VIROLOGY 196, 506–513 (1993)

Interfamilial and Intrafamilial Genomic Diversity and Molecular Phylogeny of Human T-Cell Lymphotropic Virus Type I from Papua New Guinea and the Solomon Islands[1]

VIVEK R. NERURKAR,* KI-JOON SONG,* NARUYA SAITOU,†
REBECCA R. MELLAND,* AND RICHARD YANAGIHARA*,[2]

*Laboratory of Central Nervous System Studies, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892; and †Laboratory of Evolutionary Genetics, National Institute of Genetics, Mishima 411, Japan

Received March 10, 1993; accepted June 1, 1993

To determine the interstrain genomic diversity and molecular phylogeny of the recently identified variants of human T-cell lymphotropic virus type I (HTLV-I) in Melanesia, we enzymatically amplified, then directly sequenced representative regions of the *gag*, *pol*, and *env* genes of HTLV-I strains from 10 members of four families, including one family from Papua New Guinea and three families from the Solomon Islands. When aligned and compared to a Japanese strain of HTLV-I (ATK), the Melanesian HTLV-I strains differed by 7.6 to 8.7% in the *gag*, 7.1 to 9.3% in the *pol*, and 7.3 to 8.2% in the *env* gene regions. Based on 931 nucleotides, the overall sequence divergence of the 10 Melanesian HTLV-I strains from HTLV-I ATK was 7.3 to 8.1% (68 to 75 base substitutions). The intrafamilial genetic heterogeneity among these virus strains was nil to 0.2%, while the interfamilial sequence variation between HTLV-I strains from the Solomon Islands and those from Papua New Guinea was 3.4 to 4.2%, and the genetic heterogeneity among virus strains from the three Solomon Islands families was 0.2 to 0.9%. Using the maximum parsimony and neighbor-joining methods, phylogenetic analysis indicated that the HTLV-I strains from Papua New Guinea and the Solomon Islands formed a monophyletic group and that the Melanesian and cosmopolitan strains of HTLV-I have evolved along two major geographically dependent lineages.

INTRODUCTION

Sexual intercourse and breast-feeding are the principal modes of horizontal and vertical transmission for the human retroviruses, human T-cell lymphotropic virus type I (HTLV-I), and human immunodeficiency virus type 1 (HIV-1). Sequence comparisons between strains of HIV-1 from mother–infant pairs and from sexual partners indicate a relatively high degree of intrafamilial genetic dissimilarity, in the order of 2 to 8%, compared to 6% to more than 20% among viral isolates from unrelated individuals (Burger *et al.*, 1991; Wike *et al.*, 1992; Wolinsky *et al.*, 1992). For HTLV-I, seroepidemiological data indicate familial clustering of infection (Tajima *et al.*, 1982; Kajiyama *et al.*, 1986; Yanagihara *et al.*, 1990), but sequence information on virus strains within families or between close contacts is almost nonexistent. Thus, to better understand the degree of genomic variability of HTLV-I strains within and between families, following presumed horizontal (sexual) or vertical (mother-to-child) transmission, we amplified by polymerase chain reaction (PCR) then directly sequenced selected regions of the *gag*, *pol*, and *env* genes of HTLV-I strains from 10 members of four Melanesian families. The considerable intrafamilial sequence similarity among these virus strains from Papua New Guinea and the Solomon Islands resembles that from families in equatorial Zaire (Goubau *et al.*, 1992), southern India (Nerurkar *et al.*, 1993), and from a French patient with transfusion-acquired HTLV-I myelopathy, his wife, and the blood donor (Gessain *et al.*, 1992).

MATERIALS AND METHODS

Families and virus strains

HTLV-I strains from 10 asymptomatically infected members of four Melanesian families (Table 1), including one family from Papua New Guinea and three families from the Solomon Islands, and HTLV-I strains BEL 1 and BEL 2, respectively, from two unrelated life-long residents (both women, 50- and 60-years old) of the Polynesian Outlier Bellona within the Solomon archipelago, were studied. T-cell lines harboring HTLV-I were available from three Melanesians (strains MEL 1, MEL 3, and MEL 5) and from one Polynesian (strain BEL 1) (Yanagihara *et al.*, 1991). Informed consent was obtained from all study participants.

[1] This work was presented in part at the Fifth International Conference on Retrovirology in Kumamoto, Japan, May 1992, and at the NIH Research Festival in Bethesda, Maryland, September 1992.

[2] To whom reprint requests should be addressed at the National Institutes of Health, Bldg. 36, Rm. 5B-21, Bethesda, MD 20892.

0042-6822/93 $5.00

The New Guinean family, consisting of a mother and her two sons, was part of a 260-member, recently contacted hunter–horticulturalist group (Hagahai) living along the northern banks of the Yuat River Gorge, at altitudes of 200 to 1800 m, in the fringe highlands of Madang province (Jenkins *et al.*, 1989). While the Hagahai have certain cultural practices akin to those of lowland and highland New Guinean groups, analysis of HLA-DR antigens suggests that they are distinct from both coastal and highland groups and may represent descendants of migrants who arrived more than 5000 years ago (Bhatia *et al.*, 1993).

The three unrelated Melanesian families from the Solomon Islands resided in different villages or towns (Honiara, Kolosulu, Marau) on Guadalcanal. Members of two families were born elsewhere in the Solomon Islands (two on Morovo Lagoon and one on Rendova in New Georgia, approximately 250 km west of Guadalcanal). None of these seven individuals had traveled outside the Solomon Islands and four had never left Guadalcanal.

PCR and direct sequencing of amplified products

High molecular weight DNA, extracted using a nonorganic method (Oncor, Gaithersburg, MD) from fresh, uncultured peripheral blood mononuclear cells (PBMC), from PBMC maintained in culture for 4 weeks, and from T-cell lines (Table 1), and enzymatically amplified products were handled by different personnel in separate rooms. Oligonucleotide primers for 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 *gag* (bases 1423–1444, 5′-CCATCACCAGCAGCTAGATAGC-3′ and bases 1560–1537, 5′-AGTTGCTGGTATTCTCGCCTTAAT-3′), the 3′-end of *pol* (bases 4757–4778, 5′-CCCTACAATCCAACCAGCTCAG-3′ and bases 4942–4919, 5′-GTGGTGAAGCTGCCATCGGGTTTT-3′), the 5′-end of the gp46-encoding region of *env* (bases 5228–5246, 5′-TTTATTCTTCCAGTTCTGC-3′ and bases 5596–5572, 5′-TAGGGGCTGGAGACGGCTCCTGTAT-3′) and the gp21-encoding region of *env* (bases 6068–6087, 5′-TCATAACTCCCTCATCCTGC-3′ and bases 6481–6462, 5′-CAGCCAGTCAGGACTCGATT-3′). Primers were used at a final concentration of 1 μM in a reaction mixture of 100 μl containing 2 μg of DNA and 2.5 units 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. Reaction mixtures were cycled as described previously (Nerurkar *et al.*, 1992).

Enzymatically amplified DNA, purified and concentrated using Centricon 100 microconcentrators (Amicon Division, Danvers, MA), was sequenced manually, as well as with an automated sequencer (Model 373A, Applied Biosystems, Inc., Foster City, CA), in both directions, by the dideoxy chain-termination method. Direct DNA sequencing, which obviates the need for cloning, was employed since it facilitates the rapid identification of the predominant virus genome within an infected individual and because this methodology has been shown to be well suited for molecular analysis of HIV-1 between close contacts (Burger *et al.*, 1991).

Sequence and phylogenetic analyses

Nucleotide sequences were aligned and compared with sequences of HTLV-I strains from various geographic regions, including Japan (ATK, H5, TSP-1, MT-2) (Seiki *et al.*, 1983; Tsujimoto *et al.*, 1988; Evangelista *et al.*, 1990; Gray *et al.*, 1990), the Caribbean basin (HS-35, CH) (Malik *et al.*, 1988; Paine *et al.*, 1991; Ratner *et al.*, 1991), Chile (ST) (Dekaban *et al.*, 1992) and Zaire (EL) (Paine *et al.*, 1991; Ratner *et al.*, 1991), as well as with sequences of simian T-cell lymphotropic virus type I (STLV-I) strain PtM3 from a pig-tailed macaque (*Macaca nemestrina*) from Indonesia (Watanabe *et al.*, 1985) and sequences of HTLV-II strain Mo from the United States (Shimotohno *et al.*, 1985).

The nucleotide- and amino acid-sequence alignments were facilitated by using the GAP program available on the VAX system, as part of the Genetics Computer Group, which employs the algorithm of Needleman and Wunsch (1970). To maximize our confidence in the derived topologies, we employed both the neighbor-joining (Saitou and Nei, 1987) and maximum parsimony methods (Fitch, 1977), which accommodate variable rates of genetic change, to construct unrooted phylogenetic trees. Evolutionary distances (number of nucleotide substitutions) between the sequences were estimated based on the one-parameter (Jukes and Cantor, 1969) and two-parameter methods (Kimura, 1980), and these distances were used for constructing trees by the neighbor-joining method. PAUP version 3.0 (Swofford, 1990) was used for maximum parsimony analysis and for estimating boot-strap probabilities.

Estimated branch lengths (number of substitutions per nucleotide site), obtained by using the neighbor-joining method, were multiplied by the number of nucleotides being analyzed, and the resulting values were rounded. If an internal branch was assigned a distance of zero, the branch was truncated. When the estimated number of substitutions at any branch was smaller than the minimum possible number of substitutions required by the maximum parsimony method, the minimum number was used.

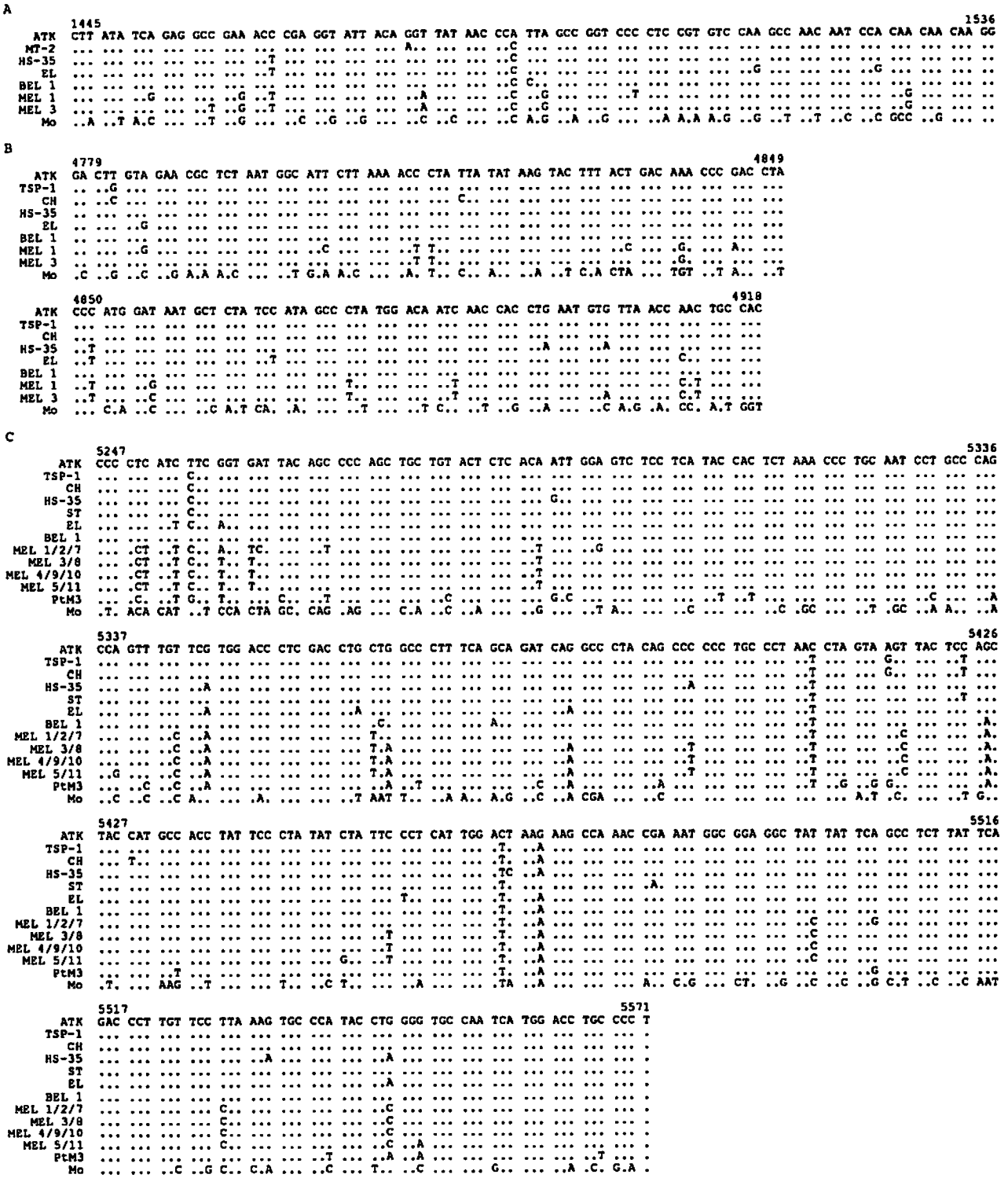

FIG. 1. Alignment and comparison of nucleotide sequences of (A) the p24-encoding region of the *gag* gene (bases 1445 to 1536), (B) the 3'-end of the *pol* gene (bases 4779 to 4918), (C) the 5'-end of the gp46-encoding region of the *env* gene (bases 5247 to 5571), and (D) the gp21-encoding region of the *env* gene (bases 6088 to 6461) in HTLV-I strains from 10 members of four Melanesian families. Family groupings with relationships are listed in Table 1. Sequence uniformity in the (A) *gag* and (B) *pol* regions among viruses from Papua New Guinea and the Solomon Islands is indicated by strains MEL 1 and MEL 3, respectively. For comparison, corresponding sequences are shown for HTLV-I strains from Japan (ATK, MT-2, TSP-1), the Caribbean basin (CH, HS-35), Chile (ST), Zaire (EL), the Polynesian Outlier Bellona (BEL 1, BEL 2), and for STLV-I strain PtM3 from Indonesia, as well as for HTLV-II strain Mo from the United States.

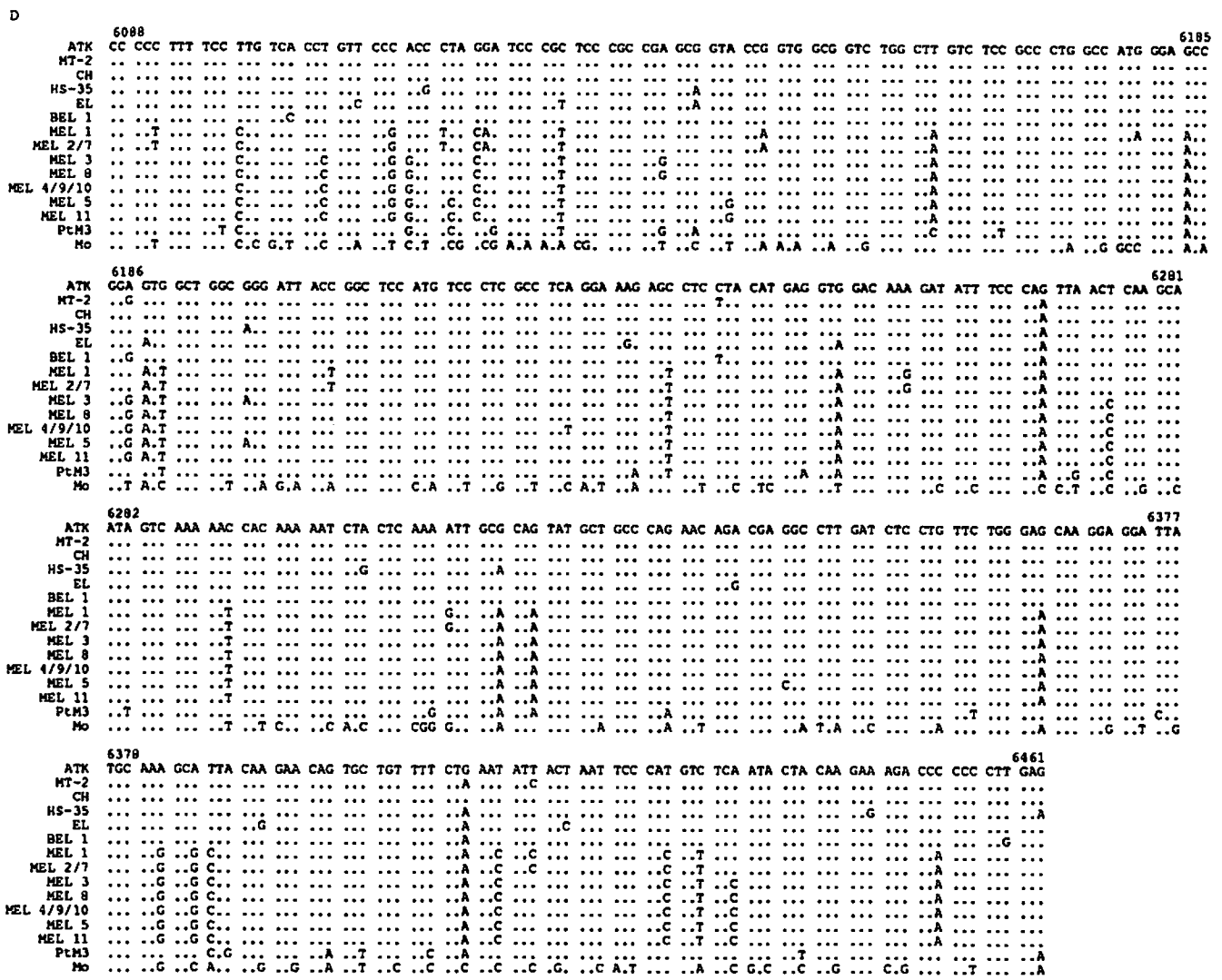

FIG. 1—*Continued*

RESULTS

Nucleotide sequence analysis of the *gag*, *pol*, and *env* genes

As verification that the sequences of the Melanesian HTLV-I strains from DNA extracted from cultured PBMC or T-cell lines did not result from selection of aberrant virus populations during the 4-week culture period or the long-term maintenance of T-cell lines in culture, identical proviral sequences were obtained from DNA extracted from fresh, uncultured PBMC. Thus, the viral sequences presented here represent those of the predominant virus population *in vivo* and do not reflect selection artifacts from *in vitro* manipulation.

Nucleotide sequences for each of the Melanesian HTLV-I strains, arranged by family group, were aligned and compared to representative strains of HTLV-I from other geographical regions. Multiple, unique base substitutions were found in the *gag*, *pol*, and *env* genes (Figs. 1A to 1D). Compared to the Japanese HTLV-I strain ATK, the Melanesian HTLV-I strains diverged by 7 to 8 of 92 bases (7.6 to 8.7%) in the *gag*, 10 to 13 of 140 bases (7.1 to 9.3%) in the *pol*, and 51 to 57 of 699 bases (7.3 to 8.2%) in the *env* gene regions (Table 1). The overall sequence similarity between the Melanesian HTLV-I strains and HTLV-I ATK was 91.9 to 92.7% (68 to 75 substitutions in 931 nucleotides), with an intrafamilial sequence heterogeneity among the Melanesian HTLV-I strains of 0 to 0.2%. Between HTLV-I strains from the Solomon Islands and those from Papua New Guinea, the interfamilial genetic variation was 3.4 to 4.2% (32- to 39-nucleotide differences), while it was 0.2 to 0.9% (2- to 8-nucleotide differences, all in the *env* gene) among virus strains from the three Solomon Islands families.

The tendency for some of the base substitutions to occur at identical positions in cosmopolitan and Melanesian HTLV-I strains suggested a nonrandom nature to the point mutations (Figs. 1A to 1D), but no differ-

TABLE 1

MELANESIAN HTLV-I STRAINS AND DNA SOURCE FOR GENE AMPLIFICATION AND SEQUENCE COMPARISON WITH COSMOPOLITAN HTLV-I (ATK)

Country of origin	Family	Relation	Age/ sex	Virus name	DNA source for PCR			Percent divergence from HTLV-I ATK		
					Uncultured PBMC	Cultured PBMC	T-cell line	gag	pol	env
Papua New Guinea	I	Son	21M	HTLV-I MEL 1	+		+	8.7	9.3	7.7
		Mother	60F	HTLV-I MEL 2	+	+		8.7	9.3	7.6
		Son	31M	HTLV-I MEL 7	+	+		8.7	9.3	7.6
Solomon Islands	II	Wife	39F	HTLV-I MEL 3		+	+	7.6	7.1	7.4
		Husband	49M	HTLV-I MEL 8		+		7.6	7.1	7.3
	III	Mother	60F	HTLV-I MEL 4	+	+		7.6	7.1	7.3
		Father	75M	HTLV-I MEL 9		+		7.6	7.1	7.3
		Daughter	13F	HTLV-I MEL 10		+		7.6	7.1	7.3
	IV	Husband	58M	HTLV-I MEL 5	+	+	+	7.6	7.1	8.2
		Wife	42F	HTLV-I MEL 11		+		7.6	7.1	7.9

ence in genetic variability was noted among the predominant virus populations presumably acquired by horizontal (husband-to-wife) or vertical (mother-to-child) transmission. Nucleotide changes for each of the virus strains consisted primarily of transitions (71 to 90%), with the majority being deoxycytosine to deoxythymidine substitutions or vice versa. Furthermore, no insertions, deletions or frame shifts were detected in the regions sequenced.

Despite the geographical proximity of Bellona to islands endemic for the Melanesian HTLV-I variants, HTLV-I strains BEL 1 and BEL 2 (which were genetically identical) were much more similar to cosmopolitan than to Melanesian strains of HTLV-I, differing by only 13 (1.4%) of 931 nucleotides from the Japanese HTLV-I strain ATK. Complete sequence identity in the 140-bp region of the *pol* gene between the Japanese HTLV-I strain ATK and the virus strains from Bellona (Fig. 1B) was, therefore, consistent with the high degree of sequence similarity elsewhere in the viral genome.

Deduced amino acid sequence analysis of the *gag*, *pol*, and *env* genes

All of the base substitutions in the 31-amino acid region of the p24 capsid-encoding *gag* gene of the Melanesian HTLV-I strains were silent. On the other hand, in the 46-amino acid region 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 ATK. Like the HTLV-I variant EL from Zaire, but unlike other cosmopolitan strains of HTLV-I from Japan, the Caribbean basin and the Americas, the Melanesian HTLV-I strains had a codon-altering substitution at position 4910, which resulted 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 6 to 7 (5.6 to 6.5%) and 4 to 7 (3.2 to 5.6%) amino acids, respectively, from HTLV-I ATK. Despite this overall degree of amino acid sequence dissimilarity, the amino-terminal neutralizing domain (comprised of amino acids 88 to 98, TrpIleLysLysProAsnArgAsnGlyGlyGly), encoded by bases 5463 to 5495, on the external envelope glycoprotein gp46 was totally conserved between Melanesian and cosmopolitan HTLV-I strains (as well as in the STLV-I strain PtM3), except for HTLV-I ATK, in which the corresponding sequence was TrpThrLysLysProAsnArgAsnGlyGlyGly. As further 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, were identical between Melanesian and cosmopolitan HTLV-I strains.

A region of variability at the 5'-end of the *env* gene, in the region of the signal peptide, resulted in a deduced sequence for amino acids 16 to 21 of ProProIleLeuSerSer and ProProIleLeuCysTyr for the HTLV-I strains from Papua New Guinea and the Solomon Islands, respectively. 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, ST, SP) and ProLeuIleLeuSerAsp (for HTLV-I strain EL). To what extent these changes are functionally important is unclear.

Phylogenetic analysis of the *gag*, *pol*, and *env* gene sequences

Unrooted phylogenetic trees constructed from the *gag* (92 bp) and *pol* (140 bp) and from the gp46 *env*

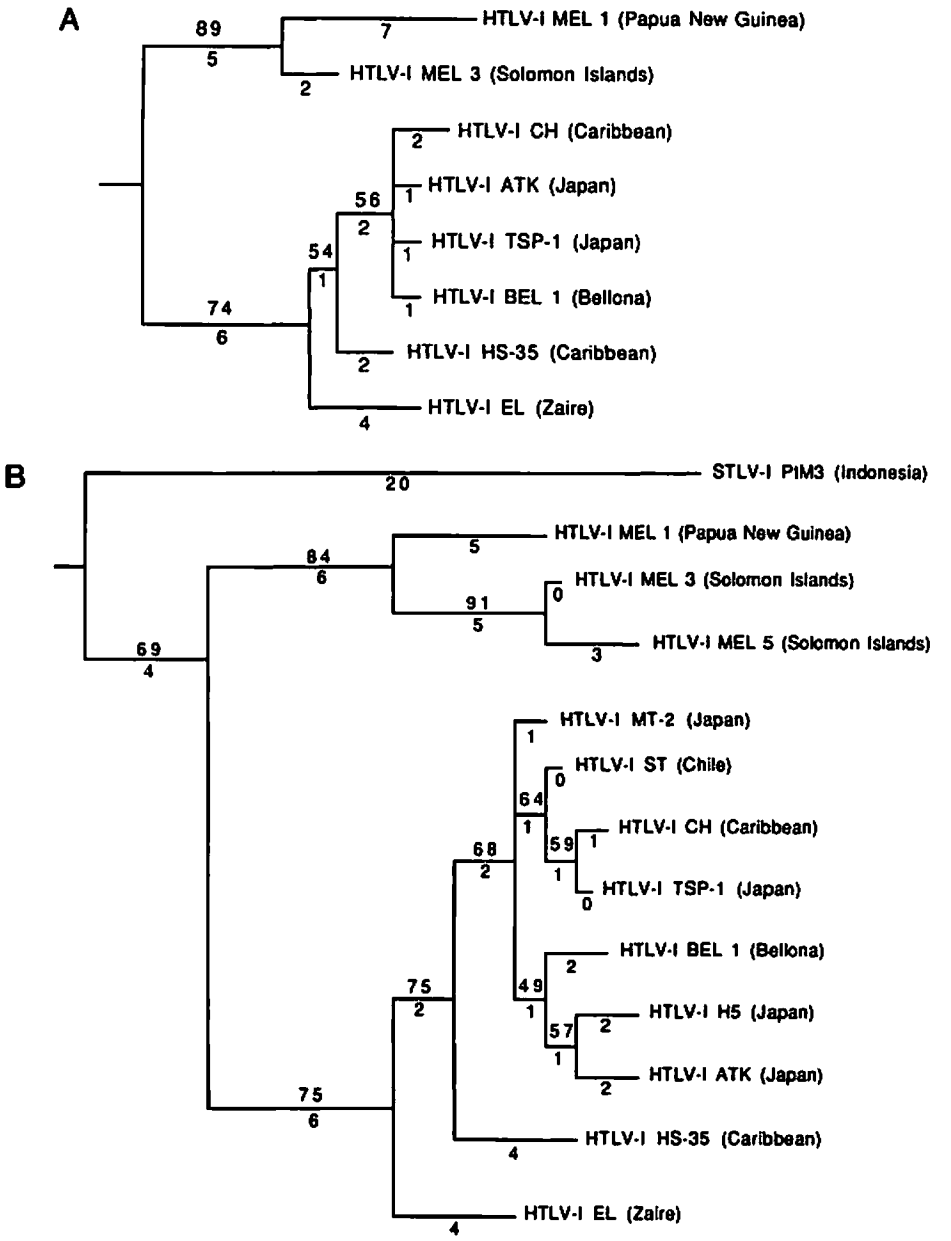

FIG. 2. Phylogenetic trees based on (A) the 92-bp region of the p24-encoding *gag* gene and 140-bp region of the *pol* gene and on (B) the 325-bp region of the gp46-encoding *env* gene of Melanesian and cosmopolitan strains of HTLV-I, and of STLV-I strain PtM3 from a pig-tailed macaque, constructed by the neighbor-joining and maximum parsimony methods (see text for details). The trees were rooted by assuming HTLV-II as the outgroup. Branch lengths (number of nucleotide substitutions) are given below each branch, and the boot-strap probabilities (in percentage), calculated from 1000 resamplings by the maximum parsimony method, are given above the internal branches. The placements of HTLV-I MEL 1 and of HTLV-I MEL 3 and HTLV-I MEL 5 on these trees are representative of the HTLV-I strains from Papua New Guinea and the Solomon Islands, respectively.

(325 bp) gene sequences of the Melanesian HTLV-I strains, and from corresponding sequences of virus strains from other geographical areas, had nearly identical branching patterns and branch lengths using either the maximum parsimony or neighbor-joining methods. The two equally parsimonious trees constructed from the *gag* and *pol* gene sequences required 115 nucleotide substitutions, while the single maximum parsimony tree based on the *env* gene sequences required 168 nucleotide substitutions. Representative trees, which were rooted by assuming HTLV-II as the outgroup, were generated by combining the results from the maximum parsimony and neighbor-joining methods (Figs. 2A and 2B). Branch lengths were identical with those of the maximum parsimony tree when the ACCTRAN option of PAUP was used.

Boot-strap probabilities (in percentage), as determined by PAUP for 1,000 resamplings, for the branches clustering Melanesian strains and cosmopolitan strains, respectively, in both composite trees were

relatively high (≥74%), demonstrating that the Melanesian and cosmopolitan strains of HTLV-I have evolved along two major geographically dependent lineages. Congruency of the phylogenetic trees (Figs. 2A and 2B) based on different gene regions indicated that the Melanesian strains of HTLV-I were monophyletic and occupied a unique position among all HTLV-I, with the Melanesian strains diverging from the common ancestor of HTLV-I prior to the divergence of cosmopolitan strains of HTLV-I.

DISCUSSION

Considering the presumed slow mutation rate of HTLV-I, a high degree of sequence similarity was not unexpected among HTLV-I strains from members of the same family. As demonstrated, the nucleotide sequences of selected regions of the *gag, pol,* and *env* genes of HTLV-I strains within each of the four Melanesian families were nearly or completely identical. These data support the notion that the virus within each family was introduced by one member and was transmitted to the others, either vertically (as from mother-to-child via virus-infected breast milk) or horizontally (as from husband-to-wife via virus-infected genital secretions). Although all three families from the Solomon Islands resided on Guadalcanal, it is possible that HTLV-I infection was acquired when some of these individuals lived on islands geographically distant to Guadalcanal. If true, this would indicate that genetically similar strains of HTLV-I exist among Melanesian groups throughout the Solomon Islands.

Our data also indicated a marked degree of sequence conservation of a linear neutralizing domain (defined by amino acids 88 to 98) on the amino-terminal external envelope glycoprotein of HTLV-I strains, regardless of geographic origin. The threonine to isoleucine difference at position 89 between HTLV-I ATK and all other HTLV-I strains, including those from Melanesia (while probably the result of sequencing error), appears to be functionally irrelevant (Palker *et al.*, 1992). Antigenic determinants for neutralization on the mid-portion of the outer envelope glycoprotein are also highly conserved among HTLV-I strains, including the Melanesian variants (Melland, 1992). Not unexpectedly then, the neutralizing epitopes of cosmopolitan and Melanesian HTLV-I strains are functionally indistinguishable as determined by cross-neutralization assays using vesicular stomatitis virus pseudotypes bearing envelope antigens of HTLV-I, indicating that HTLV-I exists as a single serotype worldwide (Hoshino *et al.*, 1993).

Recent sustained contact between the Hagahai and outsiders, beginning in late 1983, has resulted in changing patterns of communicable diseases and in the acquisition of new infectious agents, such as mumps and hepatitis B viruses (Jenkins *et al.*, 1989). However, it is unlikely that the Hagahai are a recent "virgin-soil population" for HTLV-I because of the ages of the infected individuals and the nonrandom distribution of infection (Yanagihara *et al.*, 1990). We suspect that HTLV-I variants genetically similar to the virus strains from the Hagahai occur elsewhere in Papua New Guinea and West New Guinea, and possibly elsewhere in Indonesia. Unfortunately, since there are no other HTLV-I isolates from Papua New Guinea to date, it is unclear if the HTLV-I variants from the Hagahai are truly representative of virus strains circulating among other fringe highland populations. Nevertheless, the high degree of sequence similarity among HTLV-I strains from Melanesian Solomon Islanders, as demonstrated in this study, would predict that HTLV-I strains harbored by other virus-infected fringe highland New Guinean populations are genetically similar to those found among the Hagahai. Whether or not 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, awaits further study.

Although it is unclear when HTLV-I was introduced into these 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 genetically related but distinct from each other, suggesting evolution from ancestral strains of virus introduced by one of several founder populations. 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, suggests a different source and subsequent 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.

ACKNOWLEDGMENTS

We thank Ms. Tracy C. DeLozier for technical assistance, Dr. Mark S. Godec for synthesizing oligonucleotide primers, and Mr. Gary W. Smythers of the Frederick Biomedical Supercomputing Center at the National Cancer Institute-Frederick Cancer Research and Development Center in Frederick, Maryland, for assistance with the VAX.

REFERENCES

BHATIA, K., DAVIES, R., JENKINS, C., JAZWINSKA, E., KOKI, G., and SERJEANTSON, S. W. (1993). HLA-DR RFLP-typing in four Papua New Guinea populations. *Am. J. Phys. Anthropol.*, in press.

BURGER, H., WEISER, B., FLAHERTY, K., GULLA, J., NGUYEN, P.-N., and GIBBS, R. A. (1991). Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity among close contacts. *Proc. Natl. Acad. Sci. USA* **88**, 11236–11240.

DEKABAN, G. A., KING, E. E., WATERS, D., and RICE, G. P. A. (1992).

Nucleotide sequence analysis of an HTLV-I isolate from a Chilean patient with HAM/TSP. *AIDS Res. Hum. Retroviruses* **8**, 1201–1207.

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). Nucleotide sequence analysis of a provirus derived from an individual with tropical spastic paraparesis. *Microb. Pathog.* **8**, 259–278.

FITCH, W. M. (1977). On the problem of discovering the most parsimonious tree. *Am. Natur.* **111**, 223–257.

GESSAIN, A., GALLO, R. C., and FRANCHINI, G. (1992). The low degree of HTLV-I genetic drift in vivo as a means to follow viral transmission and movement of ancient human populations. *J. Virol.* **66**, 2288–2295.

GOUBAU, P., KAZADI, K., CARTON, H., VANDAMME, A., LIU, H. F., and DESMYTER, J. (1992). The epidemiology of human T-cell lymphotropic viruses in Zaire. *Am. J. Trop. Med. Hyg.* **47**, 258. [Abstract]

GRAY, G. S., WHITE, M., BARTMAN, T., and MANN, D. (1990). Envelope gene sequence of HTLV-I isolate MT-2 and its comparison with other HTLV-I isolates. *Virology* **177**, 391–395.

HOSHINO, H., NAKAMURA, T., TANAKA, Y., MIYOSHI, I., and YANAGIHARA, R. (1993). 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.*, in press.

JENKINS, C., DIMITRAKAKIS, M., COOK, I., SANDERS, R., and STALLMAN, N. (1989). Culture change and epidemiological patterns among the Hagahai, Papua New Guinea. *Hum. Ecol.* **17**, 27–57.

JUKES, T. H., and CANTOR, C. R. (1969). Evolution of protein molecules. *In* "Mammalian Protein Metabolism" (H. N. Munro, Ed.), pp. 21–132. Academic Press, New York.

KAJIYAMA, W., KASHIWAGI, S., IKEMATSU, H., HAYASHI, J., NOMURA, H., and OKOCHI, K. (1986). Intrafamilial transmission of adult T cell leukemia virus. *J. Infect. Dis.* **154**, 851–857.

KIMURA, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120.

MALIK, K. T. A., EVEN, J., and KARPAS, A. (1988). 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.* **69**, 1695–1710.

MELLAND, R. R. (1992). 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, 70 pp.

NEEDLEMAN, S. B., and WUNSCH, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequences of two proteins. *J. Mol. Biol.* **48**, 443–453.

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). Failure to isolate human T lymphotropic virus type I and to detect variant-specific genomic sequences by polymerase chain reaction in Melanesians with indeterminate Western immunoblot. *J. Gen. Virol.* **73**, 1805–1810.

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). 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.*, in press.

PAINE, E., GARCIA, J., PHILPOTT, T. C., SHAW, G., and RATNER, L. (1991). Limited sequence variation in human T-lymphotropic virus type I isolates from North American and African patients. *Virology* **182**, 111–123.

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

RATNER, L., PHILPOTT, T., and TROWBRIDGE, D. B. (1991). Nucleotide sequence analysis of isolates of human T-lymphotropic virus type 1 of diverse geographic origins. *AIDS Res. Hum. Retroviruses* **7**, 923–941.

SAITOU, N., and NEI, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

SEIKI, M., HATTORI, S., HIRAYAMA, Y., and YOSHIDA, M. (1983). Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proc. Natl. Acad. Sci. USA* **80**, 3618–3622.

SHIMOTOHNO, K., TAKAHASHI, Y., SHIMIZU, N., GOJOBORI, T., GOLDE, D. W., CHEN, I. S. Y., MIWA, M., and SUGIMURA, T. (1985). Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: An open reading frame for the protease gene. *Proc. Natl. Acad. Sci. USA* **82**, 3101–3105.

SWOFFORD, D. L. (1990). PAUP: phylogenetic analysis using parsimony, version 3.0. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

TAJIMA, K., TOMANAGA, S., SUICHI, T., KAWAGOE, T., KOMODA, H., HINUMA, Y., ODA, T., and FUJITA, K. (1982). Epidemiological analysis of the distribution of antibody to adult T-cell leukemia virus-associated antigen (ATLA): Possible horizontal transmission of adult T-cell leukemia virus. *Jpn. J. Cancer Res.* **73**, 893–901.

TSUJIMOTO, A., TERUUCHI, T., IMAMURA, J., SHIMOTOHNO, K., MIYOSHI, I., and MIWA, M. (1988). Nucleotide sequence analysis of a provirus derived from HTLV-I-associated myelopathy (HAM). *Mol. Biol. Med.* **5**, 29–42.

WATANABE, T., SEIKI, M., TSUJIMOTO, H., MIYOSHI, I., HAYAMI, M., and YOSHIDA, M. (1985). Sequence homology of the simian retrovirus genome with human T-cell leukemia virus type I. *Virology* **144**, 59–65.

WIKE, C. M., KORBER, B. T. M., DANIELS, M. R., HUTTO, C., MUÑOZ, M., FURTADO, M., PARKS, W., SAAH, A., BULTERYS, M., KURAWIGE, J.-B., and WOLINSKY, S. M. (1992). HIV-1 sequence variation between isolates from mother-infant transmission pairs. *AIDS Res. Hum. Retroviruses* **8**, 1297–1300.

WOLINSKY, S. M., WIKE, C. M., KORBER, B. T. M., HUTTO, C., PARKS, W. P., ROSENBLUM, L. L., KUNSTMAN, K. J., FURTADO, M. R., and MUÑOZ, J. L. (1992). Selective transmission of human immunodeficiency virus type 1 from mothers to infants. *Science* **255**, 1134–1137.

YANAGIHARA, R., JENKINS, C. L., ALEXANDER, S. S., MORA, C. A., and GARRUTO, R. M. (1990). Human T lymphotropic virus type I infection in Papua New Guinea: High prevalence among the Hagahai confirmed by Western analysis. *J. Infect. Dis.* **162**, 649–654.

YANAGIHARA, R., NERURKAR, V. R., and AJDUKIEWICZ, A. B. (1991). Comparison between strains of human T lymphotropic virus type I isolated from inhabitants of the Solomon Islands and Papua New Guinea. *J. Infect. Dis.* **164**, 443–449.