R)
Run: DE
Time: 1:26 - 12:09 [h:m]
Date: 11-15-96
Comment: M13
Operator: SHYAMALA/MICHELLE

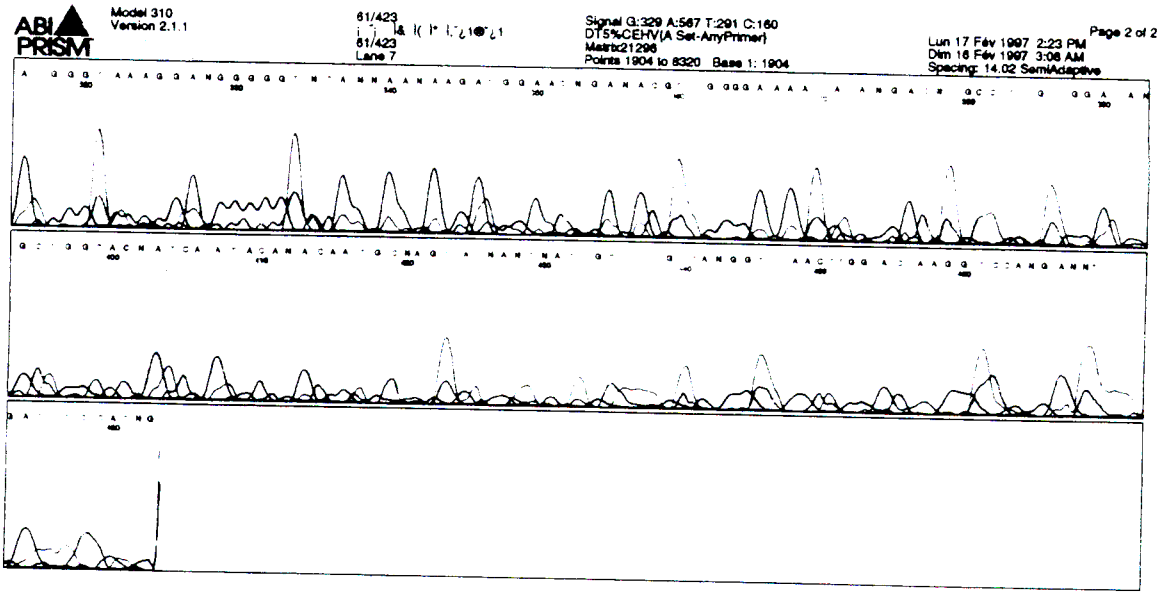
iii) 61

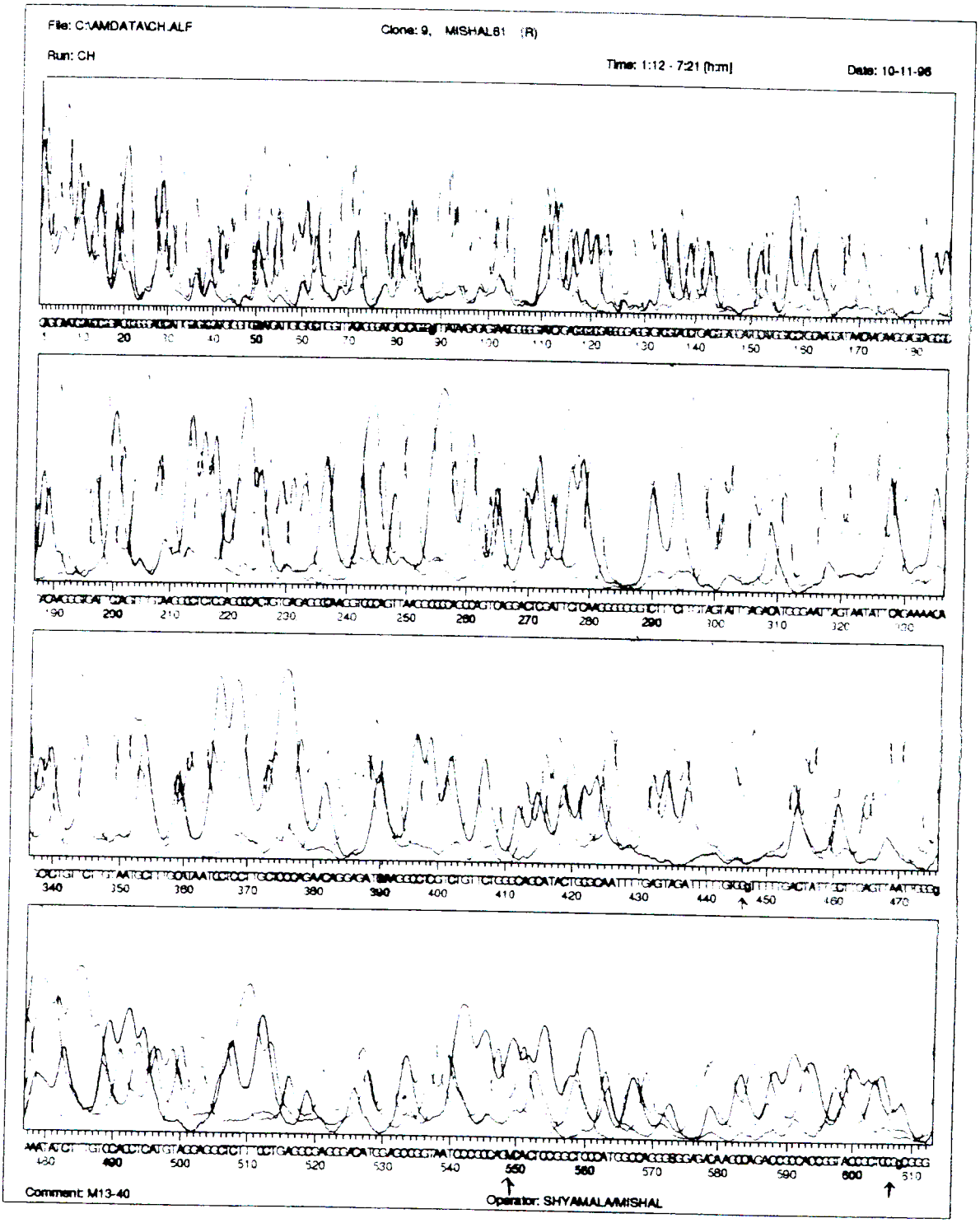

Pharmacia Biotech
AM V3.01
Page 1 of 1
File: C:\AMDATA\CH.ALF
Clone: 8, MISHAL 61 (R)
Run: CH
Time: 1:17 - 7:40 [h:m]
Date: 10-11-98
Comment: T7
Operator: SHYAMALA/MISHAL


ABI PRISM
Model 310
Version 2.1.1
61/475
61/475
Lane 8
Signal G:199 A:474 T:316 C:315
DT5%CEHV(A Set-AnyPrimer)
Matrbc21296
Points 1864 to 8280 Base 1: 1864
Page 1 of 2
Lun 17 Fév 1997 2:23 PM
Dim 16 Fév 1997 5:53 AM
Spacing: 14.22 SemiAdaptive

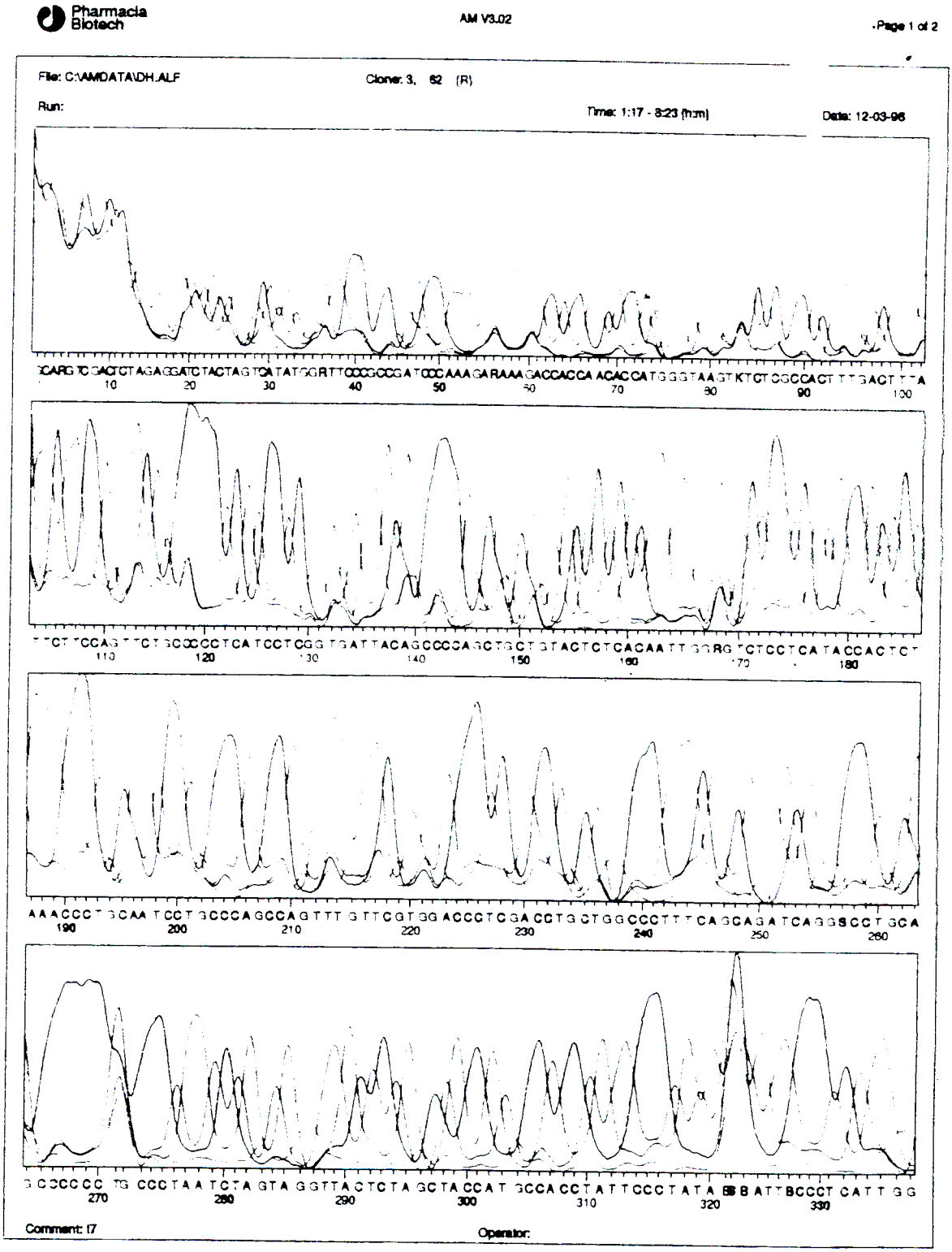

ABI PRISM
Model 310
Version 2.1.1
61/475
61/475
Lane 8
Signal G:199 A:474 T:316 C:315
DT5%CEHV(A Set-AnyPrimer)
Matrbc21296
Points 1864 to 8280 Base 1: 1864
Page 2 of 2
Lun 17 Fév 1997 2:23 PM
Dim 16 Fév 1997 5:53 AM
Spacing: 14.22 SemiAdaptive

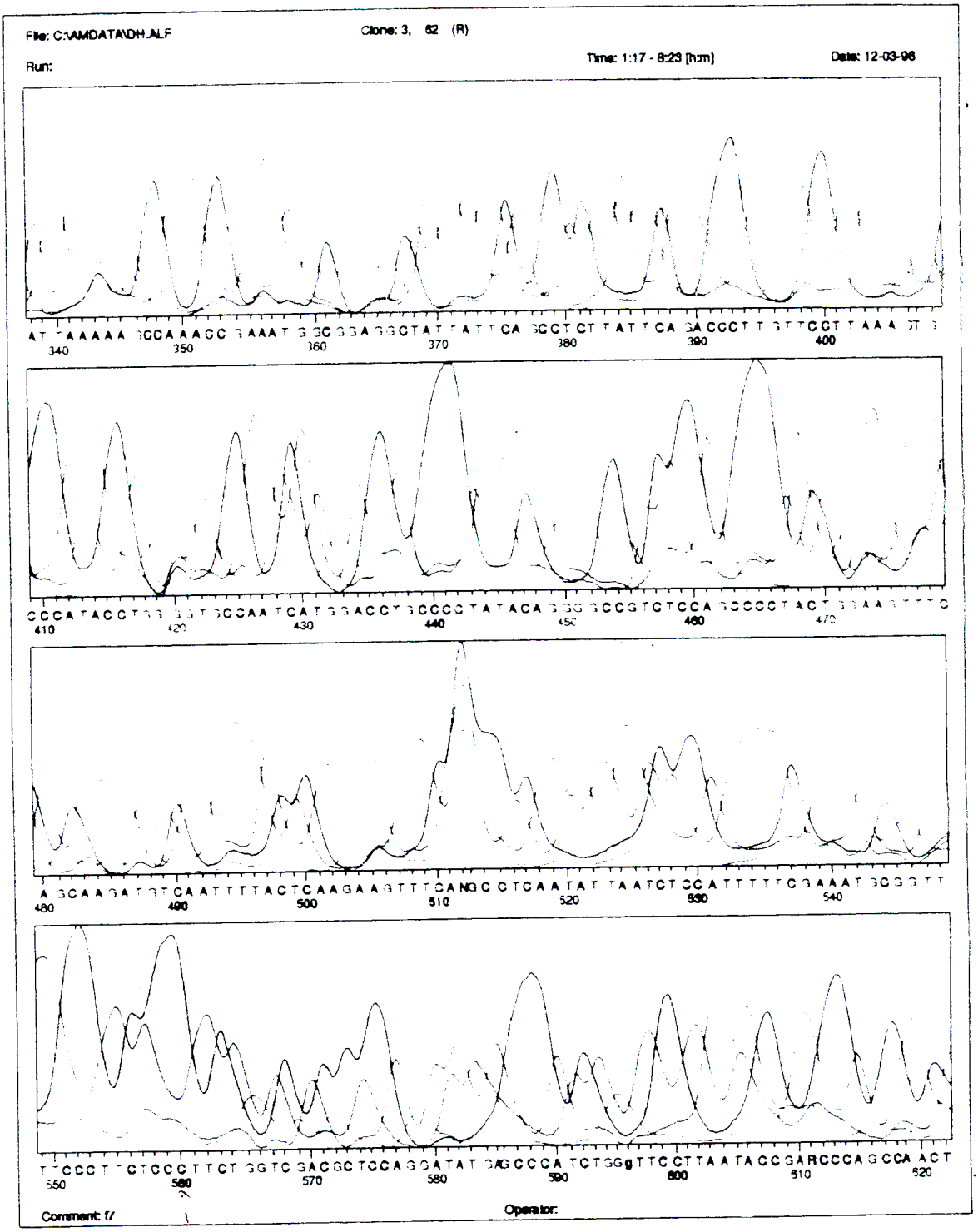

ABI PRISM
Model 310
Version 2.1.1
61/423
61/423
Lane 7
Signal G:329 A:567 T:291 C:160
DT5%CEHV(A Set-AmyPrimer)
Matrix21296
Points 1904 to 8320 Base 1: 1904
Page 1 of 2
Lun 17 Fév 1997 2:23 PM
Dim 16 Fév 1997 3:08 AM
Spacing: 14.02 SemiAdaptive

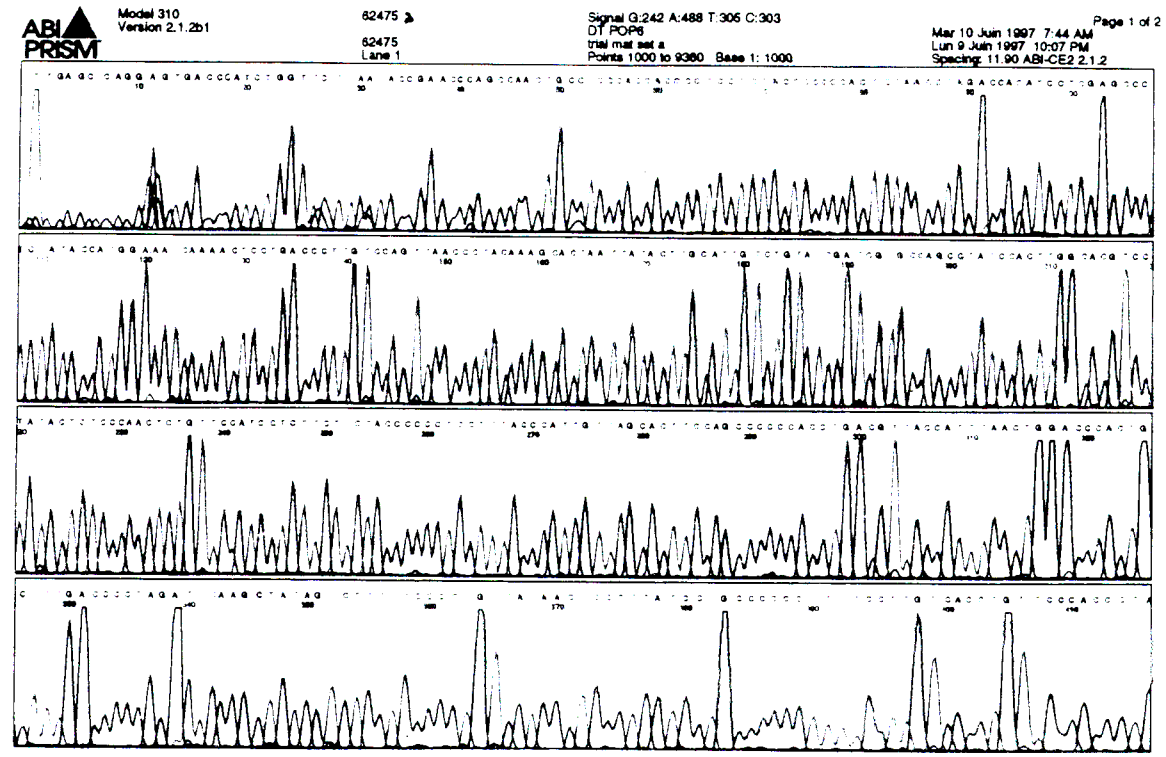

ABI PRISM
Model 310
Version 2.1.1
61/423
61/423
Lane 7
Signal G:329 A:567 T:291 C:160
DT5%CEHV(A Set-AmyPrimer)
Matrix21296
Points 1904 to 8320 Base 1: 1904
Page 2 of 2
Lun 17 Fév 1997 2:23 PM
Dim 16 Fév 1997 3:08 AM
Spacing: 14.02 SemiAdaptive

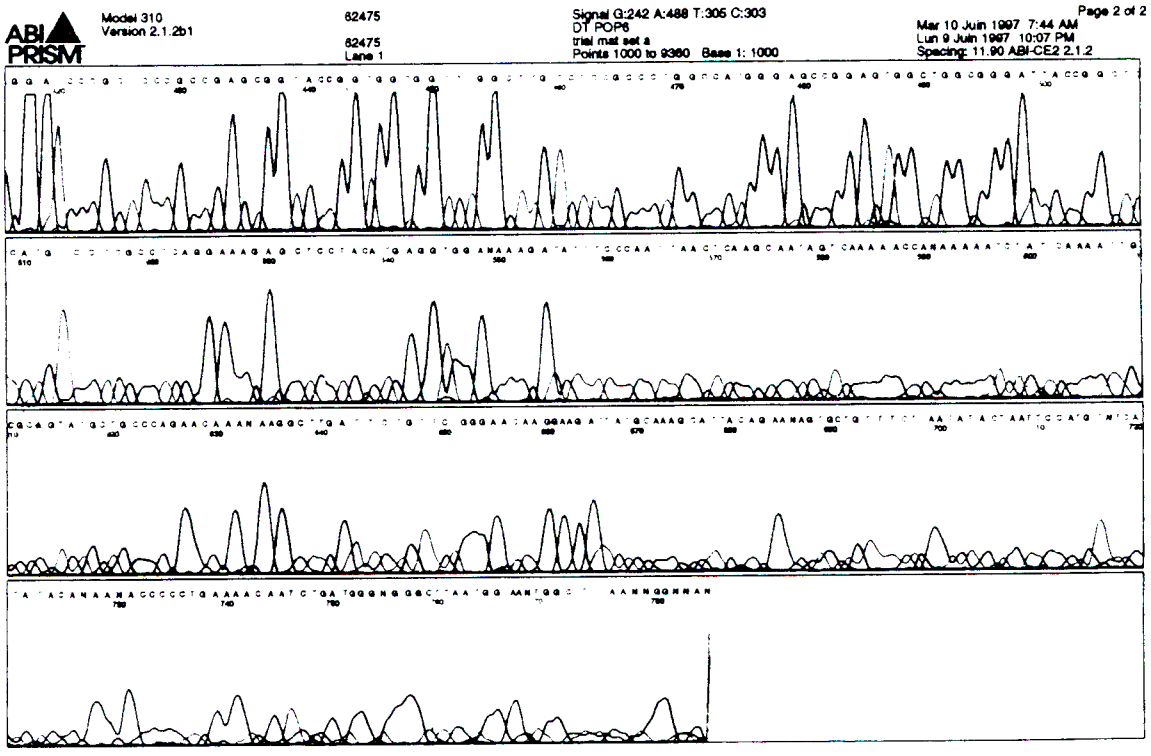

File: C:\AMDATA\CH.ALF
Clone: 9, MISHAL81 (R)
Run: CH
Time: 1:12 - 7:21 [h:m]
Date: 10-11-96
Comment: M13-40
Operator: SHYAMALA/MISHAL

iv) 62


Pharmacia Biotech
AM V3.02
Page 1 of 2
File: C:\AM\DATA\DH.ALF
Clone: 3, 62 (R)
Run:
Time: 1:17 - 8:23 (h:m)
Date: 12-03-96
Comment: I7
Operator:

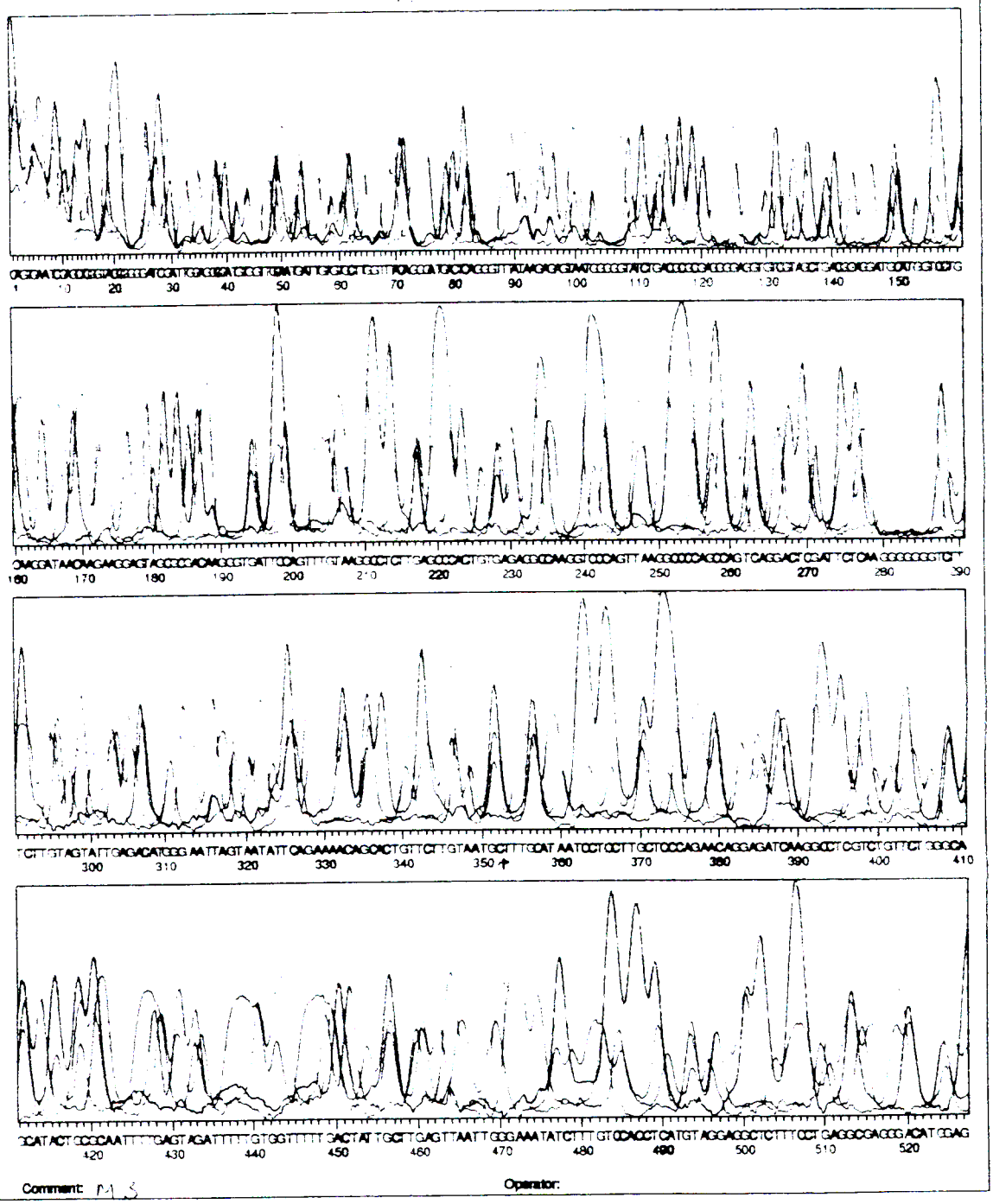

File: C:\AMDATA\DH.ALF
Clone: 3, 62 (R)
Run:
Time: 1:17 - 8:23 [h:m]
Date: 12-03-96
Comment:
Operator:


ABI PRISM
Model 310
Version 2.1.2b1
62475
62475
Lane 1
Signal G:242 A:488 T:305 C:303
DT POP6
trial mat set a
Points 1000 to 9360 Base 1: 1000
Page 1 of 2
Mar 10 Juin 1997 7:44 AM
Lun 9 Juin 1997 10:07 PM
Spacing: 11.90 ABI-CE2 2.1.2


ABI PRISM
Model 310
Version 2.1.2b1
62475
62475
Lane 1
Signal G:242 A:488 T:305 C:303
DT POP6
trial mat set a
Points 1000 to 9360 Base 1: 1000
Page 2 of 2
Mar 10 Juin 1997 7:44 AM
Lun 9 Juin 1997 10:07 PM
Spacing: 11.90 ABI-CE2 2.1.2

File: C:\AMDATA\LORNA3.ALF Clone: 2, (R)

Run: 62(4) Time: 1:25 - 14:27 [h:m] Date: 05-15-97

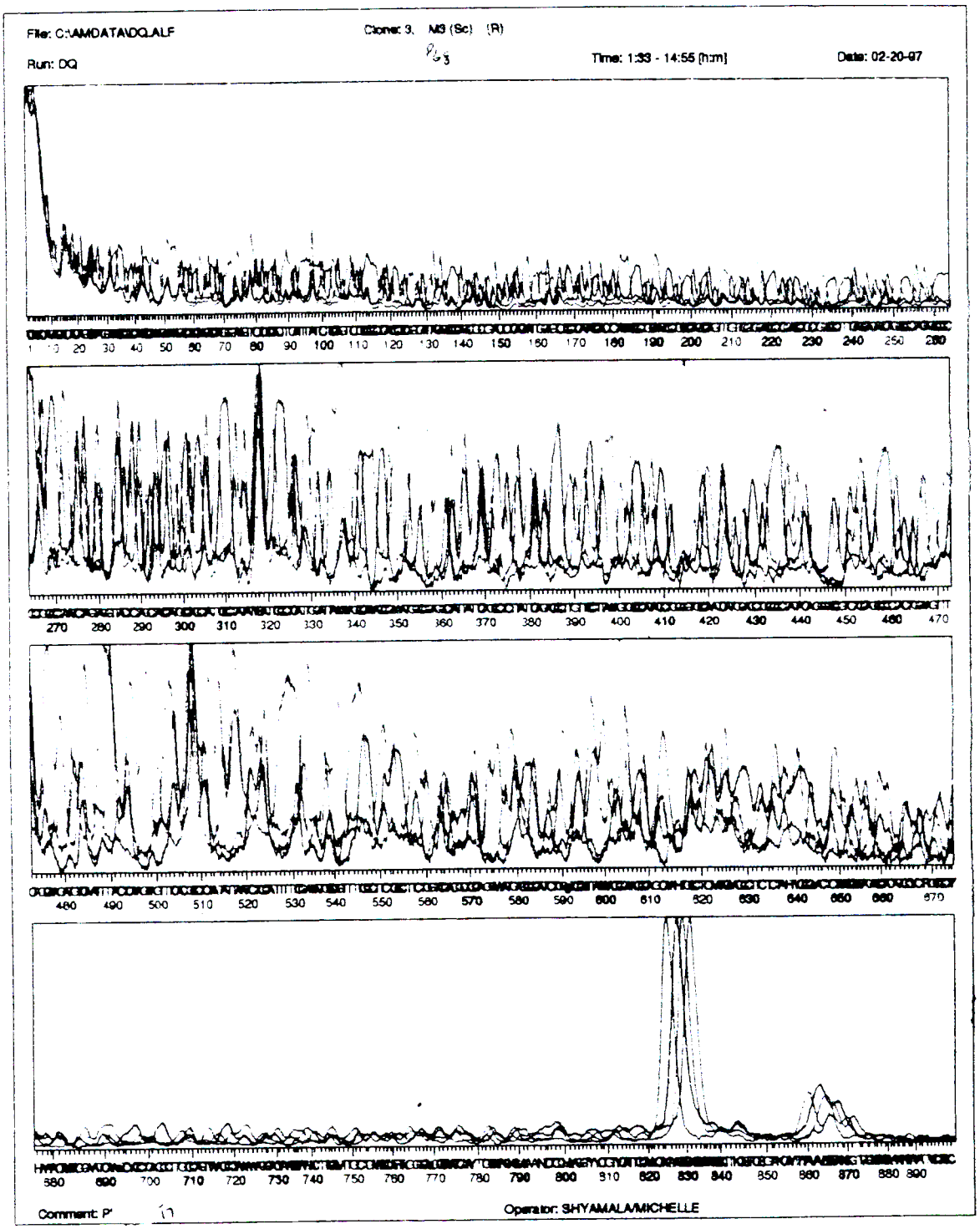


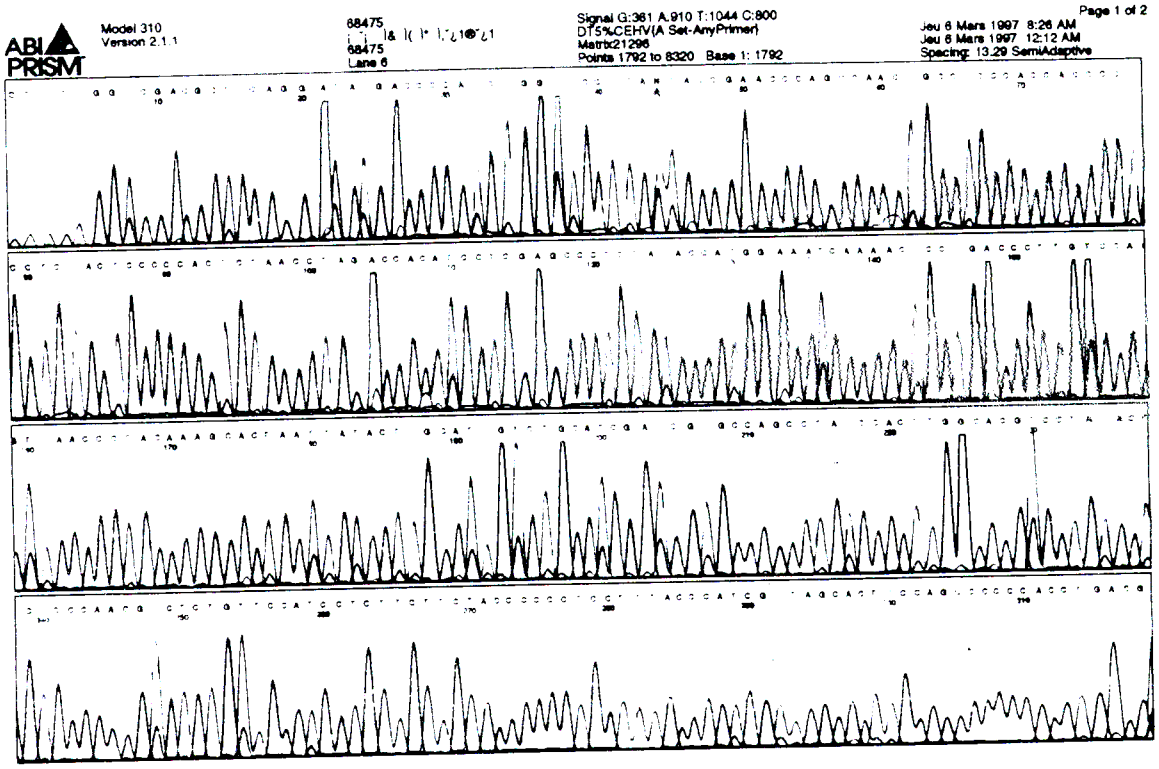
Comment: M13 Operator:

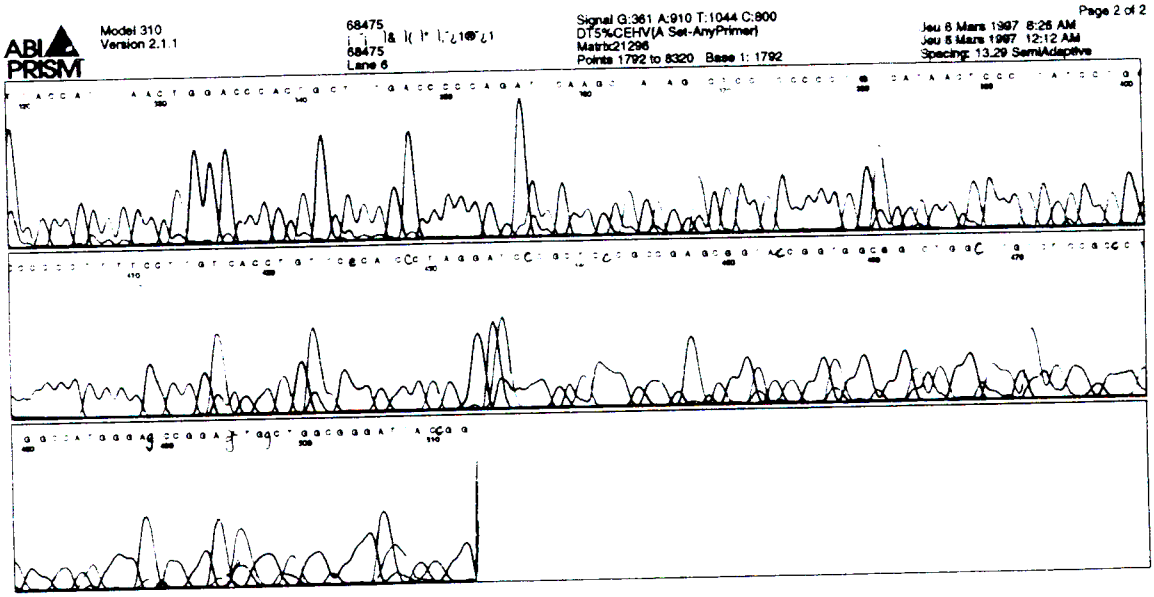
v)68

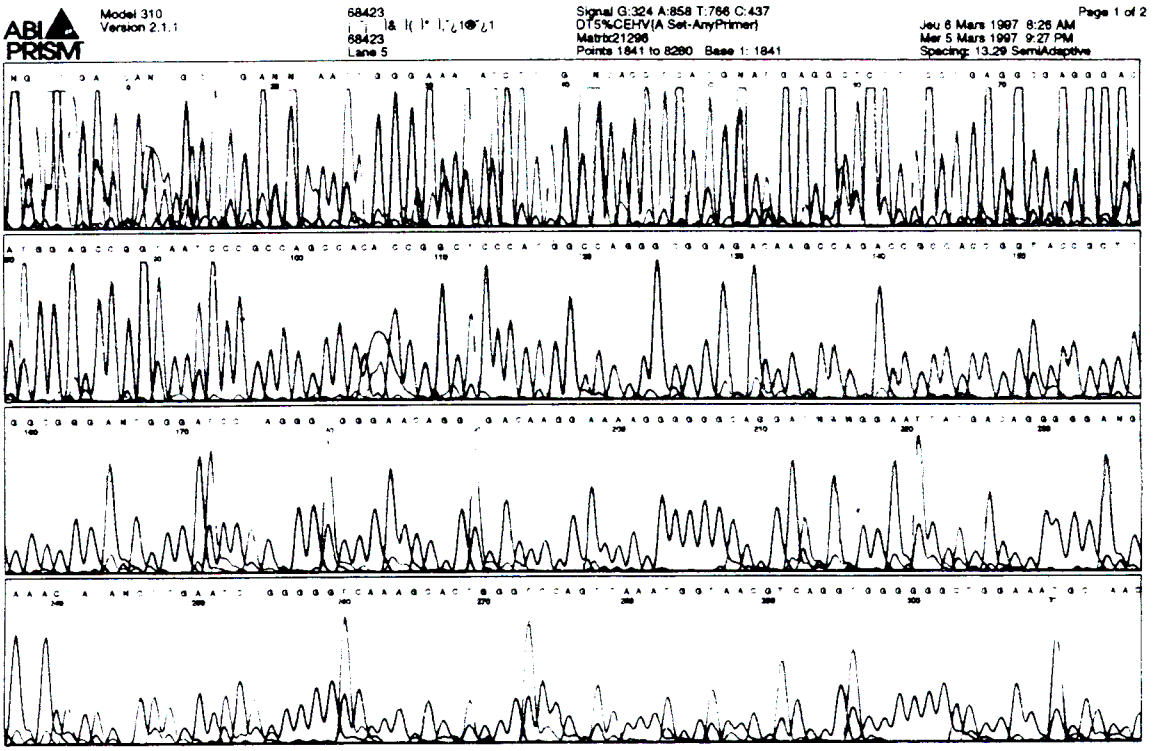
Pharmacia Biotech

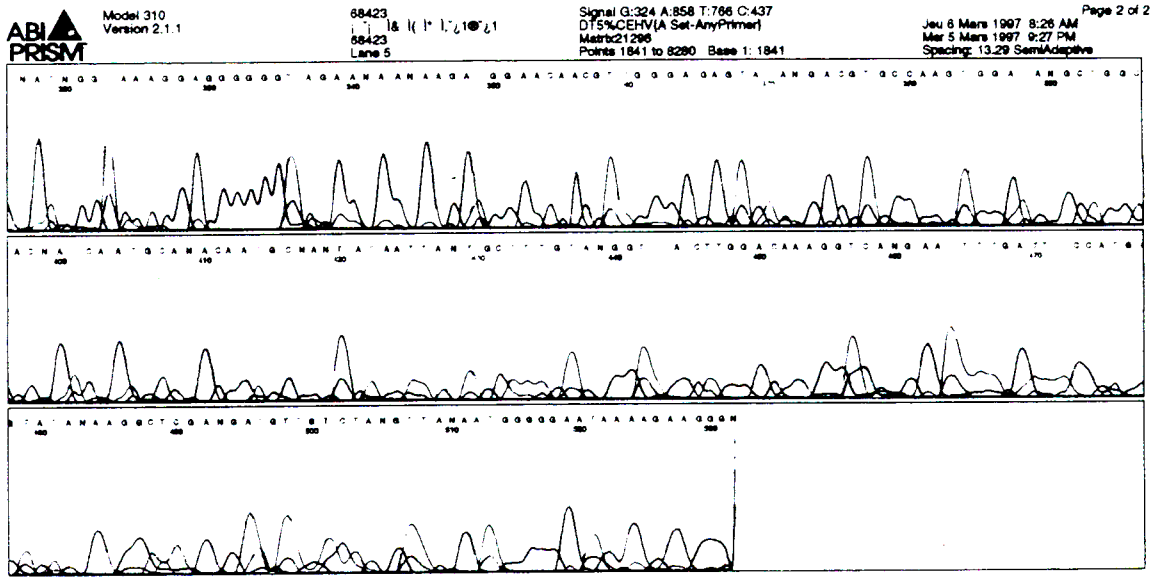
AM V3.02




ABI PRISM
Model 310
Version 2.1.1
68475
68475
Lane 6
Signal G:361 A:910 T:1044 C:800
DT5%CEHV(A Set-AmyPrimer)
Matrix:21296
Points 1792 to 8320 Base 1: 1792
Page 1 of 2
Jeu 6 Mars 1997 8:26 AM
Jeu 6 Mars 1997 12:12 AM
Spacing: 13.29 SemiAdaptive

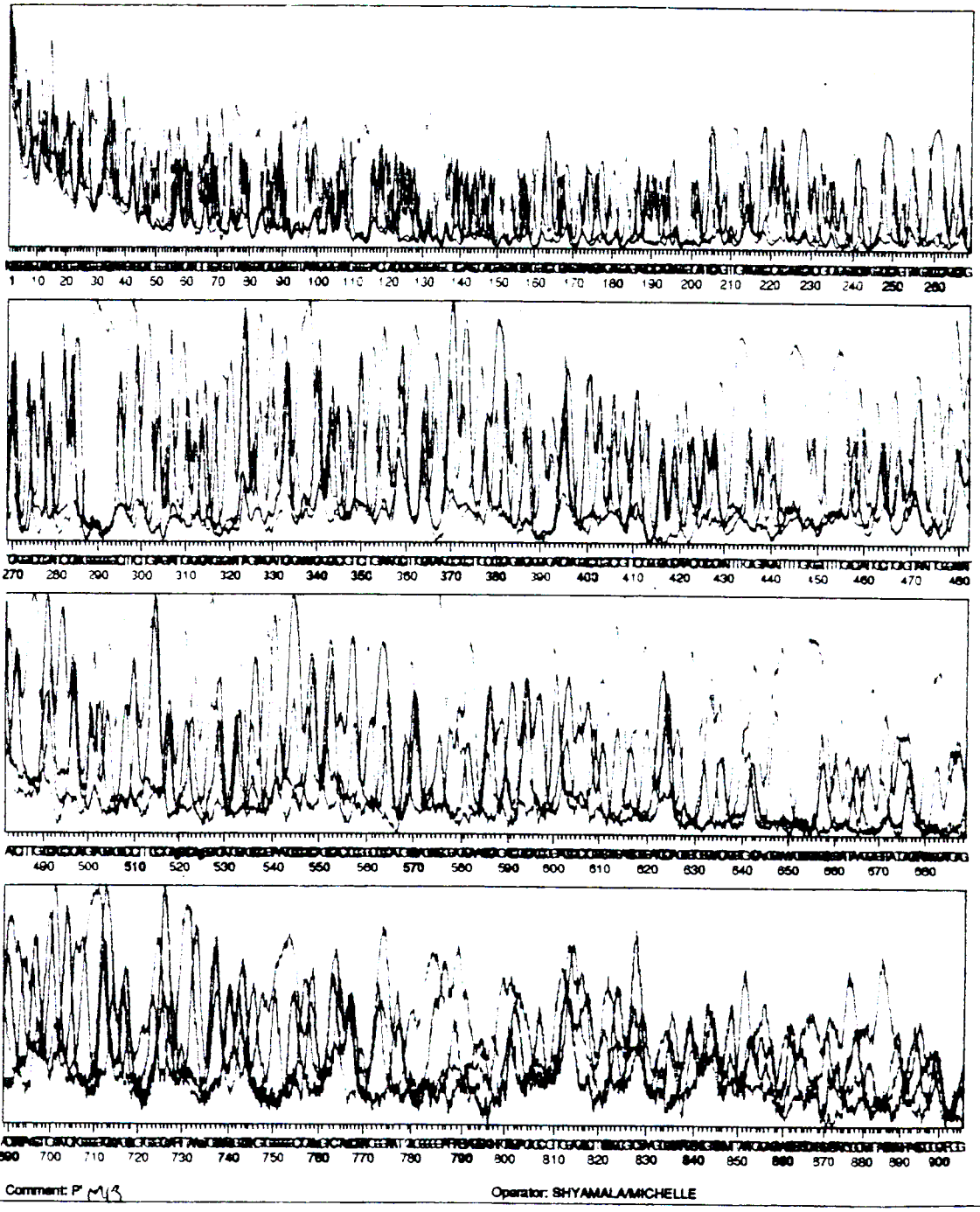

ABI PRISM
Model 310
Version 2.1.1
68475
68475
Lane 6
Signal G:361 A:910 T:1044 C:800
DT5%CEHV(A Set-AmyPrimer)
Matrix:21296
Points 1792 to 8320 Base 1: 1792
Page 2 of 2
Jeu 6 Mars 1997 8:26 AM
Jeu 6 Mars 1997 12:12 AM
Spacing: 13.29 SemiAdaptive


ABI PRISM
Model 310
Version 2.1.1
68423
68423
Lane 5
Signal G:324 A:858 T:766 C:437
DT5%CEHV(A Set-AnyPrimer)
Matrbc21296
Points 1841 to 8280 Base 1: 1841
Jeu 6 Mars 1997 8:26 AM
Mer 5 Mars 1997 9:27 PM
Spacing: 13.29 SemiAdaptive
Page 1 of 2


ABI PRISM
Model 310
Version 2.1.1
68423
68423
Lane 5
Signal G:324 A:858 T:766 C:437
DT5%CEHV(A Set-AnyPrimer)
Matrbc21296
Points 1841 to 8280 Base 1: 1841
Jeu 6 Mars 1997 8:26 AM
Mer 5 Mars 1997 9:27 PM
Spacing: 13.29 SemiAdaptive
Page 2 of 2

File: C:\AM\DATA\DQ.ALF Clone: 2. M2 (Sc) (R)

Run: DQ P29 Time: 1:24 - 14:57 [hrm] Date: 02-20-97



Comment: P M3 Operator: SHYAMALA/MICHELLE