Cellular and Molecular Biology ™ **41** (Suppl. I), S145-S161, 1995
Printed in France.

0145-5680/95
1995 Cell. mol. Biol.™

MOLECULAR PHYLOGENY AND DISSEMINATION OF HUMAN T-CELL LYMPHOTROPIC VIRUS TYPE I VIEWED WITHIN THE CONTEXT OF PRIMATE EVOLUTION AND HUMAN MIGRATION

Richard YANAGIHARA[1,✍], Naruya SAITOU[2], Vivek R. NERURKAR[3], Ki-Joon SONG[1], Ivan BASTIAN[4], Genoveffa FRANCHINI[5] and D. Carleton GAJDUSEK[1]

[1,✍] Laboratory of Central Nervous System Studies, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bldg. 36, Rm. 5B-21, Bethesda, Maryland 20892, USA
[2] Laboratory of Evolutionary Genetics, National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan
[3] Retrovirology Research Laboratory, Pacific Biomedical Research Center, University of Hawaii, Honolulu, Hawaii 96816
[4] Menzies School of Health Research, Casuarina, Darwin, Northern Territory 0811, Australia
[5] Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

Accepted March 10, 1995

Abstract - A renewed interest in the emergence and evolution of the primate T-cell lymphotropic viruses has followed the discovery of genetically distinct variants of human T-cell lymphotropic virus type I (HTLV-I) in Melanesia and Australia. Phylogenetic trees based on selected regions of the *gag*, *pol*, *env* and *pX* genes of HTLV-I from widely separated geographic regions and of simian T-cell lymphotropic virus type I (STLV-I) from African and Asian catarrhines, constructed using the neighbor-joining and maximum parsimony methods, indicated that the Australo-Melanesian and cosmopolitan strains of HTLV-I have evolved along separate geographically dependent lineages, with African STLV-I strains clustering with cosmopolitan HTLV-I strains and Asian STLV-I strains diverging from the common ancestral virus before the Australo-Melanesian HTLV-I strains. When viewed within the context of non-human primate evolution and human occupation of Australia and Melanesia, the rate of molecular change of HTLV-I and STLV-I is approximately 2.5 - 6.8 x 10^{-7} substitutions per site per year. Overall, the sequence and phylogenetic analyses are in accord with interspecies virus transmission among non-human primates, as well as between non-human primates and humans, with independent evolution of HTLV-I in Southeast Asia and in Africa, and with dissemination of HTLV-I by forced or voluntary movements of human populations. The immunosuppressive and T-cell activation properties of HTLV-I places at added risk these Australian Aboriginal and Melanesian populations, some of which are in imminent threat of infection with human immunodeficiency virus type 1.

Key words: Retroviridae, genetic diversity, phylogenetic tree, Sahul, Sunda, Melanesia, Australia, Wallacea, Pleistocene

INTRODUCTION

Sequence variants of human T-cell lymphotropic virus type I (HTLV-I), the etiologic agent of adult T-cell leukemia/lymphoma and HTLV-I-associated myelopathy/tropical spastic paraparesis, have been discovered among remote Melanesian populations in Papua New Guinea and the Solomon Islands (Gessain *et al.*, 1993, 1991; Nerurkar *et al.*, 1994b, 1993b; Saksena *et al.*, 1992; Sherman *et al.*, 1992;

Yanagihara *et al.*, 1991b, 1991c, 1990a) and among Aboriginals in Australia (Bastian *et al.*, 1993a). Unlike the so-called cosmopolitan strains of HTLV-I from Japan, India, the Middle East, the Caribbean basin, the Americas and Africa, which exhibit ≥96.5% sequence similarity among themselves (Dekaban *et al.*, 1992; Evangelista *et al.*, 1990; Gessain *et al.*, 1992; Gray *et al.*, 1990, 1989; Komurian *et al.*, 1991; Malik *et al.*, 1988; Nerurkar *et al.*, 1993a, 1995; Paine *et al.*, 1991; Ratner *et al.*, 1991; Schulz *et al.*, 1991; Shirabe *et al.*, 1990; Tsujimoto *et al.*, 1988), the Australo-Melanesian variants of HTLV-I diverge by approximately 7% from cosmopolitan strains of HTLV-I. The isolation of these genetically distinct variants of HTLV-I, particularly the viral strains from the Hagahai (Yanagihara *et al.*, 1990a, 1990b), a remote, hunter-horticulturalist group living in the fringe highlands of Papua New Guinea and having no prior contact with Africans or Japanese, has led to a renewed inquiry into the origin and evolution of HTLV-I.

To further clarify the genetic and evolutionary relationship between the geographic-specific genotypes or topotypes of HTLV-I, we aligned and compared sequences from selected regions of the *gag*, *pol*, *env* and *pX* genes of the Australo-Melanesian HTLV-I strains with HTLV-I strains from other widely separated geographic regions, as well as with corresponding sequences of simian T-cell lymphotropic virus type I (STLV-I) strains from Asian and African catarrhines. From the perspective of primate evolution and human migration, our analysis extends the conclusions drawn earlier regarding the probable interspecies virus transmission and the likely mutation rate(s) of the primate T-cell lymphotropic viruses (Ina and Gojobori, 1990). Furthermore, the accumulated data are in accord with the introduction of the ancestral strain(s) of the Australo-Melanesian HTLV-I topotype by one of several founder populations migrating from the then Southeast Asian landmass (Sunda) to the Greater Australian continent (Sahul) of present-day Australia and New Guinea, possibly during the late Pleistocene epoch 40,000 years before present.

MATERIALS AND METHODS

Gene Amplification and Nucleotide Sequencing

High-molecular weight DNA, extracted from the MSHR-1 cell line established from an asymptomatically infected Australian Aboriginal (Bastian *et al.*, 1993a) and DNA from other HTLV-I- and STLV-I-infected T-cell lines, including Brazil-R-1 from a Brazilian woman with adult T-cell leukemia (Song *et al.*, 1995), AGM22 from an African green monkey (*Cercopithecus aethiops*) from Kenya (Koralnik *et al.*, 1994) and ChM114-1 from a common chimpanzee (*Pan troglodytes*) from Sierra Leone (Tsujimoto *et al.*, 1985), were studied (Table 1). Oligonucleotide primers for polymerase chain reaction (PCR) and for direct DNA sequencing were derived from sequences of the Japanese HTLV-I strain ATK (Seiki *et al.*, 1983) for the amino-terminal p24-encoding region of the *gag* gene (bases 1423-1444, 5'-CCATCACCAGCAGCTAGATAGC-3' and bases 1560-1537, 5'-AGTTGCTGGTATTCTCGCCTTAAT-3'); the 3'-end of the *pol* gene (bases 4757-4778, 5'-CCCTACAATCCAACCAGCTCAG-3' and bases 4942-4919, 5'-GTGGTGAAGCTGCCATCGGGTTTT-3'); the amino-terminal and midportion gp46-encoding region of the *env* gene (bases 5194-5214, 5'-CCAACACCATGGGTAAGTTTC-3' and bases 5495-5476, 5'-GCCTCCGCCATTTCGGTTTG-3'; bases 5228-5246, 5'-TTTATTCTTCCAGTTCTGC-3' and bases 5495-5476, 5'-GCCTCCGCCATTTCGGTTTG-3'); the gp21-encoding region of the *env* gene (bases 6044-6067, 5'-TCAAGCTATAGTCTCCTCCCCCTG-3' and bases 6590-6613, 5'-GGGAGGTGTCGTAGCTGACGGAGG-3') and the 5'-end of the *pX* gene (bases 7358-7377, 5'-CGGATACCCAGTCTACGTGT-3' and bases 7516-7496, 5'-GAGCCGATAACGCGTCCATCG-3'). All primers were used at a final concentration of 1 μM in a reaction mixture of 100 μl comprised of 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, and 0.2 mM each dNTP and containing 0.5 μg to 1 μg of DNA and 2.5 units of *Thermus aquaticus* DNA polymerase (Perkin-Elmer/Cetus, Norwalk, CT). Reaction mixtures were cycled as described previously (Nerurkar *et al.*, 1994b). Enzymatically amplified DNA, purified using Centricon-100 microconcentrators (Amicon Division, Danvers, MA), were sequenced 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)(Nerurkar *et al.*, 1994b). Ambiguities were resolved by manual sequencing using the Sequenase version 2.0 DNA sequencing kit (U.S. Biochemicals, Cleveland, OH).

Sequence and phylogenetic Analyses

Nucleotide and deduced amino acid sequences of the *gag*, *pol*, *env* and *pX* gene regions of HTLV-I strains from various geographic regions, as well as sequences of STLV-I strains from Asian and African catarrhines (Table 1), were aligned and compared with sequences of HTLV-I strain ATK (Seiki *et al.*, 1983). Sequence alignments were facilitated by using the software package available on the VAX computer system, as part of the Genetics Computer Group.

Table 1 *HTLV-I and STLV-I strains used in genetic and phylogenetic analyses*

Virus	Strain	Host Species	Country or Place of Origin
HTLV-I	MEL 1	Homo sapiens	Papua New Guinea
	MEL 2		Papua New Guinea
	MEL 3		Solomon Islands
	MEL 4		Solomon Islands
	MEL 5		Solomon Islands
	MEL 6		Solomon Islands
	MSHR-1		Australia
	EL		Zaire
	HS-35		Caribbean
	CH		Caribbean
	ATK		Japan
	MT-2		Japan
	CMCH 13		India
	BEL 1		Bellona Island
	Brazil-R-1		Brazil
STLV-I	Si-2	Macaca fuscata	Japan
	JM86		Japan
	Matsu		Japan
	MM39-83	Macaca mulatta	Southern Asia
	AGM22	Cercopithecus aethiops	Kenya
	KIA	Papio cynocephalus	South Africa
	ChM114-1	Pan troglodytes	Sierra Leone

Rather than relying on the unweighted pair-group method of assortment (UPGMA), which assumes a constant evolutionary rate, we employed the neighbor-joining method (Saitou and Nei, 1987) and the maximum parsimony method (Swofford, 1993; Fitch, 1977) to construct phylogenetic trees, since both methods accommodate variable rates of evolutionary change. Evolutionary distances between the sequences (number of nucleotide substitutions) were estimated by the one-parameter (Jukes and Cantor, 1969) and two-parameter methods (Kimura, 1980) and these distances were used to construct neighbor-joining trees. Multiple options of PAUP (phylogenetic analysis using parsimony), version 3.1.1 (Swofford, 1993), were used for maximum parsimony analysis. Maximum parsimony and neighbor-joining trees were evaluated statistically by calculating bootstrap probabilities for 1,000 iterations using PAUP and NJBOOT2, respectively.

RESULTS AND DISCUSSION

Genetic Diversity of the primate T-Cell lymphotropic Viruses

Pairwise comparison of 579 nucleotides spanning selected regions of the *gag*, *pol*, *env* and *pX* genes of the Australo-Melanesian HTLV-I strains and representative cosmopolitan HTLV-I strains from Japan, India, Africa, the Caribbean and the Americas indicated a clear demarcation, with the Australo-Melanesian strains differing from the cosmopolitan strains by 6.0% to 7.4% and 3.7% to 5.8% at the nucleotide and amino acid levels, respectively (Table 2). The sequence variability between the HTLV-I strains from Melanesians of Papua New Guinea and the Solomon Islands and from Aboriginals of Australia was 3.3% to 3.8%, while the cosmopolitan HTLV-I strains, including those from equatorial Zaire, differed among themselves by 0.5% to 3.4%. Similarly, for the 955-nucleotide gp46- and gp21-encoding regions of the *env* gene, the Australo-Melanesian HTLV-I strains diverged by 6.6% to 7.5% from cosmopolitan HTLV-I strains, and the interstrain nucleotide variability of the Australo-

Table 2 *Nucleotide and amino acid homologies between HTLV-I and STLV-I isolates from various geographic areas*

		Percent Homology of Nucleotide and Amino Acid Sequences																
Virus strain		HTLV-I ATK	HTLV-I CMCH13	HTLV-I BEL1	HTLV-I MT-2	HTLV-I Brazil R-1	HTLV-I CH	HTLV-I HS-35	HTLV-I EL	STLV-I KIA	STLV-I CbM114-1	STLV-I AGM22	HTLV-I MEL5	HTLV-I MEL1	HTLV-I MSHR-1	STLV-I JM86	STLV-I Sr-2	STLV-I Mata	STLV-I MM39-83
HTLV-I	ATK		99,0	97,9	97,9	98,4	97,9	98,4	97,4	97,4	97,9	97,4	95,3	94,8	94,8	94,8	94,2	94,2	92,2
HTLV-I	CMCH13	98,8		97,9	99,0	99,5	99,0	99,5	98,4	98,4	99,0	98,4	96,3	95,8	95,8	95,3	94,8	94,8	93,2
HTLV-I	BEL1	98,6	98,8		96,9	97,4	96,9	97,4	96,3	96,3	96,9	96,3	95,3	94,8	94,8	94,8	94,2	94,2	92,2
HTLV-I	MT-2	98,6	99,1	98,6		98,4	97,9	98,4	97,4	97,4	97,9	98,4	95,3	94,8	94,8	94,2	93,7	93,7	92,2
HTLV-I	Brazil-R-1	98,3	99,1	98,3	99,0		99,5	99,0	97,9	97,9	98,4	97,9	95,8	95,3	95,3	95,8	95,3	95,3	93,7
HTLV-I	CH	98,3	99,5	98,3	98,6	99,0		98,4	97,4	97,4	97,9	97,4	95,3	94,8	94,8	95,3	94,8	94,8	93,2
HTLV-I	HS-35	97,8	98,3	97,8	98,1	98,1	97,8		97,9	97,9	98,4	97,9	95,8	95,3	95,3	95,8	95,3	95,3	92,7
HTLV-I	EL	96,6	97,1	96,6	96,9	96,9	96,6	97,1		99,0	98,4	97,9	96,3	96,3	95,8	95,3	94,8	94,8	93,2
STLV-I	KIA	96,6	97,1	96,6	96,9	96,9	96,6	97,1	97,9		98,4	97,9	96,9	96,9	96,3	95,3	94,8	94,8	93,7
STLV-I	CbM114-1	96,2	96,7	96,2	96,9	96,6	96,2	96,7	97,6	97,2		98,4	96,3	95,8	95,8	95,3	94,8	94,8	93,2
STLV-I	AGM22	95,9	96,4	95,9	96,6	96,6	95,9	96,4	97,2	97,6	96,9		95,8	95,3	95,3	94,8	94,2	94,2	92,7
HTLV-I	MEL5	93,4	94,0	93,8	93,8	93,8	93,4	94,5	94,0	94,5	93,8	93,4		97,4	98,4	96,9	96,3	96,3	93,7
HTLV-I	MEL1	93,1	93,6	93,4	93,4	93,4	93,1	93,6	93,8	93,6	92,8	92,8	96,6		97,9	95,8	95,3	95,3	92,7
HTLV-I	MSHR-1	92,6	93,1	93,1	93,1	93,1	92,8	93,4	92,9	93,1	92,8	92,1	96,7	96,2		96,3	95,8	95,8	93,2
STLV-I	JM86	89,8	90,2	90,2	90,0	90,3	90,0	91,2	90,2	90,5	89,3	89,5	90,7	90,2	89,6		99,5	99,5	93,2
STLV-I	Sr-2	89,1	89,5	89,5	89,3	89,6	89,3	90,5	89,8	90,2	89,0	89,1	90,0	89,8	89,0	98,6		100	93,7
STLV-I	Mata	89,0	89,3	89,3	89,1	89,5	89,1	90,3	89,6	89,8	88,8	89,0	89,8	89,6	88,8	98,5	99,8		93,7
STLV-I	MM39-83	88,1	88,6	88,4	88,8	88,6	88,8	88,8	88,8	89,1	88,6	88,8	88,3	88,1	87,1	86,7	87,1	87,2	

Pairwise homologies based on 579-bp sequencesof the *gag*, *pol*, *env* and *pX* genes are presented as a triangular matrix. Nucleotide homologies are presented in the lower half and amino acid homologies in the upper half (estimated by DISTANCE).

Melanesian HTLV-I strains was 2.7% to 3.6%.

African STLV-I strains AGM22, KIA and ChM114-1 were closely related to cosmopolitan HTLV-I strains, particularly to the Zairian HTLV-I strain EL, diverging by 2.1% to 2.8% in the 579-nucleotide region. In addition, despite originating from three catarrhine genera, the sequence similarities among the African STLV-I strains were 96.9% to 97.6% and 97.9% to 98.4% at the nucleotide and amino acid levels, respectively. By contrast, STLV-I strains Si-2, Matsu and JM86 from Japanese macaque (*Macaca fuscata*) and MM39-83 from rhesus macaque (*Macaca mulatta*) were almost equally distant (approximately 10% sequence divergence) from HTLV-I strains originating from Japan and India, as well as from Australia and Melanesia. Moreover, strain MM39-83 differed by 12.9% to 13.3% from the Japanese macaque STLV-I strains. These observations indicate that the HTLV-I strains in Japan and India could not have evolved recently from STLV-I from Japanese and rhesus macaques.

Regional genetic differences between STLV-I and HTLV-I in Asia and Africa may be attributed to one or more of the following scenarios: virus transmission between non-human primates and from non-human primates to humans may have occurred (and may still occur) more frequently in Africa than in Asia; the common ancestor of STLV-I may have emerged outside of Africa, with subsequent independent evolution in Asia and Africa; the rates of nucleotide substitution for STLV-I from Asian and African non-human primates may not be constant or similar; and as in Africa, virus variants that genetically link Asian STLV-I and Australo-Melanesian HTLV-I may have existed (or may still exist) in Southeast Asia. Still another scenario is that the ancestral STLV-I may have emerged in Africa, but has yet to be identified.

Conservation of the neutralizing Epitopes of human T-Cell lymphotropic Virus Type I

Unlike the marked genetic hypervariability of the envelope proteins of human immunodeficiency virus type 1 (HIV-1), the antigenic determinants for neutralization on the amino-terminal and central regions of the gp46 external envelope glycoprotein of HTLV-I exhibit a remarkable degree of sequence conservation (Dekaban *et al.*, 1992; Evangelista *et al.*, 1990; Gray *et al.*, 1990; Malik *et al.*, 1988; Melland 1992; Paine *et al.*, 1991; Ratner *et al.*, 1991; Schulz *et al.*, 1991; Sherman *et al.*, 1993; Tsujimoto *et al.*, 1988). Serological studies indicate that HTLV-I-infected Melanesians from Papua New Guinea and the Solomon Islands exhibit robust reactivities to synthetic peptides and recombinant proteins derived from immunodominant epitopes of cosmopolitan HTLV-I strains (Lal *et al.*, 1992a, 1992b; Yanagihara *et al.*, 1991a, 1991b, 1990b), indicating that some of these antigenic domains are shared among all HTLV-I strains.

As evidence for the structural conservation of functionally important domains, the amino acid sequences of the cleavage site on the gp61 envelope precursor protein, encoded by bases 6126 to 6143, and of the immunosuppressive region on the transmembrane envelope protein, encoded by bases 6330 to 6407, are identical between Australo-Melanesian and cosmopolitan HTLV-I strains (Bastian *et al.*, 1993a; Gessain *et al.*, 1993, 1991; Nerurkar *et al.*, 1993b). Similarly, the deduced amino acid sequences of the neutralizing epitope-spanning domains on the external envelope glycoprotein of the Australo-Melanesian HTLV-I strains are identical to those of cosmopolitan HTLV-I strains from Japan, India, Iran, Zaire, the Caribbean, the Americas and the Polynesian Outlier Bellona (Fig. 1). The threonine to isoleucine difference at position 89 between HTLV-I ATK and nearly all other HTLV-I strains, including those from Australia and Melanesia, appears to be functionally irrelevant (Palker *et al.*, 1992). Not unexpectedly then, the neutralizing epitopes of the cosmopolitan and Melanesian HTLV-I strains are functionally indistinguishable as determined by cross-neutralization assays (Benson *et al.*, 1994; Hoshino *et al.*, 1993). In further support, HTLV-I immune globulin prepared from asymptomatically infected Japanese carriers is effective in protecting against infection with the Melanesian HTLV-I variant (Tanaka *et al.*, 1993).

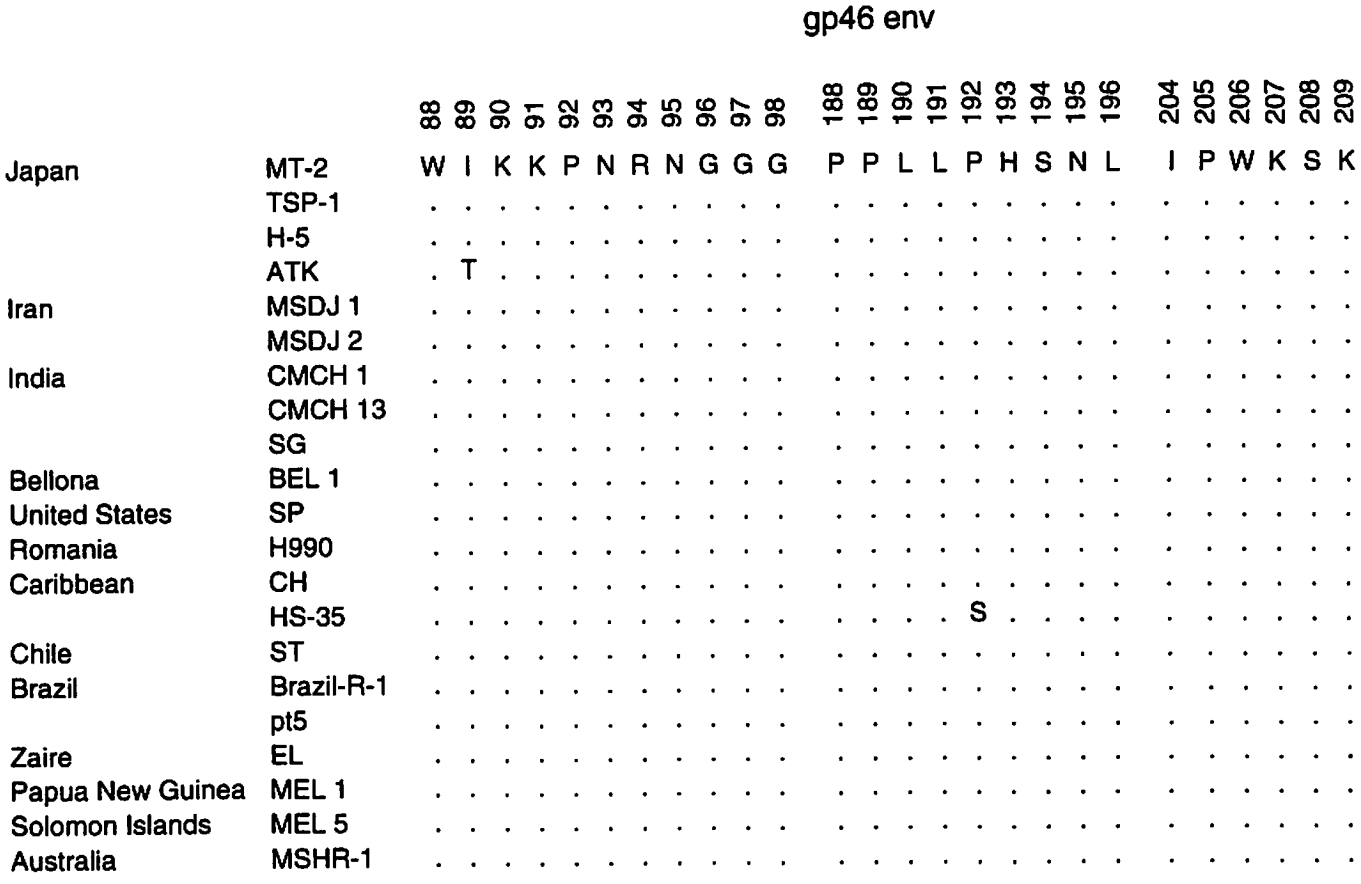

Fig. 1 *Comparison between the deduced amino acid sequences of the principal neutralizing domains on the gp46 external envelope glycoprotein of cosmopolitan and Australo-Melanesian strains of human T-cell lymphotropic virus type I.* Single letter amino acid code: G: glycine; H: histidine; I: isoleucine; K: lysine; L: leucine; N: asparagine; P: proline; R: arginine; S: serine; T: threonine; W: tryptophan

Sequence and functional conservation of the neutralizing epitopes on the external envelope glycoprotein of HTLV-I, irrespective of their geographic origin, may account for the absence of dual infection with cosmopolitan and Australo-Melanesian strains of HTLV-I in the same individual. Stability of the antigenic determinants for neutralization in HTLV-I strains identified to date suggests that HTLV-I may exist as a single serotype worldwide. Therefore, either genetically engineered recombinant protein- or synthetic peptide-based subunit vaccines might confer protective immunity against all HTLV-I strains. Moreover, the considerable conservation of functionally important domains on the envelope glycoproteins of HTLV-I suggests that they are under intense genetic constraint, that few amino acid substitutions are compatible with preservation of virus infectivity and replication, and that HTLV-I is a virus of great antiquity which arose from a common ancestor (Pique *et al.*, 1990; Yanagihara, 1994, 1992).

Identification of geographic-specific Genotypes of HTLV-I

Originally proposed for St. Louis encephalitis virus (Trent *et al.*, 1981), the concept of viral genotypes associated with specific geographic regions has been applied to other viruses, including dengue virus types 1 and 2 (Blok *et al.*, 1991; Rico-Hesse, 1990), poliovirus type 1 (Rico-Hesse *et al.*, 1987), Japanese encephalitis virus (Chen W.-R. *et al.*, 1992, 1990), human papillomavirus type 16 (Ho L. *et al.*, 1993; Chan *et al.*, 1992), hepatitis A virus (Robertson *et al.*, 1992), rabiesvirus (Smith *et al.*,

1992) and Ross River virus (Lindsay *et al.*, 1993). Accordingly, oligonucleotide primers derived from sequences unique to the gp46- and gp21-encoding regions of the *env* gene of the Australo-Melanesian variants of HTLV-I have been employed to discriminate HTLV-I strains into two major geographic-specific genotypes or topotypes, designated Australo-Melanesian and cosmopolitan (Nerurkar *et al.*, 1994a; Yanagihara, 1994, 1992; Nerurkar and Yanagihara, 1992). Each topotype of HTLV-I, which differs by approximately 7% at the nucleotide level, can be further classified into subtypes, which differ by approximately 3.5% (Table 3). In turn, individual viral strains within each subtype exhibit sequence similarity of ≥98%.

Table 3 *Geographic-specific genotypes or topotypes and subtypes of HTLV-I*

I. Australo-Melanesian topotype
A. Papua New Guinean
B. Solomon Islander
C. Australian Aboriginal
II. Cosmopolitan topotype
A. Zairian
B. Afroindoamerasian
1. West African
2. Caribbean
3. North American
4. South American
5. Japanese
6. Indian
7. Middle Eastern

Since the degree of genomic diversity between the HTLV-I subtypes within the Australo-Melanesian topotype is virtually identical, their temporal separation must be nearly the same, assuming similar rates of genetic change. Following this logic, it is improbable that the forced migration of more than 100,000 Melanesians (40,000 from the Solomon Islands) to Queensland, Australia, between 1870 and 1914, to work as indentured slaves alongside Aboriginals in mines and sugar cane plantations (Corris, 1973) is the basis for the present-day high prevalences of HTLV-I infection among Melanesians and Australian Aboriginals (Asher *et al.*, 1988; Bastian *et al.*, 1993b; Garruto *et al.*, 1990; Yanagihara *et al.*, 1990b). Similarly, the considerable genetic difference between the Australo-Melanesian and cosmopolitan topotypes of HTLV-I argues against the recent introduction of HTLV-I into Australia and Melanesia either by way of European explorers or other emigrees, such as the male pearl divers from southern Japan who migrated to Darwin and to Broome along the northern and northwestern coasts of Australia, respectively, beginning circa 1870 (Bain, 1982). In all likelihood, the existence and evolutionary separation of HTLV-I among Melanesians and Australian Aboriginals can be measured in millenia, rather than in centuries or decades.

Phylogenetic Relationship among the primate T-Cell lymphotropic Viruses

Previously published phylogenies of HTLV-I and STLV-I have relied heavily or exclusively on UPGMA (Gessain *et al.*, 1993, 1992; Ratner *et al.*, 1991; Saksena *et al.*, 1992; Sherman *et al.*, 1992), a method which assumes a constant evolutionary rate. Since the neighbor-joining method, which is based on the principle of minimum evolution, is superior to UPGMA (whether or not the actual rate of genetic change is constant) (Saitou and Nei, 1987) and since it is the most efficient method of producing the correct tree, when compared to the Fitch-Margoliash, maximum parsimony and maximum-likelihood methods of phylogenetic tree construction (Saitou and Imanishi, 1989), we chose this as the primary method to determine the phylogenetic relationship between the primate T-cell lymphotropic viruses.

Phylogenetic trees based on *gag* (92 bp), *pol* (140 bp), *env* (229 bp) and *pX* (118 bp) gene sequences (Fig. 2) and on *env* (955 bp) gene sequences encoding the amino-terminal and middle regions of the gp46 envelope glycoprotein and nearly the entire gp21 transmembrane protein (Fig. 3) of HTLV-I and STLV-I strains were constructed using the neighbor-joining and maximum parsimony methods. Neighbor-joining trees for each data set, obtained

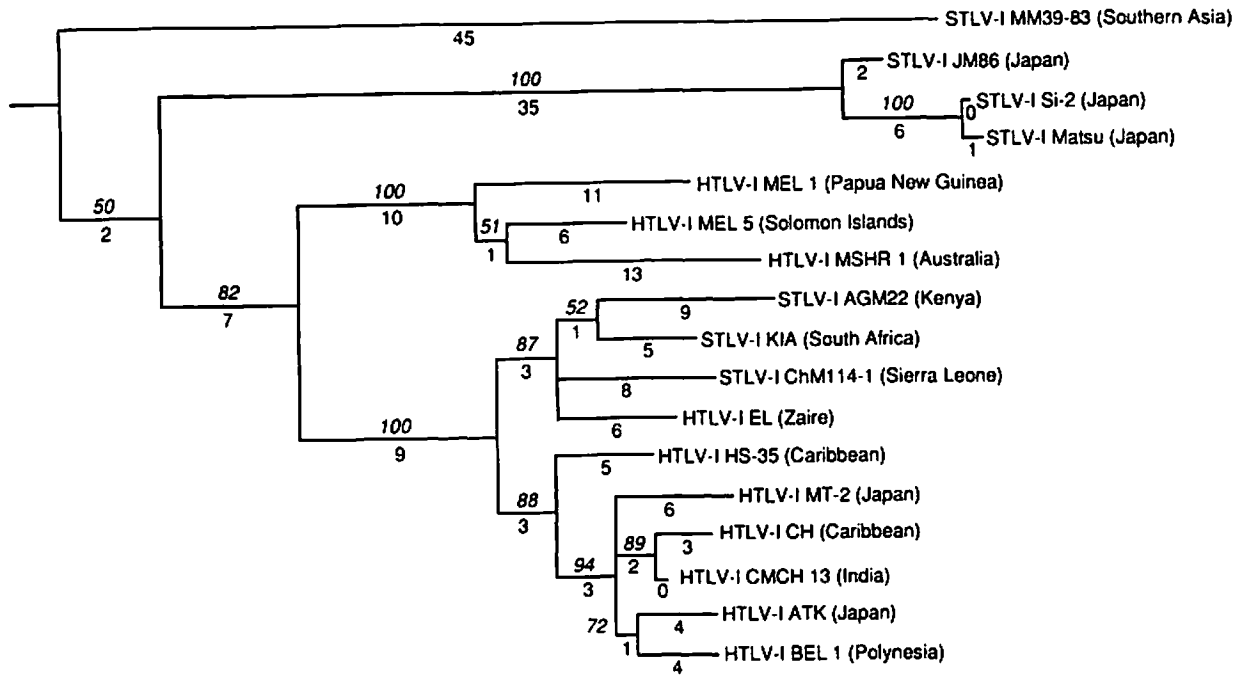

Fig. 2 *Phylogenetic tree based on 579 bp from selected regions of the gag, pol, env and pX genes of HTLV-I strains from various geographical areas, including Japan* (ATK, MT-2), *India* (CMCH 13), *Zaire* (EL), *the Polynesian Outlier Bellona* (BEL 1), *the Caribbean basin* (CH, HS-35), *Papua New Guinea* (MEL 1), *the Solomon Islands* (MEL 5) *and Australia* (MSHR-1), *and corresponding sequences of STLV-I strains from Asia* (MM39-83, JM86, Si-2, Matsu) *and Africa* (AGM22, KIA, ChM114-1), *constructed by the neighbor-joining method.* The tree was rooted by assuming HTLV-II strain Mo (Shimotohno *et al.*, 1985) as the outgroup. Branch lengths, given below each branch, are proportional to the estimated number of nucleotide substitutions, and bootstrap probabilities (in percentages), as determined for 1,000 resamplings by NJBOOT2, are given (in italics) above or beside the internal branches. The unavailability of *gag* and *pol* sequences for STLV-I strain PtM3 (Watanabe *et al.*, 1985) from an Indonesian pig-tailed macaque precluded its inclusion in this tree. GenBank accession numbers: U12101, U12102, U12103 and U11555 for AGM22; U12104, U12105, U12106 and U11556 for BEL 1; U12110, U12111, U12112 and U11561 for CMCH 13; U12116, U12117, U12118 and U11564 for ChM114-1; L20648, L20649, L20650 and L20652 for KIA; U12119, U12120, U12121 and M92818 for MSHR-1.

from evolutionary distances based on the one- and two-parameter methods and rooted by assuming HTLV-II (Shimotohno *et al.*, 1985) as the outgroup, were identical in their branching patterns, and the branch lengths were nearly the same (Figs. 2 and 3). By the maximum parsimony method, three equally parsimonious trees, requiring 381 nucleotide substitutions and with consistency indices of 0.795, were constructed from the 579-nucleotide sites.

In the neighbor-joining trees based on each data set, boot-strap probabilities (in percentages), as determined for 1,000 resamplings by NJBOOT2, for the respective nodes clustering the Australo-Melanesian HTLV-I strains and the cosmopolitan HTLV-I strains, including STLV-I strains from Africa, were ≥99%, demonstrating that the Australo-Melanesian and cosmopolitan strains of HTLV-I have evolved along separate geographically dependent lineages (Figs. 2 and 3). Two distinct lineages were also evident in the maximum parsimony trees based on a 522-nucleotide region of the gp21 transmembrane-encoding *env* gene for 36 strains of HTLV-I (data not shown). None of the HTLV-I strains were closely related to any of the Asian STLV-I strains, whereas the African STLV-I strains AGM22, KIA and ChM114-1 and the HTLV-I variant EL from equatorial Zaire clustered together in 87% of the bootstraps (Fig. 2).

Congruency of the phylogenetic trees, based on different gene regions using the neighbor-joining

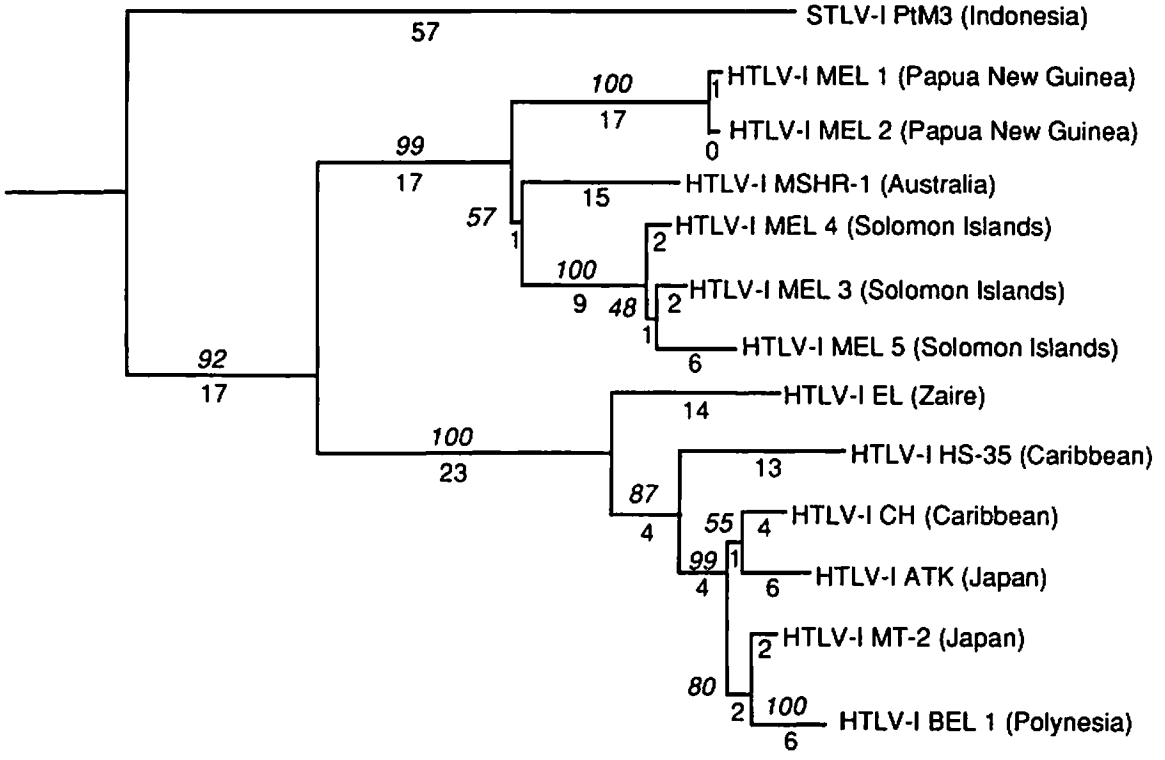

Fig. 3 *Neighbor-joining tree based on 955 bp from the gp46- and gp21-encoding regions of the env gene of Australo-Melanesian HTLV-I strains* (MEL 1 to MEL 6 and MSHR-1) *and cosmopolitan HTLV-I strains from representative geographic areas, including Japan* (ATK, MT-2), Zaire (EL), *the Polynesian Outlier Bellona* (BEL 1) *and the Caribbean basin* (CH, HS-35), *as well as the STLV-I strain PtM3 from a pig-tailed macaque* (*Macaca nemestrina*) *from Indonesia* (Watanabe *et al.*, 1985). The tree was rooted by assuming HTLV-II strain Mo (Shimotohno *et al.*, 1985) as the outgroup. Branch lengths, given below each branch, are proportional to the estimated number of nucleotide substitutions, and bootstrap probabilities (in percentages), as determined for 1,000 resamplings by NJBOOT2, are shown (in italics) above or beside the internal branches.

and maximum parsimony methods, validated the evolutionary relationship between the Australo-Melanesian and cosmopolitan strains of HTLV-I and the African and Asian strains of STLV-I. The failure of HTLV-I and STLV-I to segregate along evolutionary lines of their catarrhine hosts is consistent with interspecies virus transmission among non-human primates and between non-human primates and humans in the distant past and possibly more recently (Ina and Gojobori, 1990; Koralnik *et al.*, 1994; Song *et al.*, 1994). The degree of sequence diversity between the Australo-Melanesian and cosmopolitan strains of HTLV-I and between the Asian and African strains of STLV-I appears to be in accord with a slow accumulation of mutations over a prolonged period of time.

Estimation of Mutation Rate(s) for the primate T-Cell lymphotropic Viruses

Marked conservation in the tax-reading frame and greater heterogeneity in the rex-reading frame of the *pX* gene of HTLV-I were verified among the Australo-Melanesian HTLV-I strains and among the Asian and African STLV-I strains. Based on the maximum diversity of the *pX* gene, Ina and Gojobori (1990) concluded that the mutation rate of HTLV-I was somewhat faster than that of nuclear genes but much slower, by several orders of magnitude, than that of other RNA viruses, such as influenza A virus and HIV-1. Our estimates of the mutation rates for HTLV-I and STLV-I are in keeping with this prediction. That is, if, as judged by the archaeological record, human occupation in Australia and Melanesia

occurred 50,000 years before present, then the rate of molecular change for the 579-nucleotides spanning selected regions of the *gag*, *pol*, *env* and *pX* genes of HTLV-I is 6.8×10^{-7} substitutions per site per year, and the divergence between STLV-I strains from Japanese and rhesus macaques occurred 110,000 years before present.

Alternatively, if, as indicated by paleontological data, blood protein polymorphisms and restriction-enzyme and sequence analyses of mitochondrial DNA, Japanese and rhesus macaques diverged 0.3 to 1.8 million years ago (Hayasaka *et al.*, 1988a, 1988b; Melnick *et al.*, 1993; Melnick and Kidd, 1985), then the rate of genetic change for the 579-nucleotide region is $0.4 - 2.5 \times 10^{-7}$ substitutions per site per year, given the evolutionary distance (as determined by the two-parameter method) of 0.15 between the STLV-I strains from rhesus and Japanese macaque. Furthermore, if the nucleotide substitution rates of STLV-I and HTLV-I are similar, then HTLV-I diverged 140,000 to 850,000 years before present, given an evolutionary distance of 0.068 between the Australo-Melanesian and cosmopolitan HTLV-I strains. In this scenario, virus transmission from macaques to humans must have occurred at an early stage of modern human evolution.

Since the geographic separation between Japan and mainland Asia occurred after the last glaciation approximately 11,000 years ago (Melnick and Kidd, 1985), virus transmission between Japanese and rhesus macaques must have occurred before that time. In either case, the accumulated data support an archaic presence of STLV-I in Asia, probably predating macaque speciation, with interspecies transmission of the common ancestral strain of STLV-I long before the present-day habitats of their primate hosts were established in Africa and Asia (Song *et al.*, 1994).

Although these are conservative estimates for the nucleotide substitution rates of HTLV-I and STLV-I, even more liberal estimations would still be considerably slower than the rate of approximately 10^{-3} substitutions per site per year calculated for HIV-1. The marked disparity in the rates of molecular change between HTLV-I and HIV-1 can be accounted for by overall differences in their replication strategies, particularly the frequency of virus replication, as well as the fidelity or error proneness of their reverse transcriptase and the load of circulating free virus under selective pressure by the host immune response.

Implicit in the estimations of molecular change for HTLV-I and STLV-I is that they have been constant. In actuality, the tendency for some of the nucleotide changes to occur at identical positions in the *gag*, *pol* and *env* genes of the cosmopolitan and Australo-Melanesian HTLV-I strains and of the Asian and African STLV-I strains suggests a non-random nature to the molecular alterations. Thus, the rates of genetic change for HTLV-I and STLV-I may have varied considerably over time, and the accumulation of mutations may have occurred (and may be occurring) in a non-linear fashion. Based on this premise, the calculation of mutation rates may not be very meaningful, unless one accepts that any estimation of molecular change is merely an average of multiple rates. Studies aimed at identifying and characterizing the selective influences on and the mechanisms of genetic variation within virus populations may have more profound and immediate implications for predicting disease development and for understanding cellular tropism amd tissue targeting.

Peopling of Melanesia and the greater Australian Continent

As evidenced by radiocarbon dating of samples from a rockshelter on Buka Island off the northeastern coast of Bougainville, human occupation of the outer islands of New Guinea and of the Solomon Islands, as far as San Cristobal, dates to the late Pleistocene, approximately 30,000 years ago (Allen *et al.*, 1988; Wickler and Spriggs, 1988). Archeological sites of ancient human settlements, such as in the Huon peninsula of Papua New Guinea and near Arnhem Land in northern Australia, are somewhat older, dating to 50,000 years before present (Birdsell, 1977; Groube *et al.*, 1986; Jones, 1973; Roberts *et al.*, 1990).

Several millenia following these early settlements, Austronesians began to people the smaller islands in Melanesia and elsewhere in the Pacific beginning 1,000 to 5,000 years ago (Serjeantson, 1989, 1985; Serjeantson *et al.*, 1982). These populations, which were isolated by virtue of geography, culture and/or genetics from western influence, were first contacted, principally by Spain and Portugal, during the age of global exploration in the 16th century. Initial and later contacts, beginning at the end of the 18th century, with seamen on American, Australian and European vessels and their crews comprised of individuals from China, India and Africa, gave ample opportunity for the introduction of genes (and infectious diseases) into these insular communities. Although evidence for genetic admixture from these sources is limited, the occurrence of Machado-Joseph disease, an autosomal dominant neurological disorder characterized by limb spasticity, dystonia, dysarthria and ataxia (Rosenberg, 1978), among Aboriginals of Groote Eylandt, off the coast of Arnhem Land in the Gulf of Carpenteria (Burt *et al.*, 1993), is a probable example of genetic transfer from outsiders, in this case possibly from the Macassan trepang fishermen of Portuguese Timor.

Based on the analysis of gene frequency data for multiple polymorphic loci, including blood group antigens, serum proteins and HLA haplotypes (Nei and Roychoudhury, 1993; Serjeantson, 1989, 1985; Serjeantson *et al.*, 1982) and of morphological features, including hair texture, skin color and craniofacial measurements (Hanihara, 1993), the highland New Guineans and Australian Aboriginals are related. On the other hand, coastal New Guinean groups and inhabitants of the smaller islands within Melanesia, Micronesia and Polynesia are more closely related to Southeast Asians who migrated to these areas within the past 3,500 years (Bellwood, 1989: Chen LZ *et al.*, 1992). Despite its geographic proximity to Guadalcanal where the Australo-Melanesian topotype of HTLV-I is endemic, inhabitants of the neighboring Polynesian Outlier Bellona are infected with the cosmopolitan topotype of HTLV-I, suggesting a different source of infection and more recent introduction of the virus.

Emergence and early Evolution of HTLV-I

Although the accumulated genetic data are incomplete, they do permit a unifying hypothesis on the emergence of HTLV-I. Prior to the discovery of the Australo-Melanesian strains of HTLV-I, the long-standing dogma held that HTLV-I originated in Africa, but a strictly African origin of HTLV-I is inconsistent with our phylogenetic analysis. Instead, given the unique phylogenetic positions of the Australo-Melanesian and cosmopolitan strains of HTLV-I, it is likely that HTLV-I evolved independently in the Southeast Asia landmass of Sunda and in Africa. In this regard, the emergence and early evolution of HTLV-I is in accord with the hypothesis, based on recent analysis of human mitochondrial DNA sequence diversity in present-day human populations, that modern humans evolved from very isolated precursor populations in widely separated regions rather than from a single population in Africa (Harpending *et al.*, 1993). Explosive expansions in early human populations are felt to have occurred in different geographic regions during the Pleistocene epoch. One of these population expansions might have occurred in Sunda, where the common ancestor of the Australo-Melanesian HTLV-I topotype probably evolved. Two and possibly more of these ancestral virus variants may then have been introduced by one of several founder groups, possibly as early as 40,000 years ago, when, by intention or not, the hunter-gatherer Australoids left Sunda and arrived in the Greater Australian continent of Sahul, either via Sulawesi and Seram or via Bali and Timor (Birdsell, 1977). Similarly, cosmopolitan strains of HTLV-I may have evolved in isolated populations of early humans in Asia and Africa, while more recent migrations of modern humans may account for the present genetic relatedness between cosmopolitan HTLV-I strains from different geographic areas.

Human Migration and global Dissemination of HTLV-I

Like other viruses, the global dissemination of HTLV-I has been a dynamic process, with examples

of recent geographic spread through forced or voluntary movements of human populations (Gessain *et al.*, 1992; Yanagihara, 1994, 1992). Thus, as in other parts of the world, the distribution of HTLV-I in Oceania is not uniform, ranging from extremely high prevalences of infection in southwestern Japan to absolutely no evidence for infection in many Pacific island communities. Specifically, studies conducted on sera collected during the 1960s to 1980s from indigenous populations in Micronesia (Mariana Islands, Caroline Islands), Polynesia (Cook Islands, French Polynesia, Marquesas, Anuta, Tikopia, American and Western Samoa) and certain regions in Melanesia (Fiji and New Caledonia) have failed to disclose antibodies against HTLV-I (Asher *et al.*, 1988; Brindle *et al.*, 1988; Garruto *et al.*, 1990; Nicholson *et al.*, 1992; Morillon *et al.*, 1991; Tajima *et al.*, 1991).

Mention has already been made about the genetic evidence for the probable contact between Australian Aboriginals and Portuguese. By virtue of their expansive maritime explorations beginning in the 15th century, the Portuguese have been unfairly incriminated in the spread of HTLV-I. That they were the first Europeans to chart the unknown territories in the Indian and Pacific Oceans, including the coastal regions of New Guinea and Japan, is incontestable. However, high prevalences of HTLV-I infection have not been documented among the Portuguese (Cardoso *et al.*, 1989). Nevertheless, one hypothesis traces the origin of HTLV-I in Japan to the Portuguese seafarers and their African or Indian (Tamil) slaves, who frequented southwestern Japan during the 16th century (Gallo *et al.*, 1983; Kantha, 1986; Wong-Staal and Gallo, 1985). Heated debate continues to surround this issue, with the antithetical view claiming an ancient presence of HTLV-I in Japan, predating contact with Europeans, Africans or Indians (Hinuma, 1986; Ishida and Hinuma, 1986). Consistent with this latter hypothesis is the documented high prevalence of HTLV-I infection among the Aboriginals (Ainu) of Hokkaido in northern Japan (Ishida *et al.*, 1985). Both views may be correct. Recent sequence and phylogenetic analyses of HTLV-I strains from Japan and the Ryuukyu Islands, indicating two distinct genetic subtypes (Miura *et al.*, 1994; Ureta Vidal *et al.*, 1994), are consistent with the introduction of HTLV-I into Japan at two or more periods in the recent or distant past.

Absence of uniformity in the geographic distribution of HTLV-I, even in populations having decades- or centuries-long cohabitation with carrier groups, implies that HTLV-I is not efficiently transmitted by the sexual route and that long-term maintenance of the virus within a given population is dependent on complex interactions between environmental factors and social, behavioral and cultural practices. The comparative importance of these factors and the possible role of other infectious agents as co-factors within a given population are unknown. In this regard, the rapidly declining prevalences of HTLV-I infection in southwestern Japan (Oguma *et al.*, 1992) and the near disappearance of HTLV-I infection among third- and fourth-generation Japanese Americans in Hawaii (Blattner *et al.*, 1986; Ho GYF *et al.*, 1991) speaks to the relative ease with which HTLV-I can be eliminated from populations having unrestricted patterns of movement and marriage and whose hygienic and sociocultural milieu are compatible with early cessation of breastfeeding.

Recent population movements from the Middle East to Kerala in southern India and to Gujarat in western India, approximately 1000 to 1300 years ago (Balkrishnan, 1978; Undevia *et al.*, 1972), and the migration of more than 500,000 Indians to the Caribbean basin between 1838 to 1917 to toil as indentured laborers following the abolition of slavery by the British (Dutt, 1993), may be responsible for the demonstrated sequence similarity between HTLV-I strains from the Middle East and India and from India and the Caribbean islands (Nerurkar *et al.*, 1993a, 1995). The close sequence similarity among HTLV-I strains from western and central Africa (Central African Republic, Mauritania, Guinea Bissau), the Caribbean basin (Martinique, Guadeloupe, French Guyana, Haiti, Jamaica) and

South America (Brazil, Peru) may also be a consequence of the African slave trade to the New World, beginning in the 16th century (Gessain *et al.*, 1992; Song *et al.*, 1995).

The discovery of molecular variants of HTLV-I among Melanesians in Papua New Guinea and the Solomon Islands and among Aboriginals in Australia has provided an augmented perspective on the evolution of the primate T-cell lymphotropic viruses. Questions, however, remain about their geographic distribution, as well as about the existence of other genetically distinct variants of HTLV-I in other regions. We are thus in active pursuit of delineating if Wallace's Line (Wallace, 1860) or Weber's Line (Mayr, 1944), which separates the oriental zoogeographic fauna from that of Australia, forms the westernmost geographic border of the Australo-Melanesian topotype of HTLV-I. Such studies might provide further insights into the emergence of the Australo-Melanesian variants of HTLV-I by identifying viral strains that link the Australo-Melanesian and cosmopolitan topotypes of HTLV-I and the Australo-Melanesian HTLV-I strains with Asian STLV-I strains. Thus far, investigations of selected human populations occupying the presumed route of migration from Sunda to Sahul, including Ternate and Hiri in the Moluccas and Roti, Timor, Flores and Alor in the Lesser Sunda Islands, have failed to disclose evidence for HTLV-I infection (Song *et al.*, unpublished observations). Additional investigations in Wallacea and intensified attempts to verify HTLV-I infection among Javanese transmigrants to West New Guinea (Irian Jaya) (Anthony *et al.*, 1992), among coastal groups in Papua New Guinea and among lifelong residents of Vanuatu may yield additional clues about the emergence, evolution and dissemination of HTLV-I.

REFERENCES

Allen, J., Gosden, C., Jones, R. and White, J.P., Pleistocene dates for the human occupation of New Ireland, northern Melanesia. *Nature* 1988, **331**: 707-709.

Anthony, R.L., Jennings, G.B., Sie, A., Andersen, E.M. and Bangs, M.J., Javanese transmigrants acquire antibodies that recognize HTLV-I antigens within six months after resettlement in Irian Jaya, Indonesia. *Am. J. trop. Med. Hyg.* 1992, **47**: 139 (abstract 119).

Asher, D.M., Goudsmit, J., Pomeroy, K.L., Garruto, R.M., Bakker, M., Ono, S.G., Elliott, N., Harris, K., Askins, H., Eldadah, Z., Goldstein, A.D. and Gajdusek, D.C., Antibodies to HTLV-I in populations of the southwestern Pacific. *J. med. Virol.* 1988, **26**: 339-351.

Bain, M.A., *Full Fathom Five*. Artlook Books, Perth, 1982.

Balkrishnan, V., A preliminary study of genetic distances among some populations of the Indian sub-continent. *J. hum. Evol.* 1978, **7**: 67-75.

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

Bastian, I., Hinuma, Y. and Doherty, R.R., HTLV-I among Northern Territory Aborigines. *Med. J. Aust.* 1993b, **159**: 12-16.

Bellwood, P.S., The colonization of the Pacific: some current hypotheses. In: *The Colonization of the Pacific: A genetic Trail*, Hill, A.V.S. and Serjeantson, S.W. (eds.), Clarendon Press, Oxford, 1989, pp. 1-59.

Benson, J., Tschachler, E., Gessain, A., Yanagihara, R., Gallo, R.C. and Franchini, G., Cross-neutralizing antibodies against cosmopolitan and Melanesian strains of human T-cell leukemia/lymphotropic virus type I in inhabitants of Africa and the Solomon Islands. *AIDS Res. hum. Retrovir.* 1994, **10**: 91-96.

Birdsell, J.B., The recalibration of a paradigm for the first peopling of greater Australia. In: *Sunda and Sahul. Prehistoric Studies in Southeast Asia, Melanesia and Australia*, Allen, J., Golson, J. and Jones, R. (eds.), Acad. Press, New York, 1977, pp. 113-167.

Blattner, W.A., Nomura, A., Clark, J.W., Ho, G.Y.F., Nakao, Y., Gallo, R. and Robert-Guroff, M., Modes of transmission and evidence for viral latency from studies of human T-cell lymphotropic virus type I in Japanese migrant populations in Hawaii. *Proc. natn. Acad. Sci.* USA 1986, **83**: 4895-4898.

Blok, J., Gibbs, A.J., McWilliam, S.M. and Vitarana, U.T., NS 1 gene sequences from eight dengue-2 viruses and their evolutionary relationship with other dengue-2 viruses. *Arch. Virol.* 1991, **118**: 209-223.

Brindle, R.J., Eglin, R.P., Parsons, A.J., Hill, A.V.S. and Selkon, J.B., HTLV-I, HIV-I, hepatitis B and hepatitis delta in the Pacific and southeast Asia: a serological survey. *Epidemiol. Infect.* 1988, **100**: 153-156.

Burt, T., Blumbergs, P. and Currie, B., A dominant hereditary ataxia resembling Machado-Joseph disease in Arnhem Land, Australia. *Neurology* 1993, **43**: 1750-1752.

Cardoso, E.A., Robert-Guroff, M., Franchini, G., Gartner, S., Moura-Nunes, J.F., Gallo, R.C. and Terrinha, A.M., Seroprevalence of HTLV-I in Portugal and evidence of double retrovirus infection of a healthy donor. *Int. J. Cancer* 1989, **43**: 195-200.

Chan, S.-Y., Ho, L., Ong, C.-K., Chow, V., Drescher, B., Dürst, M., ter Meulen, J., Villa, L., Luande, J., Mgaya, H.N. and Bernard, H.-U., Molecular variants of human

papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. *J. Virol.* 1992, **66**: 2057-2066.

Chen, L.Z., Easteal, S., Board, P.G. and Kirk, R.L., Genetic affinities of Oceanic populations based on RFLP and haplotype analysis of genetic loci on three chromosomes. *Hum. Biol.* 1992, **64**: 1-15.

Chen, W.-R., Tesh, R.B. and Rico-Hesse, R., Genetic variation of Japanese encephalitis virus in nature. *J. Gen. Virol.* 1990, **71**: 2915-2922.

Chen, W.-R., Rico-Hesse, R. and Tesh, R.B., A new genotype of Japanese encephalitis virus from Indonesia. *Am. J. trop. Med. Hyg.* 1992, **47**: 61-69.

Corris, P., *Passage, Port and Plantation: A History of Solomon Islands Labour Migration, 1870-1914.* Melbourne University Press, Melbourne, 1973, 201 p.

Dekaban, G.A., King, E.E., Waters, D. and Rice, G.P.A., Nucleotide sequence analysis of an HTLV-I isolate from a Chilean patient with HAM/TSP. *AIDS Res. hum. Retrovir.* 1992, **8**: 1201-1207.

Dutt, E., Indians in the Caribbean diaspora. *India Abroad* 1993, **23**(1): 18-23.

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., Nucleotide sequence analysis of a provirus derived from an individual with tropical spastic paraparesis. *Microbiol. Pathogen.* 1990, **8**: 259-278.

Fitch, W.M., On the problem of discovering the most parsimonious tree. *Am. Natural.* 1977, **111**: 223-257.

Gallo, R.C., Sliski, A. and Wong-Staal, F., Origin of human T-cell leukemia-lymphoma virus. *Lancet* 1983, **ii**: 962-963.

Garruto, R.M., Slover, M., Yanagihara, R., Mora, C.A., Alexander, S.S., Asher, D.M., Rodgers-Johnson, P. and Gajdusek, D.C., High prevalence of human T-lymphotropic virus type I infection in isolated populations of the western Pacific confirmed by Western immunoblot. *Am. J. hum. Biol.* 1990, **2**: 439-447.

Gessain, A., Boeri, E., Yanagihara, R., Gallo, R.C. and Franchini, G., Complete nucleotide sequence of a highly divergent human T-cell leukemia (lymphotropic) virus type I (HTLV-I) variant from Melanesia: genetic and phylogenetic relationship with HTLV-I strains from other geographical regions. *J. Virol.* 1993, **67**: 1015-1023.

Gessain, A., Gallo, R.C. and Franchini, G., The low degree of human T-cell leukemia/lymphoma virus type I genetic drift *in vivo* as a means to follow viral transmission and movement of ancient human populations. *J. Virol.* 1992, **66**: 2288-2295.

Gessain, A., Yanagihara, R., Franchini, G., Garruto, R,M., Jenkins, C.L., Ajdukiewicz, A.B., Gallo, R.C. and Gajdusek, D.C., Highly divergent molecular variants of human T lymphotropic virus type I from isolated populations in Papua New Guinea and the Solomon Islands. *Proc. natn. Acad. Sci.* USA 1991, **88**: 7694-7698.

Gray, G.S., Bartman, T. and White, M., Nucleotide sequence of the core (gag) gene from HTLV-1 isolate MT-2. *Nucl. Acids Res.* 1989, **17**: 7998.

Gray, G.S., White, M., Bartman, T. and Mann, D., Envelope gene sequence of HTLV-I isolate MT-2 and its comparison with other HTLV-I isolates. *Virology* 1990, **177**: 391-395.

Groube, L., Chappell, J., Muke, J. and Price, D., A 40,000 year old occupation site at Huon Peninsula, Papua New Guinea. *Nature* 1986, **324**: 453-455.

Hanihara, T., Population prehistory of East Asia and the Pacific as viewed from craniofacial morphology: the basic populations in East Asia, VII. *Am. J. physiol. Anthropol.* 1993, **91**: 173-187.

Harpending, H.C., Sherry, S.T., Rogers, A.R. and Stoneking, M., The genetic structure of ancient human populations. *Curr. Anthropol.* 1993, **34**: 483-496.

Hayasaka, K., Gojobori, T. and Horai, S., Molecular phylogeny and evolution of primate mitochondrial DNA. *Mol. Biol. Evol.* 1988b, **5**: 626-644.

Hayasaka, K., Horai, S., Gojobori, T., Shotake, T., Nozawa, K. and Matsunaga, E., Phylogenetic relationships among Japanese, rhesus, Formosan, and crab-eating monkeys, inferred from restriction-enzyme analysis of mitochondrial DNAs. *Mol. Biol. Evol.* 1988a, **5**: 270-281.

Hinuma, Y., Seroepidemiology of adult T-cell leukemia virus (HTLV-I/ATLV): origin of virus carriers in Japan. *AIDS Res.* 1986, **2**: 517-522.

Ho, L., Chan, S.-Y., Burk, R.D., Das, B.C., Fujinaga, K., Icenogle, J.P., Kahn, T., Kiviat, N., Lancaster, W., Mavromara-Nazos, P., Labropoulou, V., Mitrani-Rosenbaum, S., Norrild, B., Pillai, M.R., Stoeker, J., Syrjaenen, K., Syrjaenen, S., Tay, S.-K., Villa, L.L., Wheeler, C.M., Williamson, A.-L. and Bernard, H.-U., The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J. Virol.* 1993, **67**: 6413-6423.

Ho, G.Y.F., Nomura, A.M.Y., Nelson, K., Lee, H., Polk, B.F. and Blattner, W.A., Declining seroprevalence and transmission of HTLV-I in Japanese families who immigrated to Hawaii. *Am. J. Epidemiol.* 1991, **134**: 981-987.

Hoshino, H., Nakamura, T., Tanaka, Y., Miyoshi, I. and Yanagihara, R., Functional conservation of the neutralizing domains on the external envelope glycoprotein of cosmopolitan and Melanesian strains of human T-cell leukemia/lymphoma virus type I. *J. infect. Dis.* 1993, **168**: 1368-1373.

Ina, Y. and Gojobori, T., Molecular evolution of human T-cell leukemia virus. *J. mol. Evol.* 1990, **31**: 493-499.

Ishida, T. and Hinuma, Y., The origin of Japanese HTLV-I. *Nature* 1986, **322**: 504.

Ishida, T., Yamamoto, K., Omoto, K., Iwanaga, M., Osato, T. and Hinuma, Y., Prevalence of human retrovirus in native Japanese: evidence for a possible ancient origin. *J. Infect.* 1985, **11**: 153-158.

Jones, R., Emerging picture of Pleistocene Australians. *Nature* 1973, **246**: 278-281.

Jukes, T.H. and Cantor, C.R., Evolution of protein molecules. In: *Mammalian Protein Metabolism*, Munro, H.N. (ed.), Acad. Press, New York, 1969, pp. 21-132.

Kantha, S.S., Portuguese role in spread of HTLV-I virus. *Nature* 1986, **321**: 733.

Kimura, M., A simple method for estimating rate of base substitutions through comparative studies of nucleotide sequences. *J. mol. Evol.* 1980, **16**: 111-120.

Komurian, F., Pelloquin, F. and de Thé, G., *In vivo* genomic variability of human T-cell leukemia virus type I depends more upon geography than upon pathologies. *J. Virol.* 1991, **65**: 3770-3778.

Koralnik, I.J., Boeri, E., Saxinger, W.C., Lo Monico, A., Fullen, J., Gessain, A., Guo, H.-G., Gallo, R.C., Markham, P., Kalyanaraman, V., Hirsch, V., Allan, J., Murthy, K., Alford, P., Slattery, J.P., O'Brien, S.J. and Franchini, G., Phylogenetic associations of human and simian T-cell leukemia/lymphotropic virus type I strains: evidence for interspecies transmission. *J. Virol.* 1994, **68**: 2693-2707.

Lal, R.B., Brodine, S., Kazura, J., Mbidde-Katonga, E., Yanagihara, R. and Roberts, C., Sensitivity and specificity of recombinant transmembrane glycoprotein ($r21^{env}$)-spiked Western blot for serological confirmation of human T-lymphotropic virus types I and II infection. *J. clin. Microbiol.* 1992a, **30**: 296-299.

Lal, R.B., Rudolph, D.L., Nerurkar, V.R. and Yanagihara, R., Humoral responses to the immunodominant *gag* and *env* epitopes of human T-lymphotropic virus type I among Melanesians. *Viral Immunol.* 1992b, **5**: 265-272.

Lindsay, M.D.A., Coelen, R.J. and MacKenzie, J.S., Genetic heterogeneity among isolates of Ross River virus from different geographical regions. *J. Virol.* 1993, **67**: 3576-3585.

Malik, K.T.A., Even, J. and Karpas, A., Molecular cloning and complete nucleotide sequence of an adult T cell leukaemia virus/human T cell leukaemia virus type I (ATLV/HTLV-I) isolate of Caribbean origin: relationship to other members of the ATLV/HTLV-I subgroup. *J. gen. Virol.* 1988, **69**: 1695-1710.

Mayr, E., Wallace's Line in the light of recent zoogeographic studies. *Q. Rev. Biol.* 1944, **19**: 1-14.

Melland, R.R., *Amino Acid Sequence Conservation of three immunodominant gag and env Epitopes of human T lymphotropic Virus Type I from Melanesian Families.* Master of Science dissertation. Pennsylvania State University, State College, 1992, 70 p.

Melnick, D.J., Hoelzer, G.A., Absher, R. and Ashley, M.V., mtDNA diversity in rhesus monkeys reveals overestimates of divergence time and paraphyly with neighboring species. *Mol. Biol. Evol.* 1993, **10**: 282-295.

Melnick, D.J. and Kidd, K.K., Genetic and evolutionary relationships among Asian macaques. *Int. J. Primatol.* 1985, **6**: 123-160.

Miura, T., Fukunaga, T., Igarashi, T., Yamashita, M., Ido, E., Funahashi, S.-I., Ishida, T., Washio, K., Ueda, S., Hashimoto, K., Yoshida, M., Osame, M., Singhal, B.S., Zaninovic, V., Cartier, L., Sonoda, S., Tajima, K., Ina, Y., Gojobori, T. and Hayami, M., Phylogenetic subtypes of human T-lymphotropic virus type I and their relations to the anthropological background. *Proc. natn. Acad. Sci.* USA 1994, **91**: 1124-1127.

Morillon, M., Guinet, M. and Lamy, O., Seroprevalence of HTLV I in New Caledonia (South Pacific). In: *Abstracts of the Asia-Pacific Congress of medical Virology*, Bangkok, Thailand, 1991, abstract P6-1.

Nei, M. and Roychoudhury, A.K., Evolutionary relationships of human populations on a global scale. *Mol. Biol. Evol.* 1993, **10**: 927-943.

Nerurkar, V.R, Achiron, A., Song, K.-J., Melland, R.R., Pinhas-Hamiel, O., Melamed, E., Shohat, B. and Yanagihara, R., Human T-cell lymphotropic virus type I in Iranian-born Mashhadi Jews: genetic and phylogenetic evidence for common source of infection. *J. med. Virol.* 1995, in press, accepted: november 4, 1994.

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., Sequence analysis of human T-cell lymphotropic virus type I strains from southern India: gene amplification and direct sequencing from whole blood blotted onto filter paper. *J. gen. Virol.* 1993a, **74**: 2799-2805.

Nerurkar, V.R., Song, K.-J., Saitou, N., Melland, R.R. and Yanagihara, R., Interfamilial and intrafamilial genomic diversity and molecular phylogeny of human T-cell lymphotropic virus type I from Papua New Guinea and the Solomon Islands. *Virology* 1993b, **196**: 506-513.

Nerurkar, V.R., Song, K.-J., Bastian, I.B., Garin, B., Franchini, G. and Yanagihara, R., Genotyping of human T-cell lymphotropic virus type I strains using Australo-Melanesian topotype-specific oligonucleotide primer-based polymerase chain reaction: insights into virus evolution and dissemination. *J. infect. Dis.* 1994a, **170**: 1353-1360.

Nerurkar, V.R., Song, K.-J., Melland, R.R. and Yanagihara, R., Genetic and phylogenetic analyses of human T-cell lymphotropic virus type I variants from Melanesians with and without spastic myelopathy. *Mol. Neurobiol.* 1994b, **8**: 155-173.

Nerurkar, V.R. and Yanagihara, R., Specificity of an oligonucleotide primer pair and of a single-base substitution in the amplification and detection of *env* gene sequences of HTLV-I variants from Papua New Guinea and the Solomon Islands. *AIDS Res. hum. Retrovir.* 1992, **8**: 1199-1200.

Nicholson, S.R., Efandis, T., Dimitrakakis, M., Karopoulos, A., Lee, H. and Gust, I.D., HTLV-I infection in selected populations in Australia and the western Pacific region. *Med. J. Aust.* 1992, **156**: 878-880.

Oguma, S., Imamura, Y., Kusumoto, Y., Nishimura, Y., Yamaguchi, K., Takatsuki, K., Tokudome, S. and Okuma, M., Accelerated declining tendency of human T-cell leukemia virus type I carrier rates among younger blood donors in Kumamoto, Japan. *Cancer Res.* 1992, **52**: 2620-2623.

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

Palker, T.J., Riggs, E.R., Spragion, D.E., Muir, A.J., Scearce, R.M., Randall, R.R., McAdams, M.W., McKnight, A., Clapham, P.R., Weiss, R.A. and Haynes, B.F., Mapping

of homologous, amino-terminal neutralizing regions of human T-cell lymphotropic virus type I and II gp46 envelope glycoproteins. *J. Virol.* 1992, **66**: 5879-5889.

Pique, C., Tursz, T. and Dokhelar, M.-C., Mutations introduced along the HTLV-I envelope gene results in a non-functional protein: a basis for envelope conservation? *EMBO J.* 1990, **9**: 4243-4248.

Ratner, L., Philpott, T. and Trowbridge, D.B., Nucleotide sequence analysis of isolates of human T-lymphotropic virus type I of diverse geographic origins. *AIDS Res. hum. Retrovir.* 1991, **7**: 923-941.

Rico-Hesse, R., Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* 1990, **174**: 479-493.

Rico-Hesse, R., Pallansch, M., Nottay, B.K. and Kew, O.M., Geographic distribution of wild poliovirus type 1 genotypes. *Virology* 1987, **160**: 311-322.

Roberts, R.G., Jones, R. and Smith, M.A., Thermoluminescence dating of a 50,000-year old human occupation in northern Australia. *Nature* 1990, **345**: 153-156.

Robertson, B.H., Jansen, R.W., Khanna, B., Totsuka, A., Nainan, O.V., Siegl, G., Widell, A., Margolis, H.S., Isomura, S., Ito, K., Ishizu, T., Moritsugu, Y. and Lemon, S.M., Genetic relatedness of hepatitis A virus strains recovered from different geographical regions. *J. gen. Virol.* 1992, **73**: 1365-1377.

Rosenberg, R.N., Joseph's disease. In: *The inherited Ataxias*, Kark, P., Rosenberg, R.N. and Schut, L. (eds.), Raven Press, New York, 1978, pp. 33-57.

Saitou, N. and Imanishi, T., Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining mehtods of phylogenetic tree construction in obtaining the correct tree. *Mol. Biol. Evol.* 1989, **6**: 514-525.

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

Saksena, N.K., Sherman, M.P., Yanagihara, R., Dube, D.K. and Poiesz, B.J., LTR sequence and phylogenetic analysis of a newly discovered variant of HTLV-I from Papua New Guinea. *Virology* 1992, **189**: 1-9.

Schulz, T.F., Calabró, M.-L., Hoad, J.G., Carrington, C.V.F., Matutes, E., Catovsky, D. and Weiss, R.A., HTLV-I envelope sequences from Brazil, the Caribbean, and Romania: clustering of sequences according to geographic origin and variability in an antibody epitope. *Virology* 1991, **184**: 483-491.

Seiki, M., Hattori, S., Hirayama, Y. and Yoshida, M., Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proc. natn. Acad. Sci.* USA 1983, **80**: 3618-3622.

Serjeantson, S.W., Migration and admixture in the Pacific. Insights provided by human leucocyte antigens. In: *Out of Asia. Peopling of the Americas and the Pacific*, Kirk, R. and Szathmary, E. (eds.), Australian National University Printery, Canberra, 1985, pp. 133-145.

Serjeantson, S.W., HLA genes and antigens. In: *The Colonization of the Pacific. A genetic Trail*, Hill, A.V.S. and Serjeantson, S.W. (eds.), Clarendon Press, Oxford, 1989, pp. 120-173.

Serjeantson, S.W., Ryan, D.P. and Thompson, A.R., The colonization of the Pacific: the story according to human leukocyte antigens. *Am. J. hum. Genet.* 1982, **34**: 904-918.

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., Sequence analysis of an immunogenic and neutralizing domain of the human T-cell lymphoma/leukemia virus type I gp46 surface membrane protein among various primate T-cell lymphoma/leukemia virus isolates including those from a patient with both human T-cell lymphoma/leukemia virus type I-associated myelopathy and adult T-cell leukemia. *Cancer Res.* 1993, **53**: 6067-6073.

Sherman, M.P., Saksena, N.K., Dube, D.K., Yanagihara, R. and Poiesz, B.J., Evolutionary insights on the origin of human T-cell lymphoma/leukemia virus type I (HTLV-I) derived from sequence analysis of a new HTLV-I variant from Papua New Guinea. *J. Virol.* 1992, **66**: 2556-2563.

Shimotohno, K., Takahashi, Y., Shimizu, N., Gojobori, T., Golde, D.W., Chen, I.S.Y., Miwa, M. and Sugimura, T., Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: an open reading frame for the protease gene. *Proc. natn. Acad. Sci.* USA 1985, **82**: 3101-3105.

Shirabe, S., Nakamura, T., Tsujihata, M., Nagataki, S., Seiki, M. and Yoshida, M., Retrovirus from human T-cell leukemia virus type I-associated myelopathy is the same strain as a prototype human T-cell leukemia virus type I. *Arch. Neurol.* 1990, **47**: 1258-1260.

Smith, J.S., Orciari, L.A., Yager, P.A., Seidel, H.D. and Warner, C.K., Epidemiologic and historical relationships among 87 rabies virus isolates as determined by limited sequence analysis. *J. infect. Dis.* 1992, **166**: 296-307.

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

Song, K.-J., Nerurkar, V.R., Pereira-Cortez, A.J., Yamamoto, M., Taguchi, H., Miyoshi, I. and Yanagihara, R., Sequence and phylogenetic analyses of human T-cell lymphotropic virus type I from a Brazilian woman with adult T-cell leukemia: comparison with virus strains from South America and the Caribbean basin. *Am. J. trop. Med. Hyg.* 1995, **52**: 101-108.

Swofford, D.L., PAUP: phylogenetic analysis using parsimony, version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois, 1993.

Tajima, K., Katoh, K., Hiraiwa, K.-I., Komoda, H. and Hayami, M., No anti-HTLV-I antibody positives among Polynesians in the southern Cook Islands. *Man Cult. Oceania* 1991, **7**: 63-69.

Tanaka, Y., Ishii, K., Sawada, T., Ohtsuki, Y., Hoshino, H., Yanagihara, R. and Miyoshi, I., Prophylaxis against a Melanesian variant of human T lymphotropic virus type I (HTLV-I) in rabbits using HTLV-I immune globulin

from asymptomatically infected Japanese carriers. *Blood* 1993, **82**: 3664-3667.

Trent, D.W., Grant, J.A., Vorndam, A.V. and Monath, T.P., Genetic heterogeneity among Saint Louis encephalitis virus isolates of different geographic origin. *Virology* 1981, **114**: 319-332.

Tsujimoto, H., Komuro, A., Iijima, K., Miyamoto, J., Ishikawa, K. and Hayami, M., Isolation of simian retroviruses closely related to human T-cell leukemia virus by establishment of lymphoid cell lines from various non-human primates. *Int. J. Cancer* 1985, **35**: 377-384.

Tsujimoto, A., Teruuchi, T., Imamura, J., Shimotohno, K., Miyoshi, I. and Miwa, M., Nucleotide sequence analysis of a provirus derived from HTLV-I-associated myelopathy (HAM). *Mol. Biol. Med.* 1988, **5**: 29-42.

Undevia, J.V., Blake, N.M., Kirk, R.L. and McDermid, E.M., The distribution of some enzyme group systems among Parsis and Iranis in Bombay. *Hum. Heredity* 1972, **22**: 274-282.

Ureta Vidal, A., Gessain, A., Yoshida, M., Mahieux, R., Nishioka, K., Tekaia, F., Rosen, L. and de Thé, G., Molecular epidemiology of HTLV-I in Japan: evidence for two distinct ancestral lineages with a particular geographical distribution. *AIDS Res. hum. Retrovir.* 1994, **10**: 1557-1566.

Wallace, A.R., On the zoological geography of the Malay Archipelago. *J. Linn. Soc.* London 1860, **4**: 172-184.

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

Wickler, S. and Spriggs, M., Pleistocene human occupation of the Solomon Islands, Melanesia. *Antiquity* 1988, **62**: 703-706.

Wong-Staal, F. and Gallo, R.C., Human T lymphotropic retroviruses. *Nature* 1985, **317**: 395-403.

Yanagihara, R., Human T lymphotropic virus type I infection and disease in the Pacific Basin. *Hum. Biol.* 1992, **64**: 843-854.

Yanagihara, R., Geographic-specific genotypes or topotypes of human T-cell lymphotropic virus type I as markers for early and recent migrations of human populations. *Adv. Vir. Res.* 1994, **43**: 147-186.

Yanagihara, R., Garruto, R.M., Miller, M.A., Leon-Monzon, M.E., Liberski, P.P., Gajdusek, D.C., Jenkins, C.L., Sanders, R.C. and Alpers, M.P., Isolation of HTLV-I from members of a remote tribe in New Guinea. *N. Engl. J. Med.* 1990a, **323**: 993-994.

Yanagihara, R., Jenkins, C.L., Alexander, S.S., Mora, C.A. and Garruto, R.M., Human T lymphotropic virus type I infection in Papua New Guinea: high prevalence among the Hagahai confirmed by Western analysis. *J. infect. Dis.* 1990b, **162**: 649-654.

Yanagihara, R., Jenkins, C.L., Ajdukiewicz, A.B. and Lal, R.B., Serological discrimination of HTLV I and II infection in Melanesia. *Lancet* 1991a, **337**: 617-618.

Yanagihara, R., Nerurkar, V.R. and Ajdukiewicz, A.B., Comparison between strains of human T lymphotropic virus type I isolated from inhabitants of the Solomon Islands and Papua New Guinea. *J. infect. Dis.* 1991b, **164**: 443-449.

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., Characterization of a variant of human T lymphotropic virus type I isolated from a member of a remote, recently contacted group in Papua New Guinea. *Proc. natn. Acad. Sci.* USA 1991c, **88**: 1446-1450.