code for 352 amino acids, is joined to the ATGG (5180 to 5183) of the *env* gene so that ATG is aligned in the frame. Thus, the initiation codon ATG for *env* is used for initiation of $p40^x$ translation. The other four cDNA clones showed similar restriction maps although they are also defective, supporting the conclusion that the sequenced cDNA clone represents the majority of the *pX* mRNA population. This conclusion was also consistent with the observation that the 2.1-kb mRNA did not significantly hybridize with the U5 probe. Similar splicing was also found in HTLV-II by Wachsman *et al.* (see *24a*).

The structure of the cDNA clone showed that subgenomic mRNA for $p40^x$ is formed by two-step splicing (Fig. 3A) and that one of the splice donor sites is located in the R region of LTR. This feature is unique to HTLV and might relate to the unusually long R region. We have proposed (*10*) that the ability of the long R sequence to form a secondary structure at the 3′ end of viral mRNA plays a role in transcriptional termination. In addition to this, the R sequence was suggested to regulate mRNA splicing expressing the *env* and *pX* genes, because a complementary sequence to this splicing donor site is found in the R region just after this donor site. Thus the donor site can form a stem structure with 18 bases (Fig. 3B). The secondary structure of the transcript at this splicing donor site would allow this region to compete with U1 RNA in a small nuclear ribonucleoprotein complex; the U1 complex is involved in exact RNA splicing (*25*). Alteration of the pairing ratio of U1 RNA to the stem and loop structure may control the splicing in the R region, eventually affecting the expression of the *env* and *pX* genes. For specific regulation of the *pX* expression, other mechanisms may be required.

One of the splicing acceptor sites is located in the *pol* region, 187 bases upstream from the ATG for the *env* gene. This structure indicates that the subgenomic RNA generated by single splicing is *env* mRNA (Fig. 3A). The calculated size of this spliced product, 3985 bases without the polyadenylated stretch, is consistent with that of one of the subgenomic mRNA's detected in infected cell lines. Probably, the other splicing takes place on the *env* mRNA and produces *pX* mRNA coding for $p40^x$. In this mRNA, the initiation codon, ATG for the *env*, is used to initiate the translation of $p40^x$, and only the first methionine is brought onto $p40^x$ from the *env* domain. Thus the molecular weight of the *pX* gene product was calculated as 39,482.

The key sequences for splicing found in HTLV-I are also present at corresponding positions in HTLV-II (*11, 26*), BLV (*12*), and STLV (*13*) as shown in Fig. 3, suggesting that the unusual splicing mechanisms producing the *env* and *pX* mRNA's are common to all four viruses. A minor exception occurs in BLV; the initiation codon for *pX* translation is located 44 bases downstream from that of the *env* gene, out of the frame. Thus, the *pX* and *env* genes do not share the same initiation codon. However, AIDS-associated viruses (*27*) are different from the other members of HTLV family in the organization of these critical sequences. Thus, a mechanism of the gene expression of AIDS-associated viruses seem to be different from the others.

MOTOHARU SEIKI
ATSUKO HIKIKOSHI
Department of Viral Oncology, Cancer Institute, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan
TADATSUGU TANIGUCHI
Institute for Molecular and Cellular Biology, Osaka University, Suita-shi, Osaka 565, Japan
MITSUAKI YOSHIDA
Department of Viral Oncology, Cancer Institute

References and Notes

1. B. J. Poiesz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7415 (1980); M. R. Reitz *et al.*, *ibid.* **78**, 1887 (1981).
2. Y. Hinuma *et al.*, *ibid.*, p. 6476.
3. M. Yoshida, I. Miyoshi, Y. Hinuma, *ibid.* **79**, 2031 (1982).
4. M. Robert-Guroff *et al.*, *Science* **215**, 975 (1982); V. S. Kalyanaraman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1653 (1982).
5. Y. Hinuma *et al.*, *Int. J. Cancer* **29**, 631 (1982).
6. W. A. Blattner *et al.*, *ibid.* **30**, 257 (1982); J. D. Schüpback, *et al.*, *Cancer Res.* **43**, 886 (1983).
7. W. Saxinger *et al.*, *Science* **225**, 1473 (1984).
8. M. Yoshida *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2534 (1984).
9. M. Seiki, S. Hattori, M. Yoshida, *ibid.* **79**, 6899 (1982).
10. M. Seiki *et al.*, *ibid.* **80**, 3618 (1983).
11. K. Shimotohno *et al.*, *ibid.* **81**, 6657 (1984); W. A. Haseltine *et al.*, *Science* **225**, 419 (1984).
12. N. R. Rice *et al.*, *Virology* **138**, 82 (1984); N. Sagata *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
13. A. Komuro *et al.*, *Virology* **138**, 373 (1984); T. Watanabe *et al.*, *ibid.*, in press; H.-G. Guo, F. Wong-Staal, R. C. Gallo, *Science* **223**, 1195 (1984).
14. T. Kiyokawa *et al.*, *GANN* **75**, 747 (1984); M. Miwa *et al.*, *ibid.*, p. 751; T. H. Lee *et al.*, *Science* **226**, 57 (1984); D. J. Slamon *et al.*, *ibid.*, p. 61.
15. J. G. Sodroski, C. A. Rosen, W. A. Haseltine, *Science* **225**, 381 (1984); J. Fujisawa *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
16. M. Seiki, R. Eddy, T. B. Shows, M. Yoshida, *Nature (London)* **309**, 640 (1984).
17. M. Tateno *et al.*, *J. Exp. Med.* **159**, 1105 (1984).
18. T. Watanabe, M. Seiki, M. Yoshida, *Virology* **133**, 238 (1984).
19. H. Land *et al.*, *Nucleic Acids Res.* **9**, 2251 (1981).
20. M. Casadaban and S. M. Cohen, *J. Mol. Biol.* **138**, 179 (1980).
21. P. J. Southern and P. Berg, *J. Mol. Appl. Genet.* **1**, 327 (1982).
22. P. A. Sharp, *Cell* **23**, 643 (1984).
23. T. H. Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3856 (1984).
24. W. Wachsman *et al.*, *Science* **226**, 177 (1984).
24a. W. Wachsman *et al.*, *ibid.* **228**, 1534 (1985).
25. M. R. Lerner *et al.*, *Nature (London)* **283**, 220 (1980); J. Rogers and R. Wall, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1877 (1980).
26. K. Shimotohno *et al.*, *ibid.*, in press.
27. L. Ratner *et al.*, *Nature (London)* **313**, 277 (1985); R. Sanchez-Pescador *et al.*, *Science* **227**, 484 (1985); M. A. Muesing *et al.*, *Nature (London)* **313**, 450 (1985); S. Wain-Hobson *et al.*, *Cell* **40**, 9 (1985).
28. P. Mellon *et al.*, *Cell* **27**, 279 (1981).
29. We thank G. Yamada, Institute for Molecular and Cellular Biology, Osaka University, for his useful discussion and help. Supported in part by a Grant-in-Aid for Special Project Research, Cancer-Bioscience; a Grant-in-Aid for Cancer Research, from the Ministry of Education, Science and Culture of Japan; and by a Research Grant of the Princess Takamatsu Cancer Research Fund.

12 February 1985; accepted 29 April 1985

HTLV *x*-Gene Product: Requirement for the *env* Methionine Initiation Codon

Abstract. *The human T-cell leukemia viruses (HTLV) are replication-competent retroviruses whose genomes contain* gag, pol, *and* env *genes as well as a fourth gene, termed* x, *which is believed to be the transforming gene of HTLV. The product of the* x *gene is now shown to be encoded by a 2.1-kilobase messenger RNA derived by splicing of at least two introns. By means of S_1 nuclease mapping of this RNA and nucleic acid sequence analysis of a complementary DNA clone, the complete primary structure of the* x*-gene product has been determined. It is encoded by sequences containing the* env *initiation codon and one nucleotide of the next codon spliced to the major open reading frame of the HTLV-I and HTLV-II* x *gene.*

The human T-cell leukemia viruses (HTLV-I and HTLV-II) are associated with specific T-cell malignancies in man. HTLV-I-related adult T-cell leukemia is endemic to parts of Japan, the Caribbean, and Africa; HTLV-II is associated with a single case of T-cell–variant hairy-cell leukemia (*1–5*). Both viruses will transform normal, human, peripheral blood T cells in vitro as defined by their continued proliferation in the absence of exogenous interleukin-2 (*6–9*). The mechanism of HTLV-induced T-cell transformation is unknown, although these retroviruses, and bovine leukemia virus (BLV), appear to use a mechanism distinct from that of other animal retroviruses. Molecular studies of the HTLV

genome in tumor cells from patients and transformed cell lines indicated that there are no preferential sites of viral integration (*10, 11*). DNA hybridization and nucleic acid sequence analyses of HTLV-I and HTLV-II genomes revealed that the HTLV proviral genome has no homology with known oncogenes or normal cellular sequences (*12, 13*). However, nucleic acid sequence analysis of HTLV-I by Seiki *et al.* identified a region of unknown function (designated X) that is located between *env* and the 3′ long terminal repeat (LTR) (*12*). A similar X region was identified in the HTLV-II and BLV genomes (*14–17*). Recently, we and others demonstrated that this X region of the HTLV genome contains a new gene, *x* (*18, 19*). This gene is transcribed into a 2.1-kilobase (kb) messenger RNA (mRNA) and encodes a protein of 40 or 37 kilodaltons (kD), (designated $p40^{xI}$ or $p37^{xII}$), in cells infected with HTLV-I or HTLV-II, respectively. Thus, unlike other replication-competent retroviruses, the HTLV class of retrovirus contains four genes: *gag*, *pol*, *env*, and *x*. As the *x* gene only occurs in this class of retrovirus, it is hypothesized that its product is responsible for the transforming properties of HTLV.

The exact primary structure of $p40^{xI}$ and $p37^{xII}$ is unknown. Immunoprecipitation of $p40^{xI}$ and $p37^{xII}$ with antisera directed against synthetic peptides (with the peptide sequences deduced from the HTLV-I nucleic acid sequence) demonstrated that the *x* product is encoded primarily from the major open reading frame of the X region (*19*). The size of the *x*-encoded proteins, and the fact that antisera were developed against peptides derived from sequences upstream of the first methionine codon in *x*, led us to predict that the NH_2-terminus of these proteins is derived from upstream viral sequences not contiguous with the major open reading frame. We demonstrate here an RNA processing scheme that would generate the complete *x*-encoded product. On the basis of the nucleic acid sequence of the HTLV-I and HTLV-II *x* mRNA, we have deduced the primary structure of $p40^{xI}$ and $p37^{xII}$ as well as a predicted *x* protein from BLV.

Typically, retroviral subgenomic mRNA's are composed of 5′ leader sequences extending from the cap site in the LTR into the *gag* gene, and a downstream exon such as *env* or *src* (*20–22*). Analyses of the HTLV-I and HTLV-II nucleic acid sequences did not reveal any potential initiation codon from the *gag* region that would maintain the *x* open reading frame and contribute less than 0.8 kb to the 2.1-kb *x* mRNA [as at least 1.5 kb of the mRNA would be derived from the downstream exon that contains the *x* gene sequences (*15, 18*)]. Therefore, we postulated the existence of a third exon that contains the initiation codon for the *x* gene and is located between a 5′ leader exon and the 3′ *x* gene exon. Further analysis of the HTLV-I and HTLV-II nucleic acid sequences in the region of the *pol–env* junction revealed the conservation of potential splice donor signals, located just downstream of the *env* methionine initiation codon, that would maintain the major open reading frame of *x* in a spliced mRNA.

To confirm the existence of these potential splice donor sites, we performed S_1 nuclease analysis on RNA from the HTLV-II–infected cell lines Mo-T and JLB-II, and from the HTLV-I–infected cell lines SLB-I and SLB-II. The hybridization probes (Fig. 1A) were prepared from the pH-6 clone of HTLV-II and the pHT-1 clone of HTLV-I to span the *pol-env* junction. Figure 1B shows the DNA fragments protected by hybridization to the RNA. A band indicative of a 90- and 93-nucleotide long protected fragment was observed for RNA from HTLV-I– and HTLV-II–infected cells, respectively. These coincide exactly with the location of the potential splice donor sequences in *env* at base pair 5183 in both the HTLV-I and HTLV-II proviral genomes (*12, 23*). A second band of 238

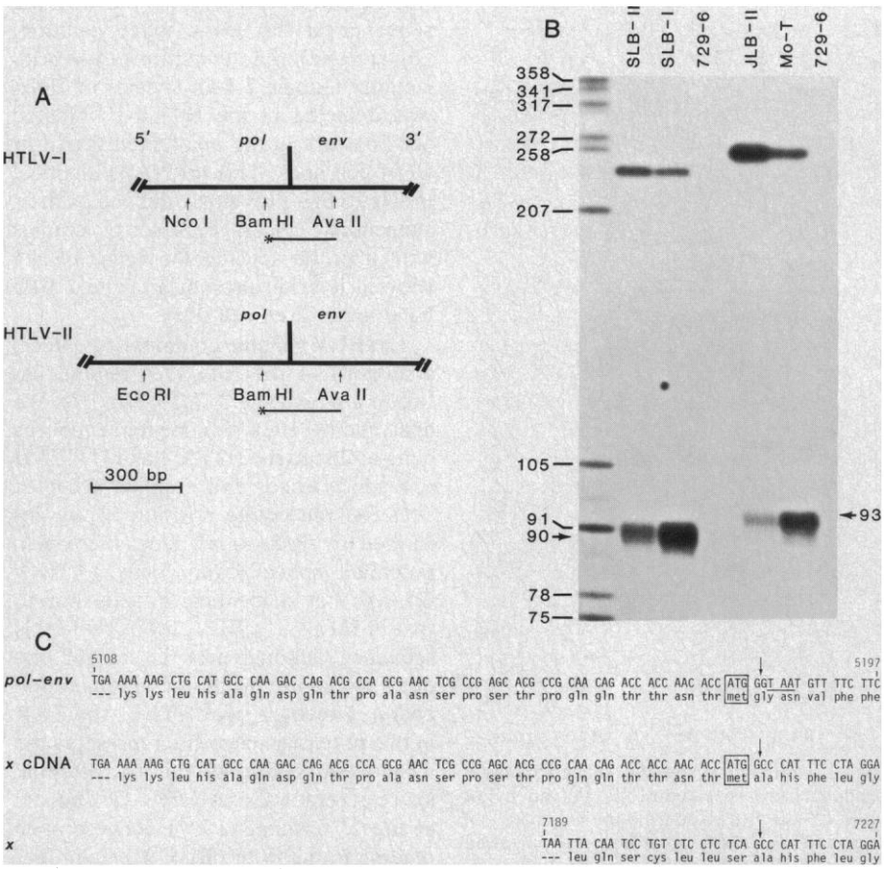

Fig. 1. Localization of the *x* gene splice donor sequences in *env*. (A) Partial restriction enzyme map of the pHT-1 (HTLV-I) and pH-6 (HTLV-II) DNA clones. The junction between *pol* and *env* is indicated as a vertical bar. The Bam HI–Ava II hybridization probe used for S_1 nuclease mapping is shown as a thin line below each map with an asterisk denoting the site at which 3′ ^{32}P end-labeling was done by means of the large fragment of DNA polymerase. (B) Hybridization-S_1 nuclease analysis was performed as described (*36, 37*). RNA was extracted from the HTLV-I–infected SLB-I and SLB-II cell lines (*18, 38*), the HTLV-II–infected Mo-T (*4, 39*) and JLB-II (*9*) cell lines, and the uninfected 729-6 B-cell line. The size of the marker DNA fragments is shown in nucleotides. The 90- and 93-nucleotide protected fragment for HTLV-I and HTLV-II, respectively, is denoted by an arrow. No protected fragment was observed after hybridization of control 729-6 RNA to the probe. (C) Nucleic acid and predicted amino acid sequence of a portion of an HTLV-II *x* cDNA clone isolated from a Mo-T cDNA library (*40*). Sequences from the *pol–env* junction and *x* open reading frame of HTLV-II are homologous to the *x* cDNA clone upstream and downstream, respectively, of the *x* cDNA splice junction. Nucleotide positions in the HTLV-II proviral genome are indicated (*23*). The methionine initiation codon of *env* and the *x* gene is denoted by a box. The arrow demarcates the respective sites for the *env* splice donor, *x* gene *env-x* splice junction, and the *x* splice acceptor. In the *pol-env* sequences the underlined TAA shows the position of the *pol* stop codon.

nucleotides for HTLV-I and 257 nucleotides for HTLV-II was observed; it corresponds to the protection, presumably by genomic or *env* mRNA, of the full length probe. No protected DNA fragments were observed when uninfected 729-6 cell RNA was hybridized with either probe and then digested with S_1 nuclease. The splice donor site occurs after a single nucleotide 3′ of the *env* initiation codon in HTLV-I and HTLV-II. A termination codon is present in the same reading frame, 36 and 24 codons upstream of the *env* initiation codon, in HTLV-I and HTLV-II, respectively. Therefore, the *x* gene uses the *env* initiation codon for its translation and the complete *x* coding sequences are formed by this initiation codon and one additional nucleotide spliced to the downstream major open reading frame of the *x* gene.

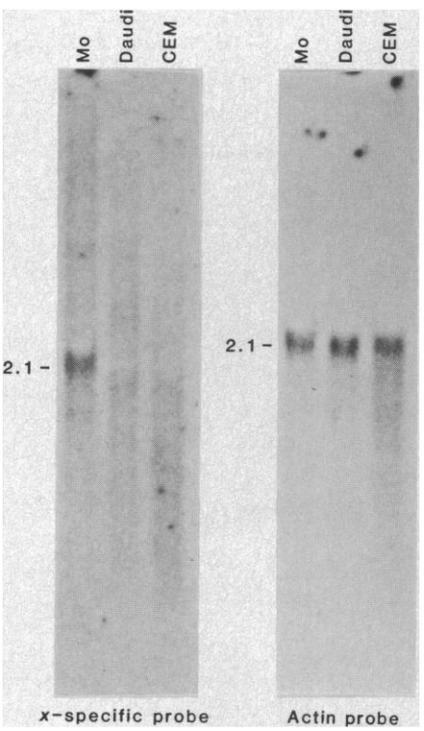

Fig. 2. Confirmation of the spliced junction sequences in *x*-specific mRNA. Polyadenylated RNA was extracted from the indicated cell lines and analyzed by Northern hybridization (*41*). The *x*-specific probe was derived by 5′ end-labeling of an oligonucleotide, 5′ AAATGGGCCATGGTGTTG 3′, which is complementary to the sequence surrounding the splice donor and acceptor junction. Under appropriate conditions of hybridization [4× standard saline citrate (SSC) at 30°C] and washing (2× SSC at 37° to 40°C), this probe hybridized only to sequences that are precisely complementary. This probe did not hybridize to noncontiguous complementary sequences, such as genomic mRNA in which the *env* splice donor and *x* splice acceptor sequences are separated. The actin probe was made from a human cDNA clone that had been isolated from a Mo-T cell cDNA library. Size is shown in kilobases. The Mo-T cells are infected with HTLV-II; the Daudi and CEM cells are uninfected.

We also isolated a complementary DNA (cDNA) clone made from mRNA of HTLV-II–infected Mo-T cells (Fig. 1C), which contains this same sequence at the splice junction of the *x* mRNA. This clone is about 400 nucleotides in size, and part of its sequence (Fig. 1C) verifies that the splicing of HTLV-II *x* mRNA involves the joining of a small segment of *env* with the downstream acceptor site of *x* at sequence locations determined by S_1 mapping.

To obtain further evidence that this splice junction is present in the 2.1-kb *x* subgenomic message, we analyzed RNA from both HTLV-II–infected and –uninfected cell lines by Northern hybridization (Fig. 2). An oligonucleotide probe was synthesized to correspond with the sequence at the *env-x* splice junction. Under appropriate conditions of hybridization, a single 2.1-kb species of RNA was detected in the HTLV-II–infected cells but not in the uninfected Daudi or CEM cell lines. This mRNA is identical in size to that previously defined as the *x* subgenomic message. Control studies with a probe specific for actin mRNA showed hybridization at a single 2.1-kb band in all three cell lines.

The BLV genome contains sequences, analogous to the HTLV X region, between *env* and the 3′ LTR (*16, 17*). We analyzed the BLV genome for sequences comparable to the HTLV-I and HTLV-II *env* splice donor and *x* splice acceptor sites. At nucleotide position 51, as described by Rice *et al.* (*16*), there is a potential splice donor site (ATGG/GTAA) that is identical in sequence to that in HTLV. In BLV, the ATG in this sequence, although near the start of *env*, does not appear to be the *env* initiation codon. However, like HTLV, the TAA in this potential splice site represents the termination codon of *pol*. A potential splice acceptor site (AG/CAAGT) occurs at the 5′ terminus of the longest open reading frame in the BLV X region [between nucleotides 2487 and 2488 in the pX2 reading frame (*16*)]. This site is preceded by pyrimidines, which is analogous to the splice acceptor site of *x* in HTLV-I and HTLV-II. Because of the similarities in sequence among BLV, HTLV-I, and HTLV-II, we believe that these potential BLV splice sites are used to generate a BLV *x* mRNA. As with HTLV-I and HTLV-II, this predicted *x* coding region would be formed by splicing of the methionine initiation codon plus a single base to the downstream *x* open reading frame. Figure 3 shows the sequences surrounding the HTLV-I and HTLV-II *x* gene splice junction as well as that predicted for BLV. The consensus sequence for this splice junction is also shown.

The first two amino acids of the translated protein, methionine-alanine, are the same for all three viruses. The predicted size of the primary translation product from HTLV-I and HTLV-II *x* mRNA is consistent with the respective 40-kD and 37-kD proteins previously demonstrated to be encoded by *x* (*19*); however, post-translational modification cannot be excluded. On the basis of the published sequence of BLV, we predict that its *x*-encoded protein is 34 kD in size. Recombinant constructs of the HTLV-I and HTLV-II *x* gene synthesized according to the predicted mRNA structure have accurately expressed either $p40^{xI}$ or $p37^{xII}$.

Retroviral mRNA's start from the cap site at the U3/R junction in the 5′ LTR and terminate at the R/U5 junction in the 3′ LTR (*24–26*). Further S_1 nuclease analysis of RNA from these HTLV-I– and HTLV-II–infected cells identified a potential splice acceptor site upstream of the *env* and *x* methionine codon at nucleotide 5183 in both HTLV-I and HTLV-II. In addition, there was at least one potential splice donor site located in the HTLV-II 5′ LTR at nucleotide 449. Therefore, the HTLV *x* mRNA consists of at least three exons and has a total size of approximately 1.8 kb if the 3′ polyadenylate sequences are excluded. The major open reading frame of the *x* gene is

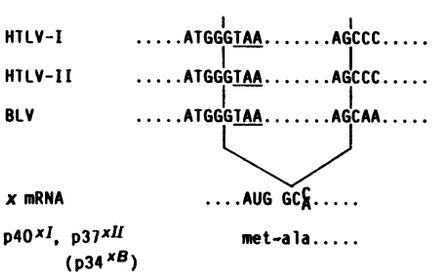

Fig. 3. Schematic representation of mRNA processing to generate the NH_2-terminal codons of the *x* gene. The sequence surrounding the *x* mRNA splice donor in *env* and the *x* mRNA splice acceptor in *x* for HTLV-I and HTLV-II as well as the predicted analogous sequences for BLV *x* mRNA are shown. The first two codons of the predicted primary translation product formed by mRNA processing are indicated. Although post-translational modifications cannot be excluded we speculate that the met-ala represents the NH_2-terminal amino acids of HTLV-I $p40^{xI}$, HTLV-II $p37^{xII}$, and $p34^{xB}$ predicted for BLV. The TAA is underlined to indicate the termination codon for *pol*. The nucleotide positions at the *env* splice donor/*x* splice acceptor junction are: 5183/7302 in HTLV-I (*12*), 5183/7214 in HTLV-II (*23*), and 51/2428 in BLV (*16*).

contained in the 1.5-kb 3′ exon of this subgenomic message. Immediately 5′ to this exon is a second exon of about 140 nucleotides that spans the *pol–env* junction and provides the initiation codon for the *x* gene. The 5′ exon appears to be a leader sequence of about 135 nucleotides. It extends from the cap site to a splice donor site located in the stem of the potential hairpin loop that has been proposed for the HTLV-I and HTLV-II LTR (*12, 27, 28*). We believe this to be the entire structure of the HTLV-II *x* subgenomic message. These data are also consistent with the observations of Seiki *et al.* (*29*) regarding the structure of HTLV-I *x* mRNA.

The functional basis for this precise scheme of mRNA processing to generate the *x* mRNA is unclear. The only other example of retroviral RNA processing in which the virus contains a gene between *env* and the 3′ LTR is Rous sarcoma virus (*30, 31*). In this case, sequences from *env* are not required to generate the *src* gene product (*32, 33*). It is possible that this type of mRNA processing, which adds only four nucleotides of coding sequence to downstream *x* sequences for HTLV, is a primary regulatory mechanism for the production of the *x* product. The function of $p40^{xI}$ and $p37^{xII}$ is unknown. These proteins may be important in viral transcription (*34, 35*). Recent *x* gene mutagenesis studies in our laboratory indicate that the *x* protein increases the efficiency of HTLV transcription (*35a*). Regulation of the expression of the *x* protein may be important for viral transcription and replication as well as for the initiation and maintenance of cellular transformation.

WILLIAM WACHSMAN
DAVID W. GOLDE
Division of Hematology-Oncology,
Department of Medicine,
UCLA School of Medicine,
Los Angeles, California 90024
PATRICIA A. TEMPLE
ELIZABETH C. ORR
STEVEN C. CLARK
Genetics Institute, Inc.,
Boston, Massachusetts
IRVIN S. Y. CHEN
Division of Hematology-Oncology,
Department of Medicine,
UCLA School of Medicine

References and Notes

1. B. J. Poiesz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7415 (1980).
2. Y. Hinuma *et al.*, *ibid.* **78**, 6476 (1981).
3. W. A. Blattner *et al.*, *Int. J. Cancer* **30**, 257 (1982).
4. A. Saxon, R. H. Stevens, D. W. Golde, *Ann. Intern. Med.* **88**, 323 (1978).
5. V. S. Kalyanaraman, M. G. Sarngadharan, M. Robert-Guroff, I. Miyoshi, D. Blayney, D. Golde, R. C. Gallo, *Science* **218**, 571 (1982).
6. I. Miyoshi, I. Kubonishi, S. Yoshimoto, T. Akagi, Y. Ohtsuki, Y. Shiraishi, K. Nagata, Y. Hinuma, *Nature (London)* **294**, 770 (1981).
7. N. Yamamoto *et al.*, *Science* **217**, 737 (1982).
8. M. Popovic *et al.*, *ibid.* **219**, 856 (1983).
9. I. S. Y. Chen, S. G. Quan, D. W. Golde, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7006 (1983).
10. M. Seiki, R. Eddy, T. B. Show, M. Yoshida, *Nature (London)* **309**, 640 (1984).
11. J. C. Gasson, personal communication.
12. M. Seiki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3618 (1983).
13. I. S. Y. Chen *et al.*, *Nature (London)* **305**, 502 (1983).
14. W. A. Haseltine *et al.*, *Science* **225**, 419 (1984).
15. K. Shimotohno *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6657 (1984).
16. N. R. Rice *et al.*, *Virology* **138**, 82 (1984).
17. N. Sagata *et al.*, *EMBO J.* **3**, 3231 (1984).
18. W. Wachsman *et al.*, *Science* **226**, 177 (1984).
19. D. J. Slamon *et al.*, *ibid.*, p. 61; T. H. Lee *et al.*, *ibid.*, p. 57.
20. P. Mellon and P. H. Duesberg, *Nature (London)* **270**, 631 (1977).
21. B. Cordell *et al.*, *Cell* **15**, 79 (1978).
22. P. B. Hackett *et al.*, *J. Virol.* **41**, 527 (1982).
23. K. Shimotohno *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
24. H. M. Temin, *Cell* **28**, 3 (1982).
25. T. Yamamoto, B. de Crombrugghe, I. Pastan, *ibid.* **22**, 787 (1980).
26. E. W. Benz, Jr., *et al.*, *Nature (London)* **288**, 665 (1980).
27. J. Sodroski *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4617 (1984).
28. N. Sagata *et al.*, *ibid.*, p. 4741.
29. M. Seiki *et al.*, *Science* **228**, 1532 (1985).
30. S. R. Weiss, H. E. Varmus, J. M. Bishop, *Cell* **12**, 983 (1977).
31. N. Quintrell *et al.*, *J. Mol. Biol.* **143**, 363 (1980).
32. A. F. Purchio *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1567 (1978).
33. K. Beemon and T. Hunter, *J. Virol.* **28**, 551 (1978).
34. J. G. Sodroski, C. A. Rosen, W. A. Haseltine, *Science* **225**, 381 (1984).
35. C. A. Rosen, J. G. Sodroski, R. Kettman, A. Burny, W. A. Haseltine, *ibid.* **227**, 320 (1985).
35a. I. S. Y. Chen *et al.*, *ibid.*, in press.
36. A. J. Berk and P. A. Sharp, *Cell* **12**, 721 (1977).
37. R. Weaver and C. Weissman, *Nucleic Acids Res.* **7**, 1175 (1979).
38. SLB-I and SLB-II are cell lines derived by infection of normal human leukocytes with HTLV-I in vitro.
39. A. Saxon *et al.*, *J. Immunol.* **120**, 777 (1978).
40. The cDNA clone was selected by hybridization to an oligonucleotide probe, 5′-CAAATCC TAGGAAATGG-3′ (complementary to nucleotides 7215–7231) (*23*), which is downstream of the *x* splice-acceptor site, and by the lack of hybridization to an oligonucleotide probe, 5′-CTCCTCTCAGCCCATTTCC-3′, which spans the *x* splice acceptor sequences (nucleotides 7204–7222) (*23*). A third hybridization probe was an HTLV-II cDNA clone previously described and located within the *x* gene (*13*). Nucleic acid sequencing of this portion of the *x* cDNA clone was performed by the standard dideoxy technique.
41. H. Lehrach *et al.*, *Biochemistry* **16**, 4743 (1977).
42. We thank J. Gasson for helpful comments on the manuscript, S. G. Quan for artwork, C. Nishikubo, S. G. Quan, A. Healy, and J. Fujii for technical assistance, and H. Merriman, B. Koers and W. Aft for editorial assistance. Supported by NCI grants CA 30388, CA 32737, CA 38597, CA 09297, and CA 16042; by grants PF-2182 and JFRA-99 from the American Cancer Society, and by a Bank of America-Giannini Foundation Fellowship (W.W.).

14 February 1985; accepted 2 May 1985

Predisposition to Hookworm Infection in Humans

Abstract. *Frequency distributions of parasitic helminths within human communities are invariably highly aggregated, the majority of worms occurring in relatively small fractions of the host populations. It has been suggested that the heavily infected individuals are predisposed to this state, not by chance, but by as yet undefined genetic, ecological, behavioral, or social factors. Analyses of individual post-treatment patterns of hookworm reinfection among 112 villagers in an endemic area of West Bengal provide quantitative evidence of predisposition to heavy infection. This observation has implications for the design of control programs based on chemotherapy because of the potential economic advantage of selective or targeted treatment as opposed to mass or blanket treatment.*

Epidemiological studies in various regions of the world have reported rates of post-treatment reinfection with hookworms and other intestinal helminths (*1*). Quantitative information on patterns of reinfection in individual patients, as opposed to groups of people, is extremely limited, however, and the evidence for predisposition to heavy infection is therefore inconclusive. This is in part a consequence of the small numbers of patients examined in previous studies (*2–5*). Here we present evidence of a statistically significant positive association between the adult worm burdens of individual patients prior to treatment (as judged by fecal egg counts) and the adult worm burdens in the same individuals following a period of reinfection.

Our analyses are based on field studies conducted in West Bengal, India, in 1968 to 1970 (*6–10*). The study area, its climate, methodological details (including methods of sample selection) and aspects of the epidemiology of hookworm parasitism are described elsewhere (*6–11*). Two groups of villagers from a rural area 40 miles (64 km) north of Calcutta were treated to remove their hookworms. Group 1 ($n = 34$) was treated with anthelmintics [bephenium hydroxynaphthoate and tetrachlorethylene (TCE)] in the dry season of 1968–1969 and then reexamined after 10, 12, 14, and 18 months. The monsoon season, a 4-month period during which conditions for frequent reinfection were optimal, began 6 months after treatment was completed. Group 2, the treated subsample of a study population of 735 villagers, was treated (with bephenium and either TCE or thiabendazole) in the same dry season and then reexamined at eight 2-month intervals (the total period included two monsoon seasons). People who provided fecal samples in at least four of