Molecular Neurobiology

ISSN0893-7648/94/8(2–3): 155–173/$7.80

Genetic and Phylogenetic Analyses of Human T-Cell Lymphotropic Virus Type I Variants from Melanesians With and Without Spastic Myelopathy

Vivek R. Nerurkar, Ki-Joon Song, Rebecca R. Melland, and Richard Yanagihara*

Laboratory of Central Nervous System Studies, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Building 36, Rm. 5B-21, Bethesda, MD 20892

Abstract

Molecular variants of human T-cell lymphotropic virus type I (HTLV-I) have been isolated recently from lifelong residents of remote Melanesian populations, including a Solomon Islander with tropical spastic paraparesis/HTLV-I-associated myelopathy (TSP/HAM) or HTLV-I myeloneuropathy. To clarify the genetic heterogeneity and molecular epidemiology of disease-associated strains of HTLV-I, we enzymatically amplified, then directly sequenced representative regions of the *gag*, *pol*, *env*, and *pX* genes of HTLV-I strains from Melanesians with and without TSP/HAM, and aligned and compared these sequences with those of HTLV-I strains from patients with TSP/HAM or adult T-cell leukemia/lymphoma and from asymptomatic carriers from widely separated and culturally disparate populations. Overall, the HTLV-I variant from the Solomon Islander with TSP/HAM, like HTLV-I strains from asymptomatically infected Melanesians, diverged by approx 7% from cosmopolitan HTLV-I strain. No disease-specific viral sequences were found. Gene phylogenies, as determined by the unweighted pair-group method of assortment and by the maximum parsimony method, indicated that the Melanesian and cosmopolitan strains of HTLV-I have evolved along separate geographically dependent lineages, one comprised of HTLV-I strains from Papua New Guinea and the Solomon Islands, and the other composed of virus strains from Japan, India, the Caribbean, Polynesia, the Americas, and Africa. The total absence of nonhuman primates in Papua New Guinea and the Solomon Islands precludes any possibility that the Melanesian HTLV-I strains have evolved recently from the simian homolog of HTLV-I.

Introduction

Disease-specific long terminal repeat (LTR) and *env* gene sequences have not been found in strains of human T-cell lymphotropic virus type I (HTLV-I) isolated from patients with tropical spastic paraparesis/HTLV-I-associated myelopathy (TSP/HAM) or adult T-cell leukemia/lymphoma (ATLL), and virus strains from asymptomatically infected carriers are genetically indistinguishable from disease-associated strains *(1–7)*. Instead, previous studies on the genomic heterogeneity of HTLV-I

*Author to whom all correspondence and reprint requests should be addressed.

Table 1
Virus Strains and DNA Source for HTLV-I Gene Amplification and Sequence Analysis

Country of origin	Age, sex	Birth	Residence	Clinical status	Virus name	DNA source for PCR: Uncultured PBMC	Cultured PBMC	T-cell line
Papua New Gulnea	21M	Madang	Madang	AC	HTLV-I MEL 1	+		+
	60F	Madang	Madang	AC	HTLV-I MEL 2	+	+	
	31M	Madang	Madang	AC	HTLV-I MEL 7	+	+	
Solomon Islands	39F	New Georgia	Guadalcanal	AC	HTLV-I MEL 3		+	+
	60F	Guadalcanal	Guadalcanal	AC	HTLV-I MEL 4	+	+	
	58M	Guadalcanal	Guadalcanal	AC	HTLV-I MEL 5	+	+	+
	38M	Guadalcanal	Guadalcanal	TSP	HTLV-I MEL 6	+	+	
	49M	New Georgia	Guadalcanal	AC	HTLV-I MEL 8		+	
	75M	Rendova	Guadalcanal	AC	HTLV-I MEL 9		+	
	13F	Guadalcanal	Guadalcanal	AC	HTLV-I MEL 10		+	
	42F	Guadalcanal	Guadalcanal	AC	HTLV-I MEL 11		+	
	60F	Guadalcanal	Guadalcanal	AC	HTLV-I MEL 12	+	+	
	60F	Bellona	Bellona	AC	HTLV-I BEL 1	+	+	+
	50F	Bellona	Bellona	AC	HTLV-I BEL 2		+	

indicate a geographic dependence, as well as a high degree of sequence conservation *(4,8–11)*. In fact, the so-called cosmopolitan strains of HTLV-I from Japan, India, the Caribbean basin, the Americas, and Africa typically exhibit ≥98.5% sequence similarity among themselves. Even the variants of HTLV-I from the equatorial region of Zaire are 96.5–98% identical to other cosmopolitan strains *(4,5,9)*. By contrast, the recently identified Melanesian variants of HTLV-I *(12,13)* diverge by approx 7% from cosmopolitan strains of HTLV-I, including those from Zaire *(14–18)*. Thus, the Melanesian HTLV-I strains differ from cosmopolitan strains even more than do strains of simian T-cell lymphotropic virus type I (STLV-I) from chimpanzees and African monkeys *(19–23)*.

The identification of a case of TSP/HAM in a lifelong resident of the Solomon Islands *(24)* provided an opportunity to determine if disease-specific viral sequences were similarly nonexistent among the Melanesian HTLV-I strains. We amplified by polymerase chain reaction (PCR), then directly sequenced selected regions of the *gag*, *pol*, *env*, and *pX* genes of HTLV-I strains from a Solomon Islander with TSP/HAM and from asymptomatically infected Melanesians from Papua New Guinea and the Solomon Islands. The virus strain from the Solomon Islander with TSP/HAM was closely related to other Melanesian HTLV-I strains, indicating that these sequence variants of HTLV-I are capable of causing disease. Recently sequence variants of HTLV-I, which are genetically related to the Melanesian strains, have been isolated from Australian Aboriginals with adult T-cell leukemia *(25,26)*. These data further support the concept that HTLV-I genomic diversity is more a function of geographic origin than of disease.

Materials and Methods

Virus Strains

HTLV-I strains from 12 Melanesians from Papua New Guinea and the Solomon Islands (Table 1) and from two women (50- and 60-yr-old) from the Polynesian Outlier Bellona, all of whom had serological evidence for HTLV-I infection, as verified by Western blot analysis (Cambridge-Biotech, Worcester, MA), were studied (Figs. 1 and 2). All but one of these individuals were asymptomatic carriers, the exception being a 38-yr-old lifelong resident of Guadalcanal who had typical HTLV-I myeloneuropathy *(24)*. T-cell lines harboring HTLV-I strains were available from three Melanesians (strains MEL 1, MEL 3, and MEL 5) and one Polynesian (strain BEL 1) *(12,13)*. Informed consent was obtained from all study participants.

Virus strains from Papua New Guinea were from members belonging to a 260-member, recently contacted hunter-horticulturalist group (Hagahai) living along the northern banks of the Yuat River Gorge, at altitudes of 200–1800 m, in the westernmost reaches of the Schrader Range in Madang province (Fig. 1 and Table 1). The Hagahai share certain cultural practices with some lowland and highland New Guinean groups *(27)*, but analysis of their HLA-DR antigens indicates that they are dis-

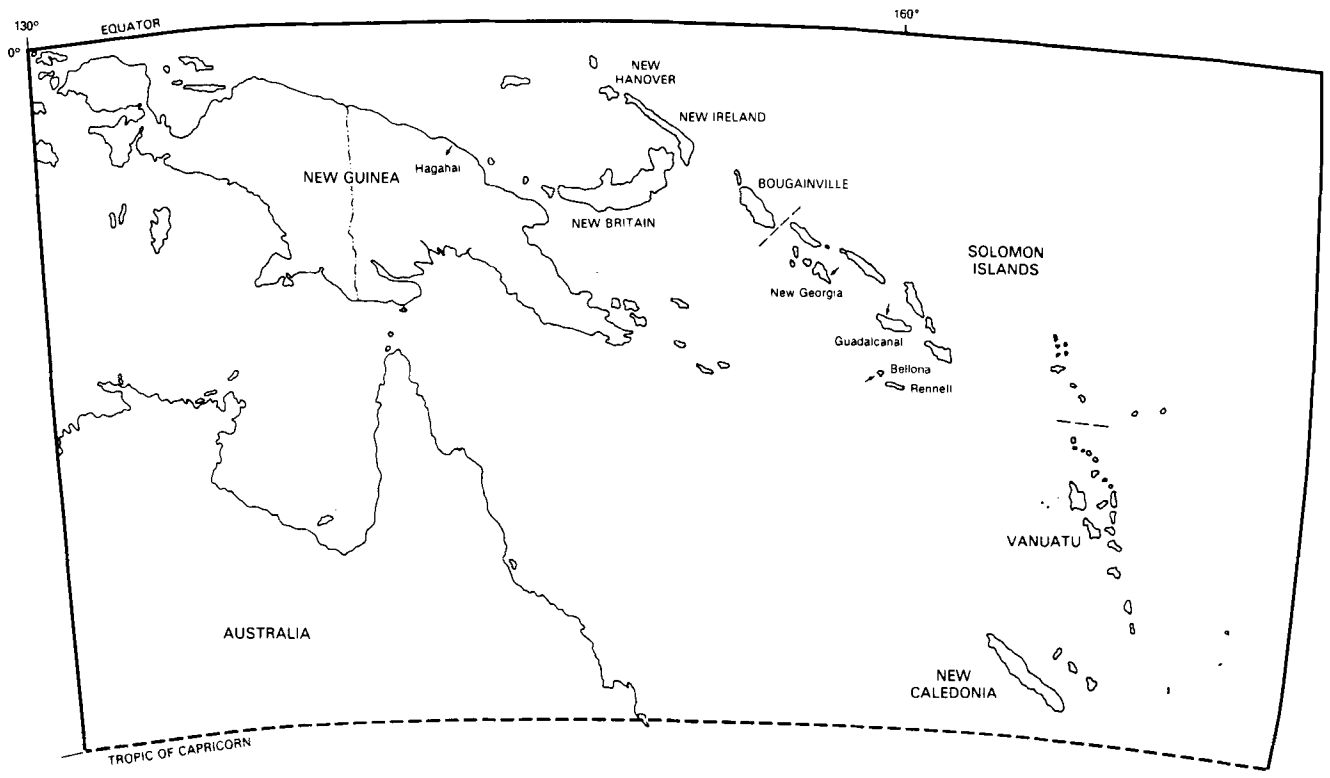


Fig. 1. Map of Melanesia, depicting the geographic relationship between New Guinea, Solomon Islands, Vanuatu, and New Caledonia. Arrows indicate where study participants resided. Blood specimens were collected after obtaining informed consent.

tinct from both coastal and highland groups and may represent descendants of an independent group of migrants, who arrived more than 5000 yr before present *(28,29)*.

Seven of the nine HTLV-I strains from the Solomon Islands were from three unrelated families residing in Honiara, Kolosulu, and Marau on Guadalcanal *(18)*. Members of two families were born elsewhere in the Solomon Islands (two on Morovo Lagoon and one on Rendova in the New Georgia group of islands, approx 250 km west of Guadalcanal). None of these nine individuals had traveled outside the Solomon Islands and six had never left Guadalcanal.

Gene Amplification

High-molecular-weight genomic DNA from a Solomon Islander with TSP/HAM and from 11 asymptomatically infected Melanesians from Papua New Guinea and the Solomon Islands was subjected to PCR. DNA was extracted using a non-organic method (Oncor, Gaithersburg, MD) from uncultured peripheral blood mononuclear cells (PBMC), from PBMC maintained in culture for 4 wk and, when available, from T-cell lines derived from infected Melanesians (Table 1). Unamplified DNA and enzymatically amplified products were handled by different personnel in separate rooms, and all specimens and reagents were dispensed into sterile tubes, using positive-displacement pipeting devices or pipeters fitted with aerosol-resistant tips.

Oligonucleotide primers for PCR and for direct DNA sequencing were derived from sequences of the Japanese HTLV strain ATK *(30)* for the B-cell immunodominant domain on the carboxy-terminal p19-encoding region of the *gag* gene (bases 1081–1100, 5'-ACTCATCCAAACCCAAGCCC-3' and bases 1278–1257, 5'-GCCTGTAGGTCTTTCATTTGC-3'); the amino-terminal p24-encoding region of the *gag* gene (bases 1423–1444, 5'-CCATCACCAG CAGCTAGATAGC-3' and bases 1560–1537, 5'-AGTTGCTGGTATTCTCGCCTTAAT-3'); the 3'-end of the *pol* gene (bases 4757–4778, 5'-CCCT ACAATCCAACCAGCTCAG-3' and bases 4942–4919, 5'-GTGGTGAAGCTGCCATCGGGTTTT-3'); the gp46-encoding region of the *env* gene (bases 5228–5246, 5'-TTTATTCTTCCAGTTCTGC-3' and bases 5596–5572, 5'-TAGGGGCTGGAGACGG CTCCTGTAT-3'); the gp21-encoding region of the

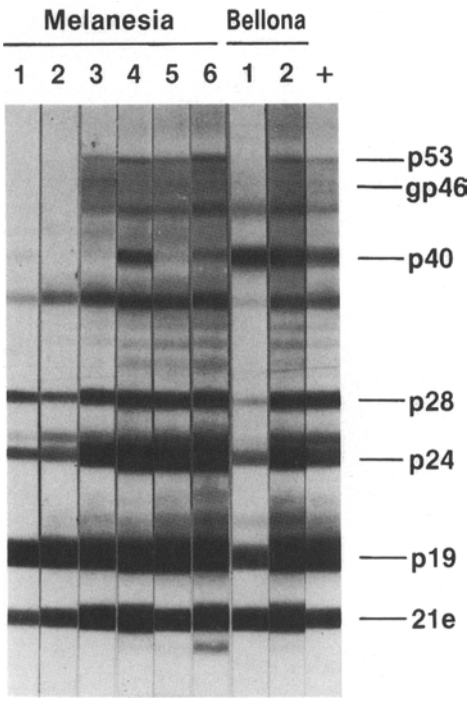


Fig. 2. Enhanced HTLV-I Western blot of sera, diluted 1:100, from six Melanesians (their corresponding virus strains were MEL 1–MEL 6) and two Polynesians (BEL 1 and BEL 2), showing reactivity to HTLV-I *gag*-encoded proteins p19 and p24, tax protein p40, and *env* gene products gp21 and/or gp46. Serum from a patient with adult T-cell leukemia/lymphoma served as the positive control (+). Viral protein-specific IgG antibodies were detected by the avidin-biotin technique. Viral proteins were prepared from the HTLV-I-infected cell line, HUT102. The purified, nonglycosylated recombinant transmembrane protein (*21e*) migrated at 16 kDa.

env gene (bases 6044–6067, 5'-TCAAGCTATA GTCTCCTCCCCCTG-3' and bases 6590–6613, 5'-GGGAGGTGTCGTAGCTGACGGAGG-3'); and the orf-II of the *pX* gene (bases 7358–7377, 5'-CGGATACCCAGTCTACGTGT-3' and bases 7516–7496, 5'-GAGCCGATAACGCGTCCATCG-3').

Primers were used at a final concentration of 1 μ*M* in a reaction mixture of 100 μL containing 2 μg of DNA and 2.5 U of *Thermus aquaticus* DNA polymerase (Perkin-Elmer/Cetus, Norwalk, CT) and comprised of 50 m*M* KCl, 10 m*M* Tris-HCl (pH 8.3), 1.5 m*M* $MgCl_2$, and 0.2 m*M* each dNTP. Using a DNA thermal cycler (Perkin-Elmer/Cetus), the mixture was initially denatured at 94°C for 5 min, then cycled 45 times at 94°C for 1 min, 55°C for 1.5 min, 72°C for 3 min, with a final extension of 7 min at 72°C before storing at 4°C *(31)*. Enzymatically amplified DNA was electrophoresed on 1.2% agarose gels to ascertain the presence of the appropriately sized fragment. As further verification, size-fractionated DNA was transferred onto nylon membranes and hybridized under high-stringency conditions using a ^{32}P-labeled full-length HTLV-I probe (Oncor), as described previously *(31)*.

Direct Sequencing of Amplified Products

Enzymatically amplified DNA samples, purified using Centricon 100 microconcentrators (Amicon Division, Danvers, MA), were sequenced directly by the dideoxy chain-termination method, using the same primers employed for PCR. Sequencing was performed in both directions using the *Taq* dye deoxy-terminator cycle sequencing kit (Applied Biosystems, Foster City, CA) on an automated sequencer (model 373A, Applied Biosystems) *(7,22)*. Sequence ambiguity was resolved by manual sequencing using the Sequenase version 2.0 DNA sequencing kit (US Biochemicals, Cleveland, OH). Direct sequencing of the PCR products was chosen because it obviated the need for cloning and provided rapid genetic information on the predominant virus genome within an infected individual *(32)*.

For manual sequencing, reactions were initiated by mixing 6 μL of heat-denatured DNA template, 2 μL of 5X concentrated Sequenase buffer (Sequenase version 2.0 DNA sequencing kit, US Biochemicals) and 2 μL (1.25–1.5 μCi) of γ-[^{32}P] ATP end-labeled sequencing primer in a 1.5-mL microcentrifuge tube, heating to 98°C for 5 min, and snap cooling on ice followed by brief centrifugation. To this annealed template-primer mixture was added 2.5 μL of dithiothreitol (DTT)-Mn^{2+} (1 μL 0.1*M* DTT, 1 μL 0.1*M* Mn^{2+}, 1 μL deionized distilled water) followed by mixing with 2 μL of Sequenase enzyme, diluted 1:8. Following centrifugation, the above template-primer mixture (3.5 μL) was dispensed into each of four tubes to which 2.5 μL aliquots of an appropriate dNTP DNA termination mix was added and incubated at 37°C for 3 min. Reactions were stopped on ice by adding 4 μL of stop solution. Samples were then heated at 80°C for 3 min and electrophoresed on 8% denaturing polyacrylamide gels in the presence of 1X TBE at 1400 V. Gels were thereafter dried and exposed to Kodak X-Omat film at –80°C for 1–3 d.

For automated sequencing, reaction premixes were prepared by mixing 4 μL of 5X TACS buffer, 1 μL of dNTP mix, 1 μL each of dye deoxyterminators

and 0.5 µL (4 U) of Ampli *Taq* DNA polymerase in 0.6-mL microcentrifuge tubes. In separate 0.6-mL tubes, 2.5 µL of autoclaved deionized distilled water, 9.5 µL of the above reaction premixes, 7 µL of heat-denatured DNA template, and 1 µL 3.2 p*M* primer were dispensed, then mixed, centrifuged, and overlaid with 40 µL of mineral oil. Samples were then cycled 25 times at 96°C for 30 s, 50°C for 15 s, and 60°C for 4 min before holding at 4°C. The sequencing products were purified using Centri-Sep™ columns (Princeton Separations, Adelphia, NJ). To the vacuum-dried eluate was added 5 µL of deionized formamide-EDTA mixture (deionized formamide 5 µL and 1 µL 50 m*M* EDTA, pH 8.0) heated at 90°C for 2 min, snap cooled on ice, and 2 µL samples were electrophoresed on 6% denaturing polyacrylamide gels, using an automated sequencer.

Sequence and Phylogenetic Analyses

Nucleotide and deduced amino acid sequences of the HTLV-I variants from the Solomon Islander with TSP/HAM and from asymptomatic Melanesian carriers were aligned and compared with those of HTLV-I strains from patients with TSP/HAM or ATLL and from asymptomatically infected individuals from various geographic regions, including strains ATK *(30)*, MT-2 *(33,34)*, H5 *(35)*, and TSP-1 *(36)* from Japan; strains BEL 1 and BEL 2 *(18)* from the Polynesian Outlier Bellona; strain 20 *(5)* from China; strains CMCH 1, CMCH 3, CMCH 4, CMCH 5, CMCH 7, and CMCH 13 *(7)* and SG *(37)* from India; strain H990 *(10)* from Romania; strains HS-35 *(38)*, 1010/3 *(38)*, CH *(4,9)*, and RD-1 (39) from the Caribbean basin; strains 1 and 4 *(5)* from Martinique; strains 5 and 7 *(5)* from Guadeloupe; strain 11 *(5)* from Haiti; strain 8 *(5)* from French Guyana; strains pt3, pt5, and pt8 *(10)* from Brazil; strain ST *(40)* from Chile; strain 10 *(5)* from Peru; strains CR *(22)*, SP *(4)*, and AJ5740 *(41)* from the US; strains EL *(4,9)*, Z23, and Z69 *(5)* from equatorial Zaire; strain CTCL-11B *(16)* from Liberia; strains 15 and 19 *(5)* from Ivory Coast; strain 1711 *(5)* from Guinea Bissau; strain 14 *(5)* from Mauritania; and strain 13 *(5)* from Central African Republic.

Sequence analysis was facilitated by using programs available on the VAX computer system, as part of the Genetics Computer Group *(42)*. Nucleotide sequences were aligned with the sequence editor PRETTY (Ann Palmenberg of UW Biophysics) and amino acid sequences were aligned with PUBLISH. Phylogenetic trees were constructed using the unweighted pair-group method of assortment *(43)*, which assumes a constant rate of evolutionary change, and the maximum parsimony method *(44)*, which accommodates variable rates of change. The branch-and-bound search of PAUP version 3.1 *(44)* was used to ensure finding the true number of parsimonious trees.

Results

Nucleotide Sequence Analysis

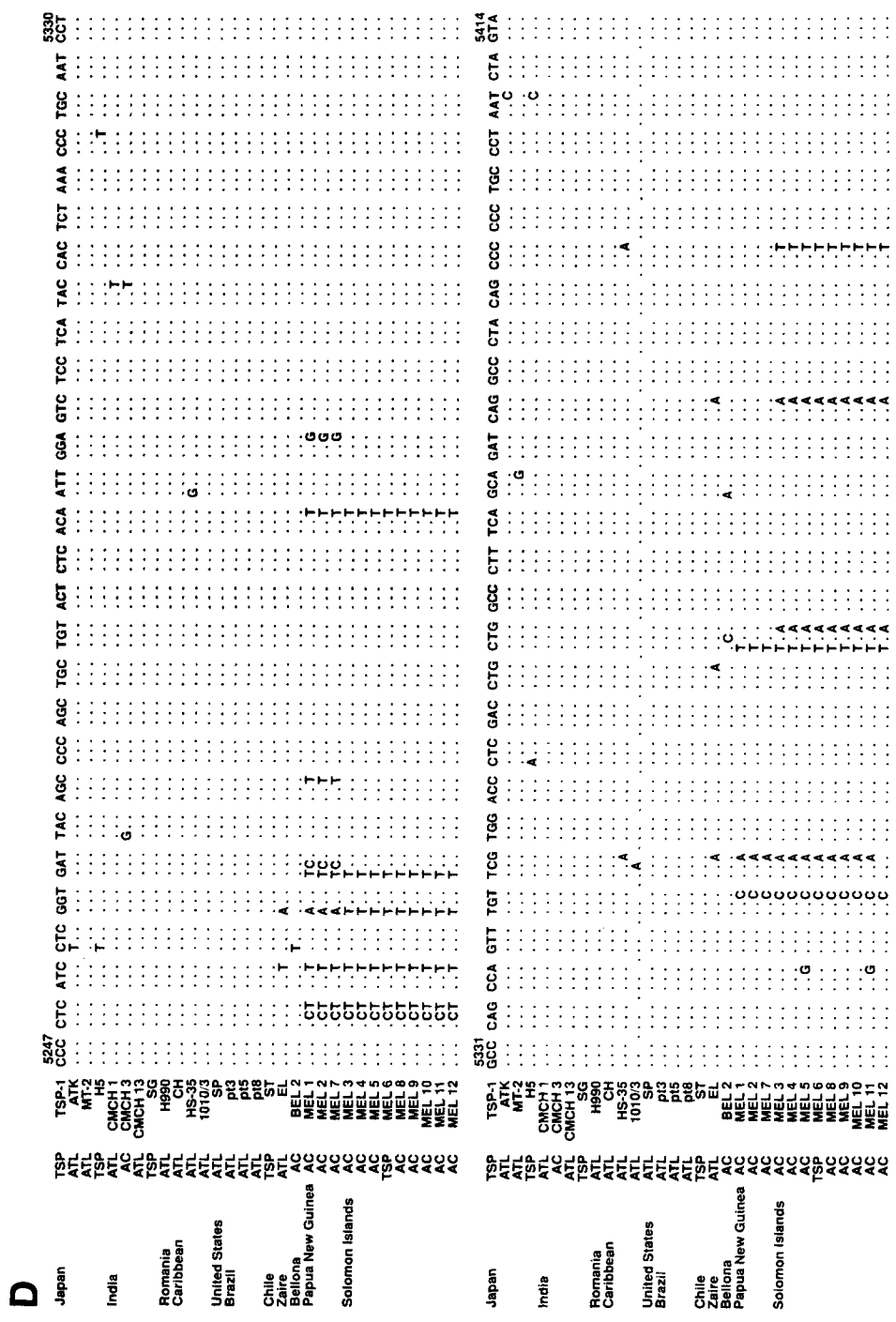
For an overall view of the genomic heterogeneity of the HTLV-I strains from Melanesians with and without spastic myelopathy, we sequenced selected regions spanning nearly the entire genome, including the p19- and p24-encoding *gag*, *pol*, gp46-, and gp21-encoding *env*, and *pX* orf-II. When aligned and compared to representative strains of HTLV-I from other geographic regions, the nucleotide sequences of the HTLV-I strains from the Solomon Islander with spastic paraparesis and from the 11 asymptomatically infected Melanesian carriers exhibited multiple base substitutions in the *gag*, *pol*, and *env* genes (Fig. 3A–F). Most of these nucleotide changes were not found among cosmopolitan HTLV-I strains and were thus unique to the Melanesian strains. Some of the base changes differentiated between virus strains from Papua New Guinea and those from the Solomon Island (e.g., position 1156, 1198, 1453, 1459, and 1501 in the *gag* gene; positions 4786, 4804, 4834, 4844, 4858, and 4903 in the *pol* gene Fig. 3A–C), and a few of the mutations were found in only one virus strain (e.g., position 6342 in the *env* gene of strain MEL 5) (Fig. 3D). The tendency for some of the base substitutions to occur at identical positions in the genes of cosmopolitan and Melanesian HTLV-I strains suggested a nonrandom nature to the mutation. In all three gene regions, nucleotide changes consisted primarily of transitions, with the majority being deoxycytosine to deoxythymidine substitutions or vice versa (Table 2). No insertions, deletions, or frame shifts were detected in the regions sequenced.

Identical proviral sequences were obtained from DNA extracted from fresh, uncultured PBMC, from PBMC maintained in culture for 4 wk, and from T-cell lines, indicating that the sequences of the Melanesian HTLV-I strains, presented here, represent those of the predominant virus population in vivo and do not reflect selection of aberrant virus populations during short- or long-term maintenance of cells in culture.

Compared to the HTLV-I strain TSP-1 (isolated from a Japanese patient with TSP/HAM), the nucle-

A
1121
1220
TSP CCA CCG CCG CCG TCA TCC CCC ACC CAC GAC CCC CCG GAT TCT GAT CCA CAA ATC CCC CCT CCC TAT GTT GAC CGT ACG GCC CCC CAA GTC CTT CCA GTC A
Japan TSP-1 . . . . . .
ATL ATK . . . . . .
ATL MT-2 . . . . . .
India ATL CMCH 1 . . . . . .
AC CMCH 4 . . . . . .
Caribbean ATL CMCH 13 . . . . . .
ATL CH . . . . . .
ATL HS-35 . . . . . .
TSP RD 1 . . . . . .
Zaire ATL EL . . . . . .
Bellona AC BEL 2 . . . . . .
Papua New Guinea AC MEL 1 . . . . . .
AC MEL 2 . . . . . .
AC MEL 7 . . . . . .
Solomon Islands AC MEL 3 . . . . . .
AC MEL 4 . . . . . .
AC MEL 5 . . . . . .
TSP MEL 6 . . . . . .
AC MEL 8 . . . . . .
AC MEL 9 . . . . . .
AC MEL 10 . . . . . .
AC MEL 11 . . . . . .
AC MEL 12 . . . . . .
B
1445
1536
TSP CTT ATA TCA GAG GCC GAA ACC CGA GGT ATT ACA GGT TAT AAC CCC TTA GCC GGT CCC CTC CGT GTC CAA GCC AAC AAT CCA CAA CAA CAA GG
Japan TSP-1 . . . . . .
ATL ATK . . . . . .
ATL MT-2 . . . . . .
India ATL CMCH 1 . . . . . .
AC CMCH 5 . . . . . .
TSP CMCH 7 . . . . . .
ATL CMCH 13 . . . . . .
Caribbean ATL CH . . . . . .
ATL HS-35 . . . . . .
TSP RD 1 . . . . . .
United States ATL CR . . . . . .
Chile TSP ST . . . . . .
Zaire ATL EL . . . . . .
Bellona AC BEL 2 . . . . . .
Papua New Guinea AC MEL 1 . . . . . .
AC MEL 2 . . . . . .
AC MEL 7 . . . . . .
Solomon Islands AC MEL 3 . . . . . .
AC MEL 4 . . . . . .
AC MEL 5 . . . . . .
TSP MEL 6 . . . . . .
AC MEL 8 . . . . . .
AC MEL 9 . . . . . .
AC MEL 10 . . . . . .
AC MEL 11 . . . . . .
AC MEL 12 . . . . . .

Fig. 3. Alignment and comparison of nucleotide sequence of **(A)** the 100-bp B-cell immunodominant epitope-spanning p19-encoding region of the *gag* gene (bases 1121–1220), **(B)** the 92-bp p24-encoding region of the *gag* gene (bases 1445–1536), **(C)** the 140-bp 3'-end of the *pol* gene (bases 4779–4918) *(continued)*.


D
5247
5330
5331
5414
TSP-1 CCC CTC ATC CTC GGT GAT TAC AGC CCC AGC TGC TGT ACT CTC ACA ATT GGA GTC TCC TCA TAC CAC TCT AAA CCC TGC AAT CCT
ATK
MT-2
H5
CMCH 1
CMCH 3
CMCH 13
SG
H990
CH
HS-35
1010/3
SP
pt3
pt5
pt8
ST
EL
BEL 2
MEL 1
MEL 2
MEL 7
MEL 3
MEL 4
MEL 5
MEL 6
MEL 8
MEL 9
MEL 10
MEL 11
MEL 12
TSP-1 GCC CAG CCA GTT TGT TCG TGG ACC CTC GAC CTG CTG GCC CTT TCA GCA GAT CAG GCC CTA CAG CCC CCC TGC CCT AAT CTA GTA
ATK
MT-2
H5
CMCH 1
CMCH 3
CMCH 13
SG
H990
CH
HS-35
1010/3
SP
pt3
pt5
pt8
ST
EL
BEL 2
MEL 1
MEL 2
MEL 7
MEL 3
MEL 4
MEL 5
MEL 6
MEL 8
MEL 9
MEL 10
MEL 11
MEL 12
Japan
India
Romania
Caribbean
United States
Brazil
Chile
Zaire
Bellona
Papua New Guinea
Solomon Islands

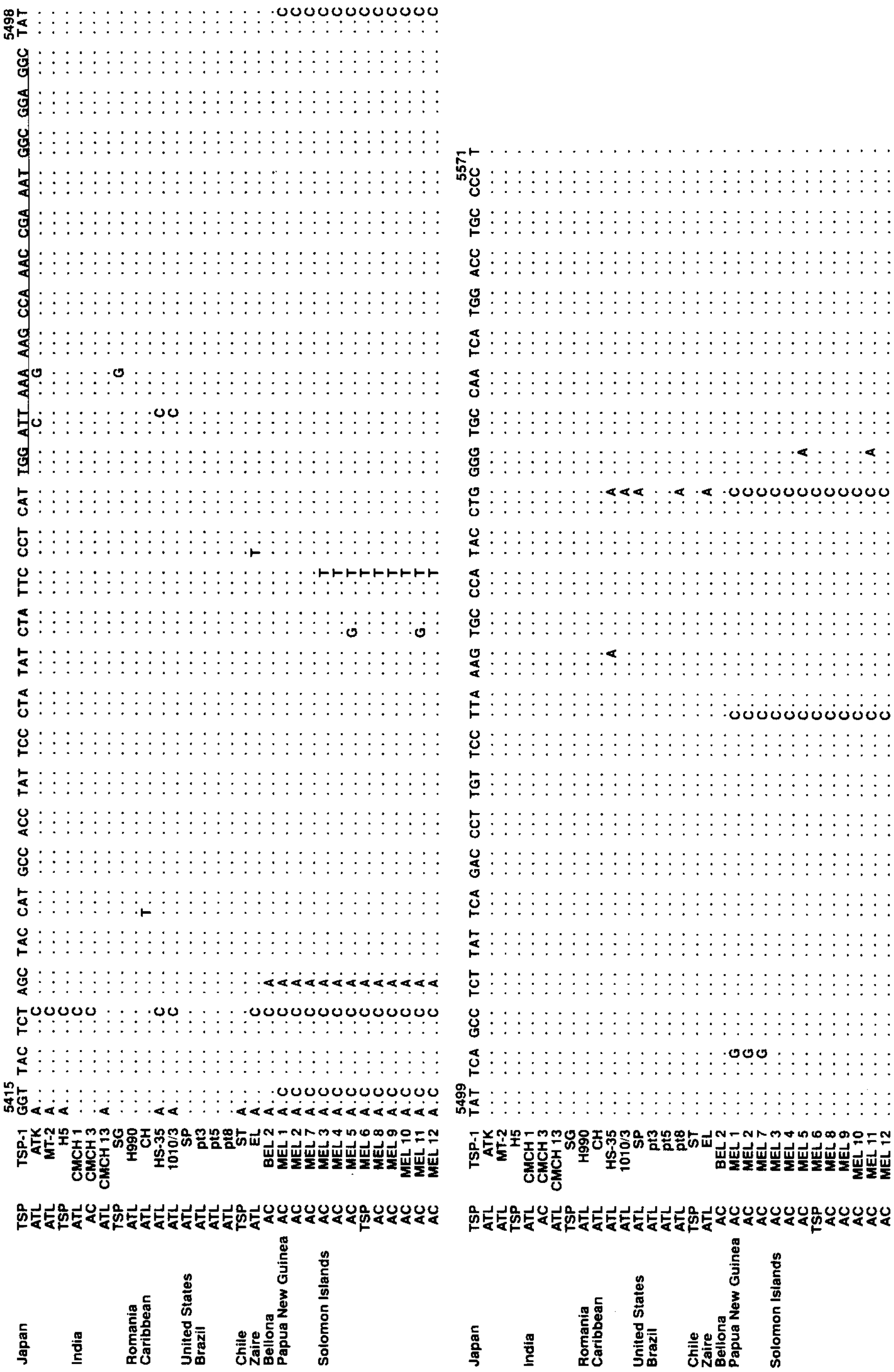


Fig. 3 (continued) **(D)** the 325-bp 5'-end of the gp46-encoding region of the *env* gene (bases 5247–5571).

E
6088
CC CCC TTT TCC TTG TCA CCT GTT CCC ACC CTA GGA TCC CGC CGA GCG GTA CCG GTG GCG GTC TGG CTT GTC TCC GCC CTG GCC ATG GGA 6185 GCC
6186
GGA GTG GCT GGC GAG ATT ACC GGC TCC ATG TCC CTC GCC TCA GGA AAG AGC CTC CTA CAT GAG GTG GAC AAA GAT ATT TCC CAA TTA ACT CAA 6281 GCA
Japan
China
Romania
Caribbean
India
United States
Martinique
Guadeloupe
French Guyana
Haiti
Brazil
Peru
Zaire
Cent African Rep
Mauritania
Guinea Bissau
Ivory Coast
Bellona
Papua New Guinea
Solomon Islands
TSP
ATL
AC
TSP-1
ATK
MT-2
H5
CMCH 1
CMCH 5
CMCH 13
SG
20
H990
CH
HS-35
1010/3
SP
1
4
5
7
8
11
pi3
pi5
pi8
10
EL
Z23
Z69
13
14
1711
15
19
BEL 2
MEL 1
MEL 2
MEL 7
MEL 3
MEL 4
MEL 5
MEL 6
MEL 8
MEL 9
MEL 10
MEL 11
MEL 12

Fig. 3 (continued) (E) the 374-bp gp21-encoding region of the *env* gene (bases 6068–6590).

F

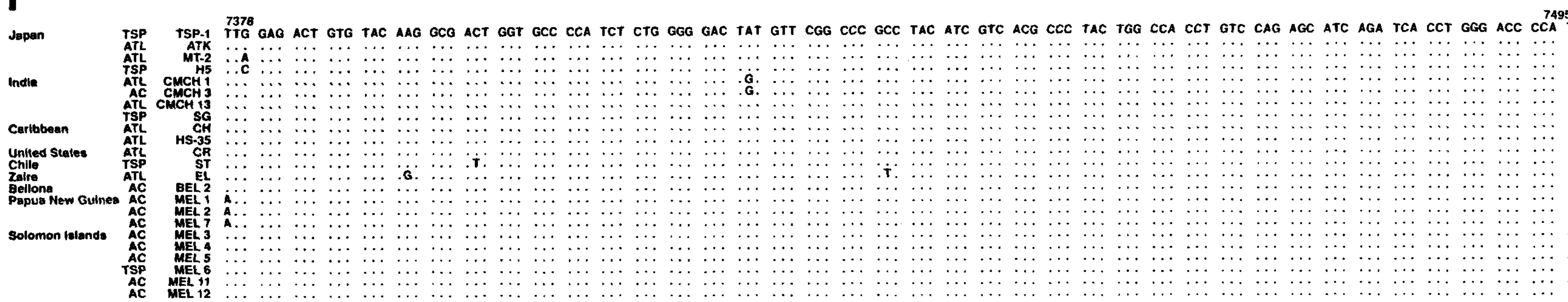


Fig. 3 (continued) **(F)** The 118-bp tax-encoding region of the *pX* gene (bases 7378–7495) in HTLV-I strains from Melanesian with and without HTLV-I myeloneuropathy. For comparison, corresponding sequences are shown for HTLV-I strains from patients with adult T-cell leukemia/lymphoma (ATL) or tropical spastic paraparesis/HTLV-I-associated myelopathy (TSP) and from asymptomatic carriers (AC) from Japan (ATK, MT-2, H5, TSP-1), India (CMCH 1, CMCH 5, CMCH 7, CMCH 13), the Caribbean basin (CH, HS-35, RD-1), the United States (CR), Chile (ST), Brazil (R-1), Zaire (EL, Z17, Z69), and the Polynesian Outlier Bellona (BEL 1, BEL 2). The B-cell epitopes on the carboxy-terminal p19 matrix protein-encoding *gag* gene and on the amino-terminal gp46 external envelope glycoprotein-encoding *env* gene are underlined.

Table 2
Sequence Comparison Between HTLV-I Strains from 12 Melanesians with a Japanese Strain of HTLV-I (TSP-1), Based on 1031 Nucleotides in Selected Regions of the *gag*, *pol*, and *env* Genes

		Percent divergence from TSP-1			Percent transitions		
Country of origin	Virus name	*gag*	*pol*	*env*	*gag*	*pol*	*env*
Papua New Guinea	HTLV-I MEL 1	5.2	10	7.6	80	79	81
	HTLV-I MEL 2	5.2	10	7.4	80	79	81
	HTLV-I MEL 7	5.2	10	7.4	80	79	81
Solomon Islands	HTLV-I MEL 3	5.7	7.9	7.3	91	82	78
	HTLV-I MEL 4	5.7	7.9	7.2	91	82	78
	HTLV-I MEL 5	5.2	7.9	8.0	90	82	79
	HTLV-I MEL 6	5.7	7.9	7.2	91	82	78
	HTLV-I MEL 8	5.7	7.9	7.2	91	82	78
	HTLV-I MEL 9	5.7	7.9	7.2	91	82	78
	HTLV-I MEL 10	5.7	7.9	7.2	91	82	78
	HTLV-I MEL 11	5.2	7.9	7.7	90	82	79
	HTLV-I MEL 12	5.2	7.9	6.9	90	82	78

otide sequence of the HTLV-I strain MEL 6 from a Melanesian with spastic paraparesis diverged by 11 of 192 bases (5.7%) in the *gag* gene, 11 of 140 bases (7.9%) in the *pol* gene, 50 of 699 bases (7.2%) in the *env* gene, and 0 of 118 base in the *pX* gene (Table 2). The overall sequence similarity between the Melanesian HTLV-I strain MEL 6 and cosmopolitan HTLV-I strain TSP-1 was 93.7% (72 substitutions in 1149 nucleotides). By contrast, when aligned and compared to the Japanese prototype HTLV-I strain ATK, few base substitutions, primarily transitions, were found in each of the gene regions of HTLV-I strains from TSP/HAM patients from other geographical regions. Between HTLV-I strains from the Solomon Islands and those from Papua New Guinea, the genetic variation was 3.4–4.2% (32- to 39-nucleotide differences), whereas it was 0.2–0.9% (2- to 8-nucleotide differences) among virus strains from the Solomon Islands. The sequence variation among the HTLV-I strains from the Solomon Islands was owing entirely to differences in the *env* gene.

In the *gag* gene region, strain TSP-1 from Japan, which was identical to some HTLV-I strains from southern India, was more closely related to strain RD-1 from Jamaica (98.4%) than to strain MEL 6 from the Solomon Islands (94.3%) (Fig. 3A,B). None of the nucleotide substitutions were shared between strains RD-1 and MEL 6.

Multiple base substitution, some resulting in amino acid changes, were found in the 5'-end of the gp46-encoding region of the *env* gene at positions 5251–5263 (Fig. 3D). The base changes in this particular region were so distinct, compared to the corresponding region in cosmopolitan HTLV-I strains, that they served as a "signature sequence" for the Melanesian HTLV-I strains. Similarly, variable regions were noted in the gp21 transmembrane protein-encoding *env* gene (e.g., positions 6113–6128 and 6410–6434) (Fig. 3E).

HTLV-I strains BEL 1 and BEL 2 from the Polynesian Outlier within the Solomon archipelago were genetically much more similar to cosmopolitan than to Melanesian strains of HTLV-I, differing by only 16 (1.4%) of 1149 nucleotides from the Japanese HTLV-I strain TSP-1. Consistent with the high degree of sequence similarity elsewhere in the viral genome, there was only one base difference between the strain TSP-1 and the strains from Bellona in the 140-bp region of the *pol* gene (Fig. 3C). However, nearly all of the base substitutions in the *gag* and *env* genes of HTLV-I strains BEL 1 and BEL 2 were distinct and were not shared with other cosmopolitan or Melanesian HTLV-I strains (Fig. 3A–D).

Sequence alignment and comparison of Melanesian HTLV-I strains with cosmopolitan strains (from the Caribbean, Chile, Zaire, Bellona, India, and Japan) disclosed differences between the Japanese HTLV-I strain TSP-1 and all other HTLV-I strains regardless of their geographic origin, suggesting a possible error in the originally reported sequence of HTLV-I TSP-1. For example, a discrepancy was noted in the gp21-encoding region of the *env* gene at position 6199 (Fig. 3E).

Deduced Amino Acid Sequence Diversity and Conservation

Base substitutions in the 31-amino acid region of the p24 capsid-encoding *gag* gene of the Melanesian

HTLV-I strains were silent. By contrast, in the 46-amino acid regions of the *pol* gene, the HTLV-I strains from Papua New Guinea and the Solomon Islands differed by four (8.7%) and two (4.4%) residues, respectively, from the Japanese HTLV-I strain TSP-1. All of the Melanesian HTLV-I strains, like the HTLV-I variant EL from Zaire, but unlike HTLV-I strains from Japan, the Caribbean, and the Americas, had an amino acid-altering substitution at position 4910, resulting in an asparagine to histidine change.

In the 108-amino acid gp46- and 124-amino acid gp21-encoding regions of the *env* gene, the Melanesian HTLV-I strains diverged by 5–6 (4.6–5.6%) and 5–7 (4.0–5.6%) amino acids, respectively, from HTLV-I TSP-1. Despite this overall degree of amino acid sequence dissimilarity, the amino-terminal neutralizing domain (comprised of amino acids 88–98, TrpIleLysLysProAsnArgAsnGlyGlyGly), encoded by bases 5463–5495, on the external envelope glycoprotein gp46 was totally conserved between Melanesian and cosmopolitan HTLV-I strains, except for HTLV-I ATK, in which the corresponding sequence was TrpThrLysLysProAsnArgAsnGlyGlyGly (Fig. 4). As further evidence for the structural conservation of functionally important domains, the amino acid sequence of the cleavage site on the gp61 envelope precursor protein, encoded by bases 6126–6143, and of the immunosuppressive region on the transmembrane envelope protein, encoded by bases 6330–6407, were identical between Melanesian and cosmopolitan HTLV-I strains.

The abovementioned region of sequence variability at the 5'-end of the *env* gene, in the region of the signal peptide, resulted in a predicted sequence for amino acids 16–21 of ProProIleLeuSerSer for HTLV-I strains MEL 1, MEL 2, and MEL 7 from Papua New Guinea and ProProIleLeuCysTyr for HTLV-I strains MEL 3, MEL 4, MEL 5, MEL 6, MEL 8, MEL 9, MEL 10, MEL 11, and MEL 12 from the Solomon Islands (Fig. 4). For cosmopolitan HTLV-I strains, the corresponding sequences were ProLeuIlePheGlyAsp (for HTLV-I strains ATK, H5, BEL 1, BEL 2), ProLeuIleLeuGlyAsp (for HTLV-I strains MT-2, CH, HS-35, 1010/3, pt5, ST, SP), and ProLeuIleLeuSerAsp for HTLV-I strain EL). It is unclear to what extent these changes are functionally important.

Phylogenetic Analysis

As determined by the unweighted pair-group method of assortment and the maximum parsimony method, and based on the premise that these gene phylogenies permit inferences to be made about virus phylogeny, evolutionary trees based on each gene region and those constructed by pooling sequences for the *gag*, *pol*, *env*, and *pX* gene regions of the Melanesian HTLV-I strains indicated that the virus strains from Papua New Guinea and the Solomon Islands formed a monophyletic group (Fig. 5A,B), that was phylogenetically distinct from a cluster formed by cosmopolitan strains of HTLV-I from Japan, the Polynesian Outlier Bellona, the Caribbean basin, and Africa. Moreover, these trees demonstrated that the Melanesian and cosmopolitan strains of HTLV-I have evolved along separate geographically dependent lineages.

Six equally parsimonious trees, requiring 637 nucleotide substitutions, were constructed from the 699 nucleotides comprising the gp46 (325-bp) and gp21 (374-bp)-encoding regions of the *env* gene. Bootstrap probabilities (in percentage) for the internal branches of the strict consensus tree, as determined for 1,000 resamplings by PAUP, were quite high (91–100%), demonstrating that the HTLV-I strains from Japan, India, and the Caribbean formed a monophyletic group and that the Melanesian HTLV-I strains evolved independently of cosmopolitan strains (Fig. 5A). The branch pattern of the UPGMA trees based on the same *env* gene sequences and on 1149 nucleotides spanning the *gag*, *pol*, *env*, and *pX* genes were nearly identical to the consensus PAUP tree, using either STLV-I strain PtM3 or HTLV-II strain Mo as the outgroup (Fig. 5B). Congruency of the phylogenetic trees, based on different gene regions by two algorithms, supported the evolutionary relationship between the HTLV-I strains from Melanesia and other geographic regions.

Discussion

In this study, we amplified and sequenced representative regions of the HTLV-I *gag*, *pol*, *env*, and *pX* gene from a Solomon Islander with TSP/HAM and from virus-infected Melanesians without disease to ascertain the overall as well as the gene-specific sequence variation. Considering the presumed slow mutation rate of HTLV-I, a high degree of sequence similarity is not unexpected among HTLV-I strains within isolated populations. As demonstrated, the nucleotide sequences of the HTLV-I strain from a Solomon Islander with TSP/HAM and virus strains from a symptomatically infected Melanesian Solomon Islanders were nearly identical. These observations are in accord with previous reports demonstrating the absence of disease-specific long terminal repeat and *env* gene sequences in HTLV-I strains from

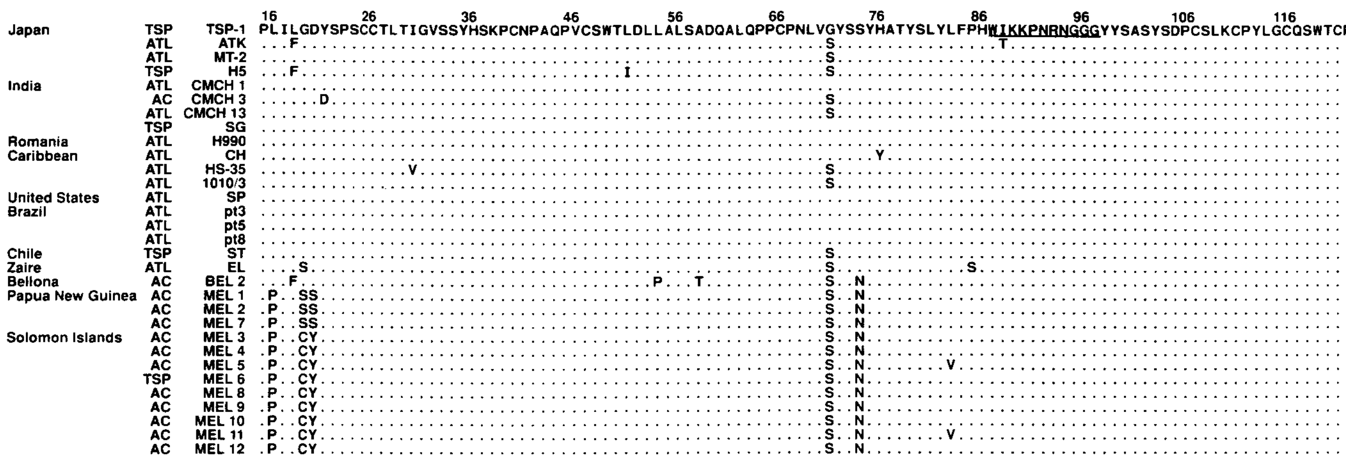


Fig. 4. Comparison of deduced or predicted amino acid sequences of the 5' end of the gp46 external envelope glycoprotein-encoding region of the *env* gene of HTLV-I strains from Melanesia, with representative strains from other geographical regions, including strains ATK, MT-2, H5, and TSP-1 from Japan; strains CMCH 1 and CMCH 13 from southern India; strains CH and HS-35 from the Caribbean basin; strain ST from Chile; strain EL from Zaire; and strain BEL 1 from the Polynesian Outlier Bellona. The amino-terminal neutralizing epitope on gp46 external envelope glycoprotein is underlined. Dots indicate homologous sequence with HTLV-I strain ATK. The single letter amino acid code was used: A, Ala; C, Cys; D, Asp; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

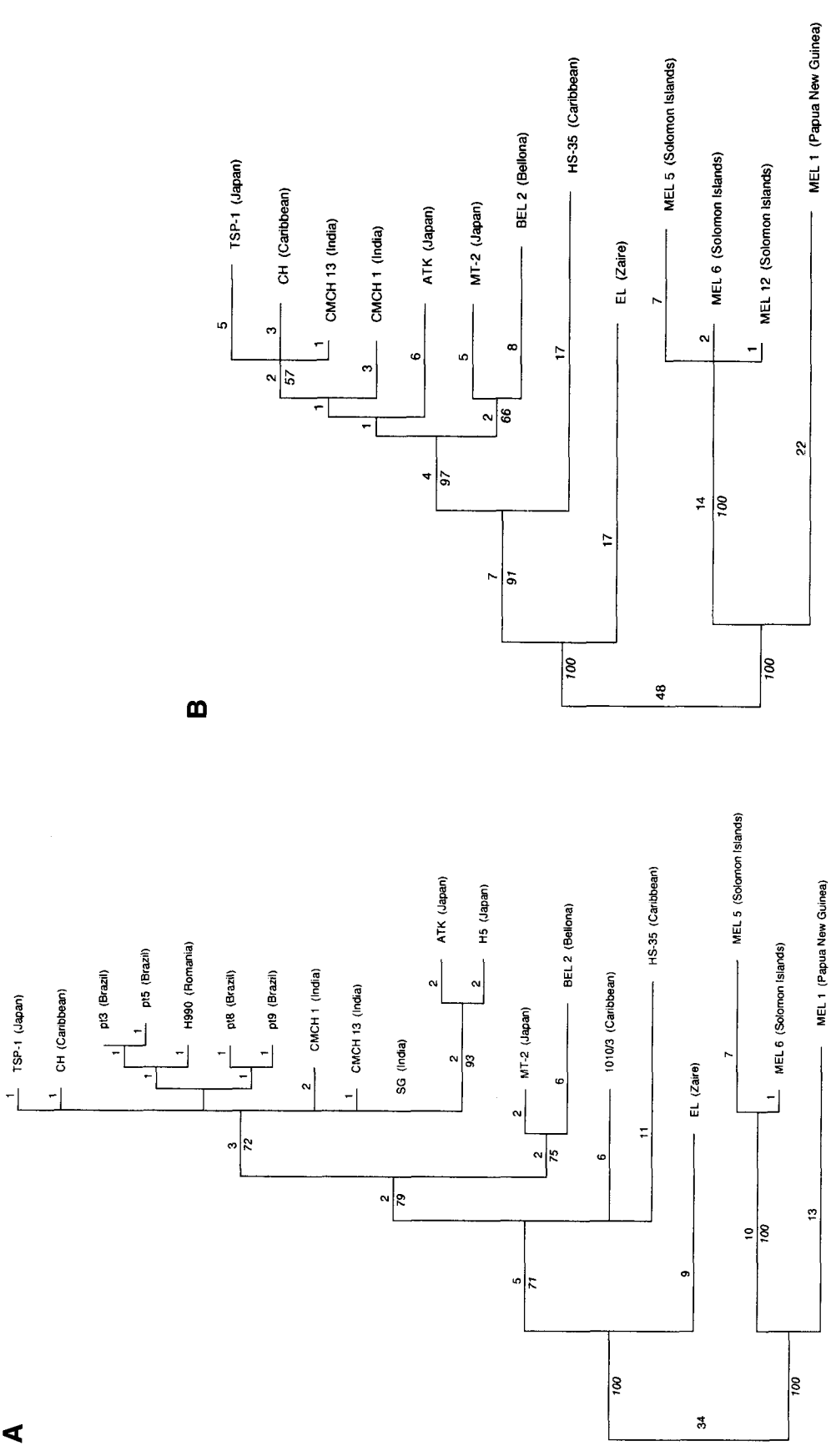


Fig. 5. Strict consensus phylogenetic trees, constructed by the maximum parsimony method, based on **(A)** both the 325-bp gp46-encoding and the 374-bp gp21-encoding regions of the *env* gene, and **(B)** 1149-bp *gag*, *pol*, *env*, and *pX* genes of HTLV-I strains from Melanesia (MEL 1 and MEL 5) and representative HTLV-I strains from other geographical regions, including Japan (ATK, MT-2, TSP-1, H5), southern India (CMCH 1 and CMCH 13), the Caribbean basin (CH, HS-35), Brazil (pt3), Zaire (EL), and the Polynesian Outlier Bellona (BEL 2). The tree was rooted by assuming (A) STLV-I strain PtM3 from Indonesia and (B) HTLV strain Mo from US as the outgroup. Branch lengths, drawn proportional to the number of nucleotide substitutions, are given above each branch, and bootstrap probabilities (in percentage), calculated from 1,000 resamplings by PAUP, are given below the internal branches.

patients with HTLV-I-caused diseases (1–7). Similarly, the virus strains from the Hagahai exhibited little sequence diversity. Thus, the considerable sequence similarity among HTLV-I strains from Papua New Guinea and the Solomon Islands is consistent with a shared common ancestral virus, which may have been introduced by one of several founder populations and was maintained by vertical (as from mother-to-child via virus-infected breast milk) and horizontal (as from husband-to-wife via virus-infected genital secretions) transmission.

Sequence analyses of HTLV-I strains indicate a geographic-dependent genetic similarity (4,8–11). Restriction analysis of enzymatically amplified LTR sequences of cosmopolitan HTLV-I strains from various geographic locales indicate the existence of three subtypes (11). In this regard, base substitutions and regions of variability unique to the Melanesian HTLV-I strains may also prove useful for direct genotyping of HTLV-I isolates (45; V. R. Nerurkar, K.-J. Song, and R. Yanagihara, unpublished observations). In addition, genotype-specific oligonucleotide primer-based PCR should facilitate studies aimed at elucidating the molecular epidemiology of the Melanesian strains of HTLV-I (46,47).

Considerable sequence conservation of neutralizing domains on the amino-terminal region and midregion of the external envelope glycoprotein have been found in HTLV-I strains, irrespective of their geographic origin (7,18,48,49). As determined by cross-neutralization assays using vesicular stomatitis virus pseudotypes bearing envelope antigens of HTLV-I, the neutralizing epitopes of cosmopolitan and Melanesian HTLV-I strains were functionally indistinguishable, suggesting that HTLV-I exists as a single serotype worldwide (50). In further support, rabbits injected intravenously with a T-cell line harboring the Melanesian HTLV-I strain MEL 5 have been protected from infection by administering human IgG purified from asymptomatically infected Japanese carriers (51). Collectively, these data infer that humoral immunity alone may afford adequate protection against HTLV-I infection and that either genetically engineered or synthetic peptide-based subunit vaccines would protect against infection with cosmopolitan and Melanesian HTLV-I strains.

Changing patterns of communicable diseases and the acquisition of new infectious agents, such as hepatitis B virus, have occurred among the Hagahai since their sustained contact with outsiders, beginning in 1984 (27). However, it is unlikely that HTLV-I has been recently introduced to the Hagahai because of the ages of the infected individuals and the nonrandom distribution of infection (52). Moreover, sera collected from other fringe-highland populations in the 1950s, at the time of first contact, indicate similarly high prevalences of HTLV-I infection. Thus, HTLV-I variants from other populations in Papua New Guinea, and possibly West New Guinea, are likely to be genetically similar to the virus strains from the Hagahai. Unfortunately, since there are no other HTLV-I isolates from Papua New Guinea to date, it is unclear if the HTLV-I variant from the Hagahai are truly representative of virus strains circulating among other fringe-highland populations, which are genetically and culturally distinct from the Hagahai and have had contact for hundreds of years with many outside groups. This would be in accord with the demonstrated high degree of sequence similarity among HTLV-I strains from Melanesian Solomon Islanders and from Aboriginals living in widely separated regions in Australia (I. B. Bastian, personal communication). At the same time, it is uncertain if HTLV-I strains circulating among coastal New Guinean populations are more closely related to strains from the Solomon Islands, than to the viral isolates from the Hagahai.

Radiocarbon dating of samples from a rock shelter on Buka Island off the northeastern coast of Bougainville establishes human occupation on the outer islands of New Guinea and the Solomon Islands, as far as San Cristobal, during the late Pleistocene epoch approx 30,000 yr ago (53,54). Thus, human settlements in some of these smaller islands within Melanesia are as old as those found on the Greater Australian continent of New Guinea and Australia (called Sahul) (55,56). Although it is uncertain when HTLV-I was introduced into these once remote Melanesian populations, our sequence and phylogenetic analyses indicate that the HTLV-I strains from Melanesians of Papua New Guinea and the Solomon Islands are more closely related to each other than to cosmopolitan HTLV-I strains. This in accord with evolution from ancestral strains of virus introduced by one of several founder populations, possibly as early as 40,000 yr ago. By contrast, the genetic similarity between HTLV-I strains from the Polynesian Outlier Bellona and other cosmopolitan HTLV-I strains, rather than to Melanesian viral strains, indicates a different and possibly more recent source of infection, with subsequent independent evolution of HTLV-I in Bellona. In any event, the total absence of nonhuman primates in Melanesia and Polynesia precludes any possibility

that HTLV-I among these genetically and culturally distinct populations evolved recently from STLV-I, the simian homolog of HTLV-I. Further molecular genetic comparisons between the Australo-Melanesian variants of HTLV-I, now underway, and studies of future isolates of type C retroviruses from India, Southeast Asia, and Indonesia will more fully clarify the emergence, evolution, and dissemination of HTLV-I in the western Pacific region.

Acknowledgments

We thank James W. Nagle and Tracy C. DeLozier for technical assistance, Mark S. Godec for synthesizing the oligonucleotide primer, Naruya Saitou for helpful advice with the phylogenetic analysis, and Gary W. Smythers of the Frederick Biomedical Supercomputing Center at the National Cancer Institute-Frederick Cancer Research and Development Center in Frederick, MD for assistance with the VAX.

References

1. Daenke S., Nightingale S., Cruickshank J. K., and Bangham C. R. M. (1990) *J. Virol.* **64,** 1278–1282.
2. De B. K., Lairmore M. D., Griffis K., Williams L. J., Villinger F., Quinn T. C., Brown C., Nzilambi N., Sugimoto M,. Araki S., and Folks T. M. (1991) *Virology* **182,** 413–419.
3. Kinoshita T., Tsujimoto A., and Shimotohno K., (1991) *Int. J. Cancer* **47,** 491–495.
4. Paine E., Garcia J., Philpott T. C., Shaw G., and Ratner L. (1991) *Virology* **182,** 111–123.
5. Gessain A., Gallo R. C., and Franchini G. (1992) *J. Virol.* **66,** 2288–2295.
6. Gonzales-Dunia D., Komurian-Pradel F., Chirinian-Syan S., de Thé G., Brahic M., and Ozden S. (1993) *AIDS Res. Hum. Retroviruses* **9,** 337–342.
7. Nerurkar V. R., Babu P. G., Song K.-J., Melland R. R., Gnanamuthu C., Saraswathi N. K., Chandy M., Godec M. S., John T. J., and Yanagihara R. (1993) *J. Gen. Virol.*, **74,** 2799–2805.
8. Komurian F., Pelloquin F., and de Thé G. (1991) *J. Virol.* **65,** 3770–3778.
9. Ratner L., Philpott T., and Trowbridge D. B. (1991) *AIDS Res. Hum. Retroviruses* **7,** 923–941.
10. Schulz T. F., Calabro M.-L., Hoad J. G., Carrington C. V. F., Matutes E., Catovsky D., and Weiss, R. A. (1991) *Virology* **184,** 483–491.
11. Komurian-Pradel F., Pelloquin F., Sonoda S., Osame M., and de Thé G. (1992) *AIDS Res. Hum. Retroviruses* **8,** 429–434.
12. Yanagihara R., Nerurkar V. R., and Ajdukiewicz A. B. (1991) *J. Infect. Dis.* **164,** 443–449.
13. Yanagihara R., Nerurkar V. R., Garruto R. M., Miller M. A., Leon-Monzon M. E., Jenkins C. L., Sanders R. C., Liberski P. P., Alpers M. P., and Gajdusek D. C. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1446–1450.
14. Gessain A., Yanagihara R., Franchini G., Garruto R. M., Jenkins C. L., Ajdukiewicz A. B., Gallo R. C., and Gajdusek D. C. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 7694–7698.
15. Saksena N. K., Sherman M. P., Yanagihara R., Dube D. K., and Poiesz B. J. (1992) *Virology* **189,** 1–9.
16. Sherman M. P., Saksena N. K., Dube D. K., Yanagihara R., and Poiesz B. J. (1992) *J. Virol.* **66,** 2556–2563.
17. Gessain A., Boeri E., Yanagihara R., Gallo R. C., and Franchini G. (1993) *J. Virol.* **67,** 1015–1023.
18. Nerurkar V. R., Song K.-J., Saitou N., Melland R. R., and Yanagihara R. (1993) *Virology* **196,** 506–513.
19. Watanabe T., Seiki M., Tsujimoto H., Miyoshi I., Hayami M., and Yoshida M. (1985) *Virology* **144,** 59–65.
20. Watanabe T., Seiki M., Hirayama Y., and Yoshida M. (1986) *Virology* **148,** 385–388.
21. Saksena N. K., Hervé V., Sherman M. P., Duran J. P., Mathiot C., Müller M., Love J. L., Luguenno B., Barré Sinoussi F., Dube D. K., and Poiesz B. J. (1993) *Virology* **192,** 312–320.
22. Song K.-J., Nerurkar V. R., Saitou N., Lazo A., Blakeslee J. R., Miyoshi I., and Yanagihara R. (1994) *Virology*, **199,** 56–66.
23. Koralnik I. J., LoMonico A., Fullen J., Saxinger C., Gessain A., Gallo R. C., Markham P., Kalyanaraman V., Boeri E., Hirsch V., Allan J., Murthy K., Guo H.-G., Alford P., and Franchini G. (1994) *J. Virol.* **68,** 2693–2707.
24. Ajdukiewicz A., Yanagihara R., Garruto R. M., Gajdusek D. C., and Alexander S. S. (1989) *N. Engl. J. Med.* **321,** 615–616.
25. Kirkland M. A., Frasca J., and Bastian I. (1991) *Aust. NZ. J. Med.* **21,** 739–741.
26. Bastian I., Gardner J., Webb D., and Gardner I. (1993) *J. Virol.* **67,** 843–851.
27. Jenkins C., Dimitrakakis M., Cook I., Sanders R., and Stallman N. (1989) *Hum. Ecol.* **17,** 27–57.
28. Bhatia K., Jenkins C., Prasad M., Koki G., and Lombange J. (1989) *Hum. Biol.* **61,** 45–64.
29. Bhatia K., Davies R., Jenkins C., Jazwinska E., Koki G., and Serjeantson S. W. (1993) *Am. J. Phys. Anthropol.*, in press.
30. Seiki M., Hattori S., Hirayama Y., and Yoshida M. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3618–3622.
31. Nerurkar V. R., Miller M. A., Leon-Monzon M. E., Ajdukiewicz A. B., Jenkins C. L., Sanders R. C., Godec M. S., Garruto R. M., and Yanagihara R. (1992) *J. Gen. Virol.* **73,** 1805–1810.
32. Burger H., Weiser B., Flaherty K., Gulla J., Nguyen P.-N., and Gibbs R. A. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 11,236–11,240.
33. Gray G. S., Bartman T., and White M. (1989) *Nucleic Acids Res.* **17,** 7998.
34. Gray G. S., White M., Bartman T., and Mann D. (1990) *Virology* **177,** 391–395.

35. Tsujimoto A., Teruuchi T., Imamura J., Shimotohno K., Miyoshi I., and Miwa M. (1988) *Mol. Biol. Med.* **5,** 29–42.
36. Evangelista A., Maroushek S., Minnigan H., Larson A., Retzel E., Haase A., Gonzalez-Dunia D., McFarlin D., Mingioli E., Jacobson S., Osame M., and Sonoda S. (1990) *Microb. Pathog.* **8,** 259–278.
37. Hashimoto K., Lalkaka J., Fujisawa J.-I., Singhal B. S., Machigashira K., Kubota R., Osame M., and Yoshida M. (1993) *AIDS Res. Hum. Retroviruses* **9,** 495–498.
38. Malik K. T. A., Even J., and Karpas A. (1988) *J. Gen. Virol.* **69,** 1695–1710.
39. Mukhopadhyaya R. and Sadaie M. R. (1993) *AIDS Res. Hum. Retroviruses* **9,** 109–112.
40. Dekaban G. A., King E. E., Waters D., and Rice G. P. A. (1992) *AIDS Res. Hum. Retroviruses* **8,** 1201–1207.
41. Dube D. K., Sherman M. P., Saksena N. K., Bryz-Gornia V., Mendelson J., Love J., Arnold C. B., Spicer T., Dube S., Glaser J. B., Williams A. E., Nishimura M., Jacobson S., Ferrer J. F., del Pimo N., Guiruelas S., and Poiesz B. J. (1993) *J. Virol.* **67,** 1175–1184.
42. Devereux J., Haeberli P., and Smithies O. (1984) *Nucleic Acids Res.* **12,** 387–395.
43. Feng D.-F. and Doolittle R. F. (1987) *J. Mol. Evol.* **25,** 351–360.
44. Swofford D. L. (1993) *PAUP: Phylogenetic Analysis Using Parsimony,* version 3.1.
45. Nerurkar V. R. and Yanagihara R. (1992) *AIDS Res. Hum. Retroviruses* **8,** 1199–1200.
46. Yanagihara R. (1992) *Hum. Biol.* **64,** 843–854.
47. Yanagihara R. (1994) *Adv. Virus Res.* **43,** 147–186.
48. Melland R. R. (1992) MS thesis, Pennsylvania State University, State College, 70 pp.
49. Sherman M. P., Dube S., Spicer T. P., Kane T. D., Love J. L., Saksena N. K., Iannone R., Gibbs C. J. Jr., Yanagihara R., Dube D. K., and Poiesz B. J. (1993) *Cancer Res.* **53,** 6067–6073.
50. Hoshino H., Nakamura T., Tanaka Y., Miyoshi I., and Yanagihara R. (1993) *J. Infect. Dis.* **168,** 1368–1373.
51. Tanaka Y., Ishii K., Sawada T., Ohtsuki Y., Hoshino H., Yanagihara R., and Miyoshi I. (1993) *Blood* **82,** 3664–3667.
52. Yanagihara R., Jenkins C. L., Alexander S. S., Mora C. A., and Garruto R. M. (1990) *J. Infect. Dis.* **162,** 649–654.
53. Allen J., Godsen C., Jones R., and White J. P. (1988) *Nature* **331,** 707–709.
54. Wickler S. and Spriggs M. (1988) *Antiquity* **62,** 703–706.
55. Jones R. (1973) *Nature* **246,** 278–281.
56. Groube L., Chappell J., Muke J., and Price D. (1986) *Nature* **324,** 453–455.