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

MICHAEL P. SHERMAN,¹ NITIN K. SAKSENA,¹ DIPAK K. DUBE,¹ RICHARD YANAGIHARA,²
AND BERNARD J. POIESZ^{1*}

Departments of Medicine and Microbiology/Immunology, State University of New York Health Science Center at Syracuse, Syracuse, New York 13210,¹ and Laboratory of Central Nervous System Studies, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland 20892²

Received 26 July 1991/Accepted 2 January 1992

Recent studies have established the presence of human T-cell lymphoma/leukemia virus type I (HTLV-I) in Melanesia. An HTLV-I strain, PNG-1, has now been isolated from a healthy member of the Hagahai, a remote, recently contacted group in Papua New Guinea. To further characterize PNG-1, we employed polymerase chain reaction amplification with subsequent cloning and sequencing of amplified products. Sequence analyses of amplified regions of *pol*, *env*, and *pX* genes of this variant indicate marked heterogeneity (approximately 7%) from prototype HTLV-I. Based on available sequence data, PNG-1 is distinct from all other known HTLV-I strains and diverges from the common ancestor of HTLV-I prior to prototype isolates. The data also suggest, therefore, that HTLV-I originated in the Indo-Malay region rather than Africa.

Human T-cell lymphoma/leukemia virus type I (HTLV-I) is a type C human retrovirus which belongs to a genus that also includes the other primate T-cell lymphoma/leukemia viruses (PTLV) HTLV-II and simian T-cell lymphoma/leukemia virus type I (STLV-I) (46). HTLV-I is the etiologic agent of adult T-cell lymphoma/leukemia (17, 33) and tropical spastic paraparesis/HTLV-I-associated myelopathy (15, 30, 35). HTLV-I is endemic in the Caribbean basin (4–6), southern Japan (23), central and western Africa (10, 12, 41), and the Seychelles Islands (36). It has been hypothesized that HTLV-I originated in Africa (13, 48, 49), but the worldwide distribution of HTLV-I remains undefined. Recently, reports of seroreactivity to HTLV-I in parts of Melanesia (2, 14, 52) have been confirmed by virus isolation (51, 53, 54). At this time, however, the existence of endemic HTLV-I infection specific to Melanesia is still questioned (3, 50).

We now report on the sequence analysis of a newly defined HTLV-I isolate, PNG-1 (53, 54), established from a member of the Hagahai who lived isolated until 1983 along the Yuat River Gorge in the fringe highlands of Papua New Guinea. These data unequivocally demonstrate that PNG-1 is a distinct variant of HTLV-I and address the question of the possible origin of HTLV-I.

The PNG-1 isolate has previously been characterized as being a member of the HTLV-I subgroup (53, 54). This isolate was further characterized by employing the polymerase chain reaction (PCR) (37, 38) with subsequent analysis by liquid hybridization of the appropriate ³²P-end-labeled oligonucleotide probe with amplified DNA as previously described (1). Comparative PCR analyses of PNG-1 with three HTLV-I isolates and an HTLV-II isolate obtained from widely separated geographic locales indicate that, while PNG-1 belongs to the HTLV-I subgroup, it varies considerably from HTLV-I MT-2 (28), HUT 102B2 (33), and HSC-CTCL-11B (11) (Table 1).

We analyzed PNG-1 by sequencing amplified DNA fragments from four regions of the HTLV-I genome (Fig. 1). The linker sequence ACAGGTACCTGCAGATCTAGA (5' to 3'), containing a restriction site for *Kpn*I, was synthesized on the 5' end of the positive-strand primers, while the linker sequence TACGAGCTCGCGAATTCATGA (5' to 3'), possessing an *Sst*I restriction site, was added to the 5' end of the negative-strand primers. Amplified DNAs and M13mp18 (26) vector DNA were digested with both *Kpn*I and *Sst*I and then ligated together with T4 ligase. After transfection into *Escherichia coli* (DH5αF'), plaques were screened by hybridization with end-labeled probes, and the DNA from each plaque was sequenced by the dideoxy method (39).

First, open reading frame II within the *pX* regions (ORF *pX*-II) of HTLV-I PNG-1 and HSC-CTCL-11B was sequenced, and the corresponding regions in HTLV-II MoT (20, 45) and two other HTLV-II isolates (RW and FF) were sequenced (Fig. 1A). Published sequences for a prototype Japanese HTLV-I isolate (ATK) (42) and a southeast Asian STLV-I isolate (PtM3) (48) are also included for comparison. While the complete proviral DNA sequences of prototype HTLV-I (42) and HTLV-II (45) vary by approximately 38%, these two isolates differ by only 15% in ORF *pX*-II. Within the HTLV-I and HTLV-II subgroups, there was minimal sequence variation and the PNG-1 sequence was almost identical to that of prototype HTLV-I. Others have also found strong sequence conservation in ORF *pX*-II, presumably because it encodes a portion of two overlapping functional proteins, Tax and Rex, which are critical for PTLV expression and replication (19, 44). Tax and Rex are translated from distinct mRNAs which both contain ORF *pX*-II message but in different reading frames (43). Hence, there are strong constraints on sequence variation in ORF *pX*-II DNA. Two variants within the HSC-CTCL-11B isolate were identified on the basis of sequence information, which was consistent with previous data indicating two proviral integrants in this cell line (11).

Prototypes HTLV-I and -II differ by 29% in the portion of the transmembrane *env* gene (gp21) analyzed, while PNG-1

* Corresponding author.

TABLE 1. PCR results when various HTLV-I/II primers were used with different HTLV-I or HTLV-II isolates^a

Primers	Probe	Region	Band appears for:				
			HTLV-I				HTLV-II MoT
			PNG-1	HSC-CTCL-11B	HUT 102B2	MT-2	
HTIL(490-515)+/HTIL(655-630)-	HTIL(564-596)+d	LTR	—	+	+	+	—
HTIG(863-886)+/HTIG(1397-1375)-	HTIG(981-1023)+d	<i>gag</i>	+ ^b	+ ^b	+ ^b	+ ^b	—
HTIG(1215-1235)+/HTIG(1393-1370)-	HTIG(1238-1277)+d	<i>gag</i>	—	+	+	+	—
HTIG(1388-1411)+/HTIG(1660-1641)-	HTIG(1451-1412)+d	<i>gag</i>	+ ^b	+	+	+	—
HTIG(1423-1444)+/HTIG(1560-1537)-	HTIG(1475-1507)+d	<i>gag</i>	+	+	+	+	—
HTIP(2801-2820)+/HTIG(3037-3018)-	HTIP(2821-2860)-d	<i>pol</i>	+ ^b	+	+ ^b	+ ^b	—
HTIP(3365-3384)+/HTIP(3483-3465)-	HTIP(3460-3426)-d	<i>pol</i>	—	+	+	+	—
HTIP(4757-4778)+/HTIP(4942-4919)-	HTIP(4870-4902)+d and	<i>pol</i>	+	+	+	+	+
HTIIP(4735-4756)+/HTIIP(4920-4897)-	HTIIP(4848-4880)+d						
	HTIP(4825-4850)+d	<i>pol</i>	+	+	+	+	—
	HTIIP(4880-4898)+d	<i>pol</i>	—	—	—	—	+
HTIE(5228-5247)+/HTIE(5596-5572)-	HTIE(5305-5271)+d	<i>env</i>	+	+	+ ^b	—	—
HTIE(5270-5292)+/HTIE(5540-5521)-	HTIE(5301-5340)+d	<i>env</i>	+ ^b	+	+ ^b	+ ^b	—
HTIE(6293-6324)+/HTIE(6527-6498)-	HTIE(6330-6368)+d	<i>env</i>	+	+	+ ^b	+ ^b	—
HTIPX(7248-7267)+/HTIPX(7406-7386)- and HTIIPX(7358-7377)+	HTIPX(7337-7476)+d and HTIIPX(7447-7468)+d	<i>pX</i>	+	+	+	+	+
/HTIPX(7516-7596)-							

^a PCR amplification and subsequent liquid hybridization using primer pairs and detectors to different regions of the HTLV genome were performed on DNA from four HTLV-I- and one HTLV-II-infected cell lines, as previously described (1). Oligonucleotides are named by a two-letter initial for HTLV (HT) followed by the number of the designated virus (I or II), an initial for the gene or region of the indicated virus with the numbered position in the genome (EMBL system for HTLV-I [42] and Shimotohno et al. system [45] for HTLV-II), a + or — to indicate the strand, and a d to indicate a detector. PNG-1, HSC-CTCL-11B, HUT 102B2, and MT-2 are cell lines containing HTLV-I isolates from an asymptomatic Papua New Guinean (53, 54), a Liberian of American slave descent who had a polyclonal T-cell lymphocytosis (11), an African-American adult T-cell lymphoma/leukemia (ATL) patient (33), and a Japanese ATL patient (28), respectively. MoT contains an HTLV-II isolate from a Caucasian American with hairy-cell leukemia (20). Input DNA from each isolate was normalized to 10,000 copies. The isolate HSC-CTCL-11B was used to verify the efficiency of each set of primers and probe. This isolate could be consistently detected at an input of 10 copies of HTLV-I DNA for all except one HTLV-I primer pair described, thereby demonstrating their sensitivity. +, band appears on liquid hybridization after 30 or 60 cycles of PCR; —, no hybridization band after 60 cycles of PCR; LTR, long terminal repeat.

^b 60 cycles.

varies from HTLV-I ATK by 9.7% (Fig. 1B). However, four transversions from A to T, which result in three amino acid changes, were conserved in all of the sequences depicted here and in 49 of 50 other previously published HTLV-I isolates irrespective of geographic origin (8, 10a, 21). Hence, these four base substitutions may represent sequencing errors of the original ATK clone.

A deletion in six independent gp21 clones of PNG-1 resulted in an altered reading frame of the transmembrane protein (Fig. 1B). It is not known whether a compensatory insertion lies outside the region sequenced. This deletion, of course, could represent a mutational event that occurred during cell culture, amplification, cloning, or sequencing. However, this change most likely happened during culture, since it was not seen in similar analyses performed on fresh peripheral blood mononuclear cells from this patient (data not shown) and since no deletions were observed in this region of gp21 when we sequenced 58 clones from 19 different HTLV-I isolates as well as 10 replicates from an HTLV-I-containing plasmid (10a). One PNG-1 gp21 clone of six contained a G substitution that is not present in the others. This change may also represent a sequencing error or, alternatively, the presence of quasispecies in the PNG-1 cell line, despite the fact that the cell line contained one integrated copy of proviral DNA by Southern blot analysis (54). The possibility of quasispecies also exists in our analyses of gp46 (Fig. 1C) and has been previously reported for other HTLV-I isolates (8, 10a, 16).

Figure 1C shows the sequence comparison of the HTLV variants from the 5' region of the *env* gene, which encodes the extracellular membrane protein gp46. We expected portions of this region to be either strongly conserved or highly

variable, since the protein it encodes determines cell tropism but is also under the selective pressure of the patient's immune system. Indeed, in Fig. 1C, one can see stretches of variability from a consensus sequence interrupted by conserved regions, for an overall variation for HTLV-II MoT and PNG-1 from prototype HTLV-I ATK of 30 and 6.9%, respectively. PNG-1 contained an insertion near the 5' end of gp46, but a deletion that restored the protein to the consensus frame occurred shortly thereafter (Fig. 1C-1). This specific region of the gp46 (positions 5250 to 5265) exhibited considerable variation in the STLV-I and HTLV-II isolates. This nucleotide sequence and its corresponding peptide may prove to be valuable in diagnosing type-specific PTLV infection by creating oligonucleotide primers for PCR or peptides for enzyme-linked immunosorbent assays and Western blot immunoassays capable of identifying individual variants. Whether these mutations change the function of this portion of the major envelope glycoprotein with respect to infectivity and pathogenicity needs to be clarified.

Two specific transitions from the Japanese (ATK) sequence at positions 5467 and 5471 (Fig. 1C-2) are also shared by all the isolates compared here and elsewhere (8, 21). Therefore, these mutations do not appear to be specific for Japanese isolates, as was previously suggested (9). Though geographic-specific mutations do exist (9), especially in gp46 (31), PNG-1 contained deoxyadenosine at position 5348, which was thought to be specific for African isolates (9), and the Liberian isolate contained deoxyadenosine at position 5534, thought to be specific for Japanese isolates (9). More Papua New Guinean isolates must be obtained and sequenced to determine the frequencies and specificities of mutations particular to viruses from this geographic region.

21

C-1 gp 46

2558

FIG. 1. Sequence analysis of amplified, cloned DNA. Primers (underlined) specific for four different regions of the HTLV-I genome and two corresponding regions of the HTLV-II genome were employed to amplify target DNA, which was subsequently cloned and sequenced. DNAs from the HTLV-I-infected cell lines PNG-1 and HSC-CTCL-11B were amplified and sequenced in specified regions of the *pol*, *pX*, and *env* (gp21 and gp46) genes. In addition, DNAs from the cell line MoT and from two HTLV-II-infected patients (RW and FF) were amplified in equivalent regions of the *pol* and *pX* genes. Four to eight clones of any particular region were sequenced. Fractions above a nucleotide change indicate the frequency of that mutation when it was not seen in all clones from an individual patient. Corresponding sequences of STLV-I PtM3 are included where data are available. An insertion site is indicated by ^, and the two deletions are represented by *. Dashed lines are regions where the sequence of the isolate cannot be determined because the primers themselves are incorporated into the amplified PCR product. Lowercase letters in the prototype HTLV-I sequence correspond to positions of variation noted in the isolates listed below HTLV-I ATK (EMBL numbering). Amino acid changes are shown for regions between the primers. (A) ORF *pX*-II, (B) gp21, (C) gp46, and (D) *pol*.

Because of the importance of reverse transcription in the retroviral life cycle, there are regions of highly conserved sequence within the *pol* gene of the PTLV (24, 34, 42, 45). However, the divergence between the entire *pol* sequences of prototypes HTLV-I and -II is 37%, very close to the variation found in the complete proviral DNA (42, 45). For this study, primer pairs (Table 1) were chosen such that a region of the *pol* genes of all PTLV isolates could be generically amplified and yet would contain intervening sequences that reflected the variation seen in the entire *pol* gene and, in turn, the total proviral DNA (22, 24, 34, 42, 45).

As can be seen in Fig. 1D, there is a 40% difference in the *pol* sequences analyzed between prototype HTLV-I ATK and HTLV-II MoT, while there is strong sequence conservation among the Japanese, Caribbean, and Liberian HTLV-I isolates and also among the HTLV-II isolates. By contrast, PNG-1 possessed four amino acid changes in the amplified region, and the sequence varied by 9.3% from the prototype Japanese HTLV-I.

To elucidate the relationships among members of the PTLV, dendrograms were constructed by using the unweighted pair-group method of association (UPGMA) pro-

FIG. 2. Dendrograms for the PTLV family of retroviruses. Relationship trees are based on the 756-bp sequence from *pol*, gp21, gp46, and ORF *pX-II* (A); the 616-bp sequence from gp21, gp46, and ORF *pX-II* (B); and the 325-bp sequence from gp46 (C). Dendrograms were constructed by using the UPGMA method (29) for comparing the divergence patterns for HTLV-II MoT; STLV-I PtM3; and HTLV-I PNG-1, HSC-CTCL-11B, HS-35 (24), EL and CH (31, 34), ATK, and H5. Sequences were aligned to maximize homology, and differences between quaspecies were averaged together. Intraisolate variation among the PNG-1 clones was only 0.07% in dendrogram A, 0.09% in dendrogram B, and 0.12% in dendrogram C. Insertions and deletions were valued at 0.5 changes each and had minimal impact on the calculated divergence. When PNG-1 was compared directly with prototype HTLV-I, there were 53.1 changes (7.0%) in dendrogram A, 39.7 changes (6.4%) in dendrogram B, and 22.4 changes (6.9%) in dendrogram C. By contrast, HTLV-I EL differed from prototype HTLV-I by 3.2% in dendrogram A, 3.3% in dendrogram B, and 3.4% in dendrogram C, while STLV-I PtM3 differed by 9.6% in dendrogram B and 10.8% in dendrogram C.

cedure (29) (Fig. 2). Previous, similar analyses have indicated that the PTLV group has apparently evolved at a constant rate from a common ancestor, with HTLV-II having diverged from the progenitor of STLV-I and HTLV-I much earlier than these two subgroups separated from each other (18, 48). The phylogenetic clustering of the PTLV retroviruses differs from the pattern of the host species, which is indicative of interspecies transmission between humans and nonhuman primates (18). Further dissection of the interrelationship of the PTLV suggested that the subgroup represented by STLV-I PtM3 from Asia branched from HTLV-I earlier than did STLV-I from Africa and that HTLV-I originated in Africa and was then disseminated around the world (9, 13, 18, 34, 48). These data raise the question as to whether STLV-I PtM3 represents a subgroup ancestral to the HTLV-I and STLV-I from Africa and thereby suggest that the PTLV could have spread from Asia to Africa, allowing speculation regarding the origin of HTLV-I.

Figure 2A is a dendrogram constructed from the sequence data from all four regions (756 bp) (Fig. 1) of the PNG-1 isolate presented here and similar information available for HTLV-II and other HTLV-I isolates. To validate the relative positions of each isolate on the dendrogram, we applied the relative rate test (40). The data indicate that PNG-1 is more closely related to HTLV-II and diverged from a common ancestor of HTLV-I prior to all other HTLV-I isolates analyzed and, in particular, HTLV-I EL from Africa. Another dendrogram (Fig. 2B) exclusive of the *pol* region (616 bp) was constructed so that comparisons with the Asian STLV-I isolate and another Japanese HTLV-I isolate, H5 (47), could be made. In this analysis, the degree of divergence for PNG-1 decreased slightly relative to rates for the

other HTLV-I isolates, but now STLV-I PtM3 can be seen branching from the PTLV prior to PNG-1. A third dendrogram (Fig. 2C) was drawn by using just the data from the gp46 *env* gene sequences (325 bp). This was done so that PNG-1 could be compared with several African and additional Japanese and American HTLV-I isolates for which only this sequence is available (9). This analysis demonstrated a temporal order of divergence within the PTLV beginning with HTLV-II and then proceeding with STLV-I PtM3, PNG-1, African HTLV-I, and then the New World and Japanese HTLV-I isolates. All three dendrograms support the hypothesis that the STLV-I PtM3 and PNG-1 isolates are ancestral to all of the other HTLV-I isolates illustrated. These data predict, then, that HTLV-I originated in the Indo-Malay region, somehow was brought to Africa, and from there spread throughout the world.

Several caveats could be raised regarding these conclusions. The first is that too little sequence is analyzed to accurately compare the evolutionary relationships of the PTLV. This seems unlikely, since sequences from various positions throughout the PTLV genome were utilized for comparison, thereby averaging out the effects of random genetic drift and selective pressures. In addition, the clustering of the isolates presented (Fig. 2) is in agreement with similar dendrograms of the PTLV family based on available kilobase or full-length sequences of a subset of the same isolates (34). In fact, data published after the submission of this manuscript, including the full-length gp21 (16) and the entire long terminal repeat (38a) sequences of the PNG-1

isolate, indicated that the relationships shown in Fig. 2 are valid.

Recent infection with STLVI or cosmopolitan prototype HTLV-I cannot account for the origin of PNG-1. First, there have never been nonhuman primates in New Guinea. Second, anthropological data indicate that the Hagahai have been isolated from outsiders for thousands of years and have had no contact with Africans or Japanese. Further, data published after the submission of this manuscript indicate that gp21 *env* sequences derived from PNG-1 and HTLV-I strains from the mother of the individual who provided PNG-1 and from four unrelated Melanesian Solomon Islanders show considerable homology to each other and significant variability from all other HTLV-I isolates (16).

It is remotely possible that the cultured PNG-1 isolate is defective and does not reflect the true PNG-1 sequence. The PNG-1 DNA utilized in this study was extracted from the CD8-positive T-cell line PNG-1, which had been cultured for 12 months (53, 54). These cells contain predominantly one full-length copy of HTLV-I proviral DNA (54), and our sequence analyses of up to eight different clones from each region examined showed minimal evidence of intrapatient variation. However, the provirus in this culture appears to be latent in that while analyses for viral RNA have not been performed, no evidence of HTLV-I reverse transcriptase or p19 could be detected in the culture by using previously described techniques (32, 33; data not shown). Further, in cocultivation experiments (25), HTLV-I PNG-1 could not be transmitted to target human T lymphocytes. It is possible that, as has been published previously (27), the cultured CD8-positive cells are actually cytotoxic for HTLV-I-expressing cells and that only latently infected cells persist in the PNG-1 culture. Nonetheless, it is doubtful that these biological characteristics of the PNG-1 culture would negate the implications of the comparative sequence analyses because, again, an independent analysis (16) of the gp21 *env* regions of PNG-1 and other Melanesian HTLV-I confirms our conclusions. Likewise, the fact that the sequences for HTLV-II MoT, HTLV-I HUT 102B2, and HTLV-I MT-2, which are all derived from cell lines cultured for over a decade, are clustered in their appropriate geographic associations (Fig. 2C) makes it unlikely that the comparatively short-term culture of PNG-1 would have resulted in significant sequence variation.

The data presented here and elsewhere (16), as well as the presence of endemic HTLV-I infection in Australian aborigines (3a), are consistent with an Indo-Malay origin for HTLV-I, but only further DNA sequence comparisons among PTLV isolates obtained worldwide will serve to establish the evolutionary relationships among all of the members of this group of retroviruses.

Nucleotide sequence accession numbers. The GenBank nucleotide sequence accession numbers are M76754 for PNG-1 pX-II, M76750 for HSC-CTCL-11B pX-II, M76753 for PNG-1 gp21, M76748 for HSC-CTCL-11B gp21, M76752 for PNG-1 gp46, M76749 for HSC-CTCL-11B gp46, M76755 for PNG-1 *pol*, and M76751 for HSC-CTCL-11B *pol*.

We acknowledge Timothy Spicer, Patricia Benz, Osama Ibrahim, Mary Rubert, Cheryl Brothers, Jayne Love, and Virginia Bryz-Gornia for technical assistance and Lori Tillis and Cheryl Brady for manuscript preparation. We also thank Michael Lane for help constructing the dendrograms and Lee Ratner for sharing his unpublished sequences.

This work was supported by grant RO1 HL43602-01.

ADDENDUM IN PROOF

During review of the manuscript a report was published in which the gp21 region of HTLV-I ATK was resequenced (T. F. Schulz, M.-L. Calabro, J. G. Hoad, C. V. F. Carrington, E. Matutes, D. Catovsky, and R. A. Weiss, *Virology* 184:483-491, 1991). The report by Schulz et al. confirms our postulate that the original ATK sequence contains errors at bases 6402, 6409, 6419, and 6432 (Fig. 1B). All of the T's at these four positions should be C's. These changes do not affect the branching pattern of the dendrograms (Fig. 2) but do indicate that HTLV-I ATK is slightly less variable than as shown.

REFERENCES

- Abbott, M. A., B. J. Poiesz, B. C. Byrne, S. Kwok, J. J. Sninsky, and G. D. Ehrlich. 1988. Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified in vitro. *J. Infect. Dis.* 158:1158-1169.
- Asher, D. M., J. Goudsmit, K. L. Pomeroy, R. M. Garruto, M. Bakker, S. G. Ono, N. Elliott, K. Harris, H. Askins, Z. Eldadah, A. D. Goldstein, and D. C. Gajdusek. 1988. Antibodies to HTLV-I in populations of the southwestern Pacific. *J. Med. Virol.* 26:339-351.
- Babona, D. V., and G. T. Nurse. 1988. HTLV-I antibodies in Papua New Guinea. *Lancet* ii:1148.
- Bastian, I. B. Personal communication.
- Blattner, W. A., D. W. Blayney, M. Robert-Guroff, V. S. Kalyanaraman, and R. C. Gallo. 1983. Epidemiology of human T-cell leukemia/lymphoma virus. *J. Infect. Dis.* 147:406-416.
- Blattner, W. A., V. S. Kalyanaraman, M. Robert-Guroff, T. A. Lister, D. A. Galton, P. S. Sarin, M. H. Crawford, D. Catovsky, M. Greaves, and R. C. Gallo. 1982. The human type C retrovirus, HTLV, in blacks from the Caribbean region, and relationship to adult T-cell leukemia/lymphoma. *Int. J. Cancer* 30:257-264.
- Catovsky, D., M. F. Greaves, M. Rose, D. A. Galton, A. W. Goiden, D. R. McCluskey, J. M. White, I. Lampert, G. Bourikas, R. Ireland, A. I. Brownell, J. M. Bridges, W. A. Blattner, and R. C. Gallo. 1982. Adult T-cell lymphoma/leukemia in blacks from the West Indies. *Lancet* i:639-643.
- Cavalli-Sforza, L. L., A. Piazza, P. Menozzi, and J. Mountain. 1988. Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proc. Natl. Acad. Sci. USA* 85:6002-6006.
- Daenke, S., S. Nightingale, J. K. Cruickshank, and C. R. M. Bangham. 1990. Sequence variants of human T-cell lymphotropic virus type I from patients with tropical spastic paraparesis and adult T-cell leukemia do not distinguish neurological from leukemic isolates. *J. Virol.* 64:1278-1282.
- De, B. K., M. D. Lairmore, K. Griffis, L. J. Williams, F. Villinger, T. C. Quinn, C. Brown, Nzilambi, M. Sugimoto, S. Araki, and T. M. Folks. 1991. Comparative analysis of nucleotide sequences of the partial envelope gene (5' domain) among human T lymphotropic virus type I (HTLV-I) isolates. *Virology* 182:413-419.
- Delaporte, E., A. Dupont, M. Peeters, R. Josse, M. Merlin, D. Schrijvers, B. Hamono, L. Bedjabaga, H. Cheringou, and F. Boyer. 1988. Epidemiology of HTLV-I on Gabon (western equatorial Africa). *Int. J. Cancer* 42:687-689.
- Ehrlich, G. D., J. Andrews, M. R. Sherman, S. G. Greenberg, and B. J. Poiesz. 1992. DNA sequence analysis of the gene encoding the HTLV-I P21E transmembrane protein reveals inter- and intrasolate genetic heterogeneity. *Virology* 186:619-627.
- Ehrlich, G. D., F. R. Davey, J. J. Kirshner, J. J. Sninsky, S. Kwok, D. J. Slamon, R. Kalish, and B. J. Poiesz. 1989. A polyclonal CD4+ and CD8+ lymphocytosis in a patient doubly infected with HTLV-I and HIV-1: a clinical and molecular analysis. *Am. J. Hematol.* 30:128-139.
- Fleming, A. F., N. Yamamoto, S. R. Bhusnurmath, R. Mahara-

- jan, J. Schneider, and G. Hunsmann. 1983. Antibodies to ATL (HTLV) in Nigerian blood donors and patients with chronic lymphatic leukemia or lymphoma. *Lancet* ii:334-335.
13. Gallo, R. C., A. Sliiski, and F. Wong-Staal. 1983. Origin of human T-cell leukaemia-lymphoma virus. *Lancet* ii:962-963.
 14. Garruto, R. M., M. Slover, R. Yanagihara, C. A. Mora, S. S. Alexander, D. M. Asher, P. Rodgers-Johnson, and D. C. Gajdusek. 1990. High prevalence of human T-lymphotropic virus type I infection in isolated populations of the western Pacific region confirmed by Western immunoblot. *Am. J. Hum. Biol.* 2:439-447.
 15. Gessain, A., J. C. Vernant, L. Maurs, F. Barin, O. Gout, A. Calender, and G. De The. 1985. Antibodies to human T-lymphotropic virus type-I in patients with tropical spastic paraparesis. *Lancet* ii:407-409.
 16. Gessain, A., R. Yanagihara, G. Franchini, R. M. Garruto, C. L. Jenkins, A. B. Ajdukiewicz, R. C. Gallo, and D. C. Gajdusek. 1991. Highly divergent molecular variants of human T-lymphotropic virus type I from isolated populations in Papua New Guinea and the Solomon Islands. *Proc. Natl. Acad. Sci. USA* 88:7694-7698.
 17. Hinuma, Y., K. Nagata, M. Hanaoka, M. Nakai, T. Matsumoto, K. I. Kinoshita, S. Shirakawa, and I. Miyoshi. 1981. Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proc. Natl. Acad. Sci. USA* 78:6476-6480.
 18. Ina, Y., and T. Gojobori. 1990. Molecular evolution of human T-cell leukemia virus. *J. Mol. Evol.* 31:493-499.
 19. Inoue, J. I., M. Yoshida, and M. Seiki. 1987. Transcriptional (p40x) and post-transcriptional (p27 x-III) regulators are required for the expression and replication of human T-cell leukemia virus type I genes. *Proc. Natl. Acad. Sci. USA* 84:3653-3657.
 20. Kalyanaraman, V. S., M. G. Sarngadharan, I. Miyoshi, D. Blayney, D. Golde, and R. C. Gallo. 1982. A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia. *Science* 218:571-573.
 21. Kinoshita, T., A. Tsujimoto, and K. Shimotohno. 1991. Sequence variations in LTR and env regions of HTLV-I do not discriminate between the virus from patients with HTLV-I associated myelopathy and adult T-cell leukemia. *Int. J. Cancer* 47:491-495.
 22. Kwok, S., D. Kellogg, G. Ehrlich, B. Poiesz, S. Bhagavati, and J. J. Sninsky. 1988. Characterization of a sequence of human T-cell leukemia virus type I from a patient with chronic progressive myelopathy. *J. Infect. Dis.* 158:1193-1197.
 23. Maeda, Y., M. Furukawa, Y. Takehara, K. Yoshimura, K. Miyamoto, T. Matsuura, Y. Morishima, K. Tajima, K. Okochi, and Y. Hinuma. 1984. Prevalence of possible adult T-cell leukemia virus-carriers among volunteer blood donors in Japan: a nation-wide study. *Int. J. Cancer* 33:717-720.
 24. Malik, K. T., J. Even, and A. Karpas. 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 ATL/HTLV-I subgroup. *J. Gen. Virol.* 69:1695-1710.
 25. Merl, S., B. Kloster, J. Moore, C. Hubbell, R. Tomar, F. Davey, D. Kalinowski, A. Planas, G. Ehrlich, D. Clark, R. Comis, and B. J. Poiesz. 1984. Efficient transformation of previously activated and dividing T lymphocytes by human T cell leukemia-lymphoma virus. *Blood* 64:967-974.
 26. Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* 19:269-276.
 27. Mitsuya, H., R. F. Jarrett, J. Cossman, O. J. Cohen, C.-S. Kao, H.-G. Guo, M. S. Reitz, and S. Broder. 1986. Infection of human T lymphotropic virus-I-specific immune T cell clones by human T lymphotropic virus-I. *J. Clin. Invest.* 78:1302-1310.
 28. Miyoshi, I., I. Kubonishi, S. Yoshimoto, T. Akagi, Y. Ohtsuki, Y. Shiraiishi, K. Nagata, and Y. Hinuma. 1981. Type C virus particles in a cord T-cell line derived by cocultivating normal human cord leukocytes and human leukaemic T cells. *Nature* (London) 294:770-771.
 29. Nei, M. 1975. Molecular population genetics and evolution, p. 199. In A. Neuberger and E. L. Tatum (ed.), *Frontiers of biology*. American Elsevier Publishing Co., Inc., New York.
 30. Osame, M., K. Usuku, S. Izumo, N. Ijichi, H. Amitani, A. Igata, M. Matsumoto, and M. Tara. 1986. HTLV-I associated myelopathy, a new entity. *Lancet* i:1031-1032.
 31. Paine, E., J. Garcia, T. C. Philpott, G. Shaw, and L. Ratner. 1991. Limited sequence variation in human T-lymphotropic virus type I isolates from North American and African patients. *Virology* 182:111-123.
 32. Papsidero, L., F. Swartzwelder, M. Sheu, R. Montagna, G. Ehrlich, S. Bhagavati, H. Dosik, J. Sninsky, and B. J. Poiesz. 1990. Immunodetection of human T-cell lymphotropic virus type I core protein in biological samples by using a monoclonal antibody immunoassay. *J. Clin. Microbiol.* 28:949-955.
 33. Poiesz, B. J., F. W. Ruscetti, A. F. Gazdar, P. A. Bunn, J. D. Minna, and R. C. Gallo. 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. USA* 77:7415-7419.
 34. Ratner, L., T. Philpott, and D. B. Trowbridge. 1991. Nucleotide sequence analysis of isolates of human lymphotropic virus type I of diverse geographical origins. *AIDS Res. Hum. Retroviruses* 7:923-941.
 35. Rodgers-Johnson, P. E., D. C. Gajdusek, O. S. C. Morgan, V. Zaninovic, P. S. Sarin, and D. S. Graham. 1985. HTLV-I and HTLV-II antibodies in tropical spastic paraparesis. *Lancet* ii:1247-1248.
 36. Roman, G. C., B. S. Schoenberg, D. L. Madden, J. L. Sever, J. Hugon, A. Ludolph, and P. S. Spencer. 1987. Human T-lymphotropic virus type I antibodies in the serum of patients with tropical spastic paraparesis in the Seychelles. *Arch. Neurol.* 44:605-607.
 37. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1987. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.
 38. Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of β -globin genomic sequences and restriction site analysis of diagnosis of sickle cell anemia. *Science* 230:1350-1354.
 - 38a. Sakseena, N. K., M. P. Sherman, R. Yanagihara, D. K. Dube, and B. J. Poiesz. Submitted for publication.
 39. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
 40. Sarich, V. M., and A. C. Wilson. 1973. Generation time and genomic evolution in primates. *Science* 179:1144-1146.
 41. Saxinger, W., W. A. Blattner, P. H. Levine, J. Clark, R. Biggar, M. Hoh, J. Moghissi, P. Jacobs, L. Wilson, R. Jacobsen, R. Crookes, M. Strong, Ansari, A. G. Dean, F. K. Nkrumah, N. Murali, and R. C. Gallo. 1984. Human T-cell leukemia virus (HTLV) antibodies in Africa. *Science* 225:1473-1476.
 42. Seiki, M., S. Hattori, Y. Hirayama, and M. Yoshida. 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.
 43. Seiki, M., A. Hikikoshi, T. Taniguchi, and M. Yoshida. 1985. Expression of the pX gene of HTLV-I: general splicing mechanism in the HTLV family. *Science* 228:1532-1534.
 44. Shaw, G. M., M. A. Gonda, G. H. Flickinger, B. H. Hahn, R. C. Gallo, and F. Wong-Staal. 1984. Genomes of evolutionarily divergent members of the human T-cell leukemia virus family (HTLV-I and HTLV-II) are highly conserved, especially in pX. *Proc. Natl. Acad. Sci. USA* 81:4544-4548.
 45. Shimotohno, K., Y. Takahashi, N. Shimizu, T. Gojobori, D. W. Golde, I. S. Y. Chen, M. Mieva, and T. Sugimura. 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.
 46. Teich, N. 1984. Taxonomy of retroviruses, p. 25-208. In R.

- Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), Molecular biology of tumor viruses: RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
47. **Tsujimoto, A., T. Teruuchi, J. Imamura, K. Shimotohno, I. Miyoshi, and M. Miwa.** 1988. Nucleotide sequence analysis of a provirus derived from HTLV-I-associated myelopathy (HAM). *Mol. Biol. Med.* **5**:29–42.
 48. **Watanabe, T., M. Seiki, Y. Hirayama, and M. Yoshida.** 1986. Human T-cell leukemia virus type I is a member of the African subtype of simian viruses (STLV). *Virology* **148**:385–388.
 49. **Watanabe, T., M. Seiki, H. Tsujimoto, I. Miyoshi, M. Hayami, and M. Yoshida.** 1985. Sequence homology of the simian retrovirus genome with human T-cell leukemia virus type I. *Virology* **144**:59–65.
 50. **Weber, J. N., N. Banatvala, S. Clayden, K. P. W. J. McAdam, S. Palmer, H. Moulds, J. Tossell, P. Dilger, R. Thorpe, and S. Amann.** 1989. HTLV-I infection in Papua New Guinea: evidence for serologic false positive. *J. Infect. Dis.* **159**:1025–1028.
 51. **Yanagihara, R., A. B. Ajdukiewicz, V. R. Nerurkar, R. M. Garruto, and D. C. Gajdusek.** 1991. Verification of HTLV-I infection in the Solomon Islands by virus isolation and gene amplification. *Jpn. J. Cancer Res.* **82**:240–244.
 52. **Yanagihara, R., C. L. Jenkins, S. S. Alexander, C. A. Mora, and R. M. Garruto.** 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.
 53. **Yanagihara, R., V. R. Nerurkar, and A. B. Ajdukiewicz.** 1991. Comparison between strains of HTLV-I isolated from inhabitants of the Solomon Islands and Papua New Guinea. *J. Infect. Dis.* **164**:443–449.
 54. **Yanagihara, R., V. R. Nerurkar, R. M. Garruto, M. A. Miller, M. E. Leon-Monzon, C. L. Jenkins, R. C. Sanders, P. P. Liberski, M. P. Alpers, and D. C. Gajdusek.** 1991. Characterization of a variant of human T-lymphotropic virus type I isolated from a healthy member of a remote, recently contacted group in Papua New Guinea. *Proc. Natl. Acad. Sci. USA* **88**:1446–1450.