



Edinburgh Research Explorer

Complete sequence of enzootic nasal tumor virus, a retrovirus associated with transmissible intranasal tumors of sheep

Citation for published version:

Cousens, C, Minguijon, E, Dalziel, R, Ortin, A, Garcia, M, Park, J, Gonzales, L, Sharp, JM & de las Heras, M 1999, 'Complete sequence of enzootic nasal tumor virus, a retrovirus associated with transmissible intranasal tumors of sheep', *J Virol*, vol. 73, no. 5, pp. 3986-93.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

J Virol

Publisher Rights Statement:

Copyright © 1999, American Society for Microbiology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Complete Sequence of Enzoitic Nasal Tumor Virus, a Retrovirus Associated with Transmissible Intranasal Tumors of Sheep

CHRISTINA COUSENS,^{1,2*} ESMERALDA MINGUIJON,³ ROBERT G. DALZIEL,² AURORA ORTIN,³
MERCEDES GARCIA,⁴ JANE PARK,¹ LORENZO GONZALEZ,⁴ J. MICHAEL SHARP,¹
AND MARCELO DE LAS HERAS³

*Moredun Research Institute¹ and Royal (Dick) School of Veterinary Studies, University of Edinburgh,²
Edinburgh, United Kingdom, and Facultad de Veterinaria, Universidad de Zaragoza, Zaragoza³ and
Servicio de Investigacion y Mejora Agraria, Derio,⁴ Spain*

Received 9 September 1998/Accepted 3 February 1999

The sequence of the complete genome of ovine enzoitic nasal tumor virus, an exogenous retrovirus associated exclusively with contagious intranasal tumors of sheep, was determined. The genome is 7,434 nucleotides long and exhibits a genetic organization characteristic of type B and D oncoviruses. Enzoitic nasal tumor virus is closely related to the Jaagsiekte sheep retrovirus and to sheep endogenous retroviruses.

Animal models of retrovirus-induced tumors of the lymphoreticular system are numerous and have contributed greatly to our understanding of oncogenesis (reviewed in reference 10). However, animal models of retrovirus-associated epithelial tumors are restricted to mouse mammary tumors, sheep pulmonary adenomatosis (SPA), and enzoitic nasal tumors (ENT) of sheep and goats. Of these, ENT and SPA are also important for economic and animal welfare considerations in affected flocks. Both have long incubation periods, but once clinical signs become apparent the diseases are invariably fatal.

There is compelling evidence that SPA (also known as Jaagsiekte or ovine pulmonary carcinoma) is caused by Jaagsiekte sheep retrovirus (JSRV) (reviewed in reference (15), although classical approaches to investigate the presence of an etiological role for this virus continue to be hampered by the lack of a cell culture system to propagate JSRV. The relationship between ENT and its associated retrovirus, enzoitic nasal tumor virus (ENTV), is less well characterized. The presence of the retrovirus was first shown by electron microscopy in intranasal tumors of sheep (7, 23, 37) and goats (4, 5, 35) and in nasal fluid produced by these tumors. Antigenic cross-reactivity of ENTV with Mason-Pfizer monkey virus (MPMV) p27 (6, 31) and recombinant JSRV capsid protein (7) indicated that ENTV has a type D retroviral capsid. However, although the complete genome of JSRV has been sequenced (38), the data for ENTV is limited to two small regions of the *gag* gene (3, 29). The close homology between JSRV and sheep endogenous retroviral sequences (SERVs) (14, 38) was a problem for the development of JSRV-specific PCRs. This was finally resolved by the identification of small areas of sequence difference between JSRV and SERVs (1, 26). A similar approach is required to develop sensitive ENTV-specific detection techniques in order to elucidate the role of ENTV in disease. This is particularly important because ENTV, like JSRV, does not replicate in tissue culture.

Here we report the cloning and sequencing of the complete ENTV genome from an ovine case of ENT and investigate its relationship to JSRV, SERVs, and other retroviruses. This data is a vital step in the study of this animal model of epithelial

cell-derived tumors and opens the way to further studies to elucidate the pathogenesis and oncogenic mechanisms of this virus.

MATERIALS AND METHODS

Source of material. Nasal fluid from a sheep with naturally occurring ENT was collected, clarified, and concentrated as described previously (6). The resultant pellet was resuspended as a 100-fold concentrate in TNE buffer (0.01 M Tris-Cl [pH 7.5], 0.1 M NaCl, 0.001 M EDTA). Samples of ENT and kidney were collected from this sheep at necropsy and snap frozen in liquid nitrogen. Tissue samples were stored at -70°C . DNA was extracted from these tissues by the method of Wu et al. (36). The sheep was verified to be negative for SPA by macroscopic and microscopic pathologic testing at necropsy.

Density gradient fractionation. ENTV was purified from 200 μl of 100-fold-concentrated nasal fluid by isopycnic centrifugation on 20 to 55% (wt/wt) sucrose gradients (16). Fractions (0.5 ml) were collected, and fractions with densities between 1.17 and 1.19 g/ml were pooled, diluted in TNE buffer and centrifuged at $100,000 \times g$ for 1 h at 4°C . The pellet was resuspended in 0.5 ml of solution D (4 M guanidinium thiocyanate, 25 mM sodium citrate [pH 7.0], 0.5% [wt/vol] sarcosyl, 0.1 M 2-mercaptoethanol) for RNA preparation.

RNA preparation and reverse transcription. Total RNA was extracted from pooled sucrose gradient fractions by the acid guanidinium thiocyanate-phenol-chloroform method (2), without adding carrier RNA or treating with DNase. The RNA was resuspended in 15 μl of 1 mM EDTA. cDNA was synthesized by random-primed or oligo(dT)-primed reverse transcription with the Moloney murine leukemia virus (H-) riboclonal cDNA synthesis system (Promega). The presence of ENTV cDNA was confirmed by *gag*-PCR (3).

Construction of cDNA libraries. *EcoRI* linkers were ligated to the cDNA by using the Riboclonal *EcoRI* adapter ligation system (Promega). Excess adapters were removed from the ligation product with Sephacryl-400 as directed by the manufacturer. All of the cDNA-linker product was ligated to lambda arms by using the λ ZapII-*EcoRI*-calf intestinal alkaline phosphatase cloning kit as described by the manufacturer (Stratagene). Gigapack II Gold packaging extract (Stratagene) was used for packaging.

Recombinant clones were selected initially with JSRV clones as probes (pJS382 [38], pJS107 [38], and pLG1 [26]) and later by using ENTV clones from the library. The selected clones were excised by using plating strain SOLR, which converts lambda ZapII to Bluescript SK(-) plasmid and allows for blue-white screening for inserts.

PCR. Gaps in the sequence were filled by PCR with primers chosen from the available ENTV sequence data. The positions of the PCR fragments relative to the ENTV genome are shown in Fig. 1. In each case, PCR products were pooled from at least four independent reactions, gel purified, and cloned into pGEMT (Promega) or pCR2.1 (Invitrogen) as specified by the manufacturers. Product a was amplified from cDNA made from ENTV RNA prepared as described above. Products b, c, and d were obtained by amplification of ENTV provirus from nasal tumor DNA.

(i) **PCR a.** Product a was amplified from ENTV cDNA with primers ATGG GAGCCCTACAACCTGG (nucleotides [nt] 3113 to 3131 [Fig. 2]) and CAG AGGATACAAGGAGAAACC (reverse of nt 3542 to 3572 [Fig. 2]). The PCR was carried out in 2.5 mM MgCl_2 -10 mM Tris (pH 9.2)-50 mM KCl-20 mM $(\text{NH}_4)_2\text{SO}_4$ -200 μM each deoxynucleoside triphosphate (dNTP)-0.125 μM each

* Corresponding author. Mailing address: Moredun Research Institute, International Research Centre, Pentlands Science Park, Bush Loan, Penicuik, Midlothian EH26 0PZ, United Kingdom. Phone: (0131) 445 5111. Fax: (0131) 445 6111. E-mail: couse@mri.sari.ac.uk.

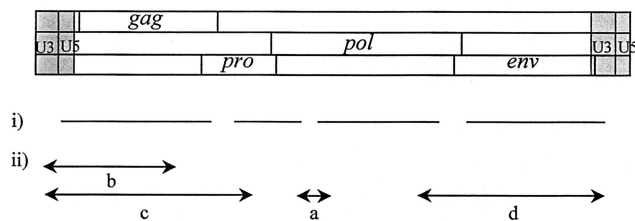


FIG. 1. Schematic diagram of the ENT virus showing (i) sequence derived from cDNA library clones and (ii) PCR products generated to complete the sequencing of ENT virus.

primer-1.25 U of *Taq*. A total of 45 cycles were used, each consisting of 1 min at 94°C, 30 s at 59°C, and 1 min at 72°C.

(ii) **PCR b.** A 250-ng portion of nasal tumor DNA was used as template for the first round of PCR with primers PL1 (TCAGGAAGTCTTAAGAGCTTTTAG) (nt 7267 to 7290 [Fig. 2]), and PG1 (ATACTGCAGCYCGATGGCCAG) (reverse of nt 1802 to 1822 [Fig. 2]). A 2- μ l volume of the product of this PCR was used in the second round of amplification with primers PL2 (GATGTCTTTTG GTTTTGCAACATG) (nt 7297 to 7320 [Fig. 2]) and PG1. The same conditions were used for both the first and second rounds of PCR as follows: 1.5 mM MgCl₂-50 mM KCl-10 mM Tris-HCl (pH 8.3)-100 μ M each dNTP-10 pmol of each primer-1.25 U of *Taq* polymerase (Boehringer). The PCR cycles were 94°C for 4 min, then 30 cycles at 94°C for 40 s, 63°C for 40 s, and 72°C for 90 s plus a progressive increment at 72°C of 10 s per cycle from cycle 11 onward, and a final extension of 72°C for 7 min. A tumor-specific product of approximately 2 kb was confirmed to originate from ENT virus by demonstration of the *gag* *Pst*I site previously shown to be a molecular marker for ENT virus (reference 3 and data not shown).

(iii) **PCR c.** Fragment c was amplified from nasal tumor DNA with primers PL3 (ATGATCTCAAGTTACTTAACCTTGC) (nt 4859 to 4880 [Fig. 2]) and NT4R (CTGTTCTGGATTGGTAGTTGG) (reverse of nt 7376 to 7399 [Fig. 2]). The PCR conditions were 100 ng of template DNA-50 mM KCl-10 mM Tris-HCl (pH 8.3)-2 mM MgCl₂-12.5 pmol of each primer-200 μ M each dNTP-1 U of *Taq* polymerase. A total of 40 cycles of 94°C for 1 min, 59°C for 1 min, and 72°C for 2 min were performed, with a final extension of 5 min at 72°C.

(iv) **PCR d.** Fragment d was amplified with PCR primers NTG4F (CTCATG GCTCCCTTCATCTG) and PL4 (AAGCAAGTTAAGTAACTTGAGATC) with nasal tumor DNA as the template. The PCR conditions were 100 ng of template DNA-50 mM KCl-10 mM Tris-HCl (pH 8.3)-1.5 mM MgCl₂-12.5 pmol of each primer-200 μ M each dNTP-1 U of *Taq* polymerase. A total of 40 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 2.5 min were performed, with a final extension of 5 min at 72°C.

Sequence analysis and phylogeny. Each clone was sequenced in the forward and reverse orientations with a Li-Cor automated sequencer. The clones were ordered initially by alignment with JSRV by using the Genetics Computer Group (GCG) Bestfit program (13). The sequence data were compiled with the GCG fragment assembly package (13). The sequence of ENT virus was confirmed by sequencing at least two overlapping clones. Sequences obtained by PCR amplification were confirmed by sequencing at least three independent clones. As would be expected for sequences of an uncloned virus population, albeit isolated from a single animal, some differences were found. Where differences occurred, the consensus sequence is shown.

The GCG programs Eclustalw and Clustree were used for phylogenetic analysis. Percent divergence figures between all pairs of sequences were calculated and used to generate unrooted phylogenetic trees by the neighbor-joining method of Saitou and Nei (30).

Nucleotide sequence accession numbers. The nucleotide sequence of the complete genome of ovine ENT virus will appear in the EMBL Nucleotide Sequence Database under accession no. Y16627.

RESULTS

A total of 39 clones were selected by screening of the ENT virus cDNA libraries. Sequencing of these clones generated more than 75% of the ENT virus genome. The remaining gaps, as identified by alignment with the JSRV genome, were generated by PCR (Fig. 1). Gap a in the *pol* gene was filled by PCR on cDNA from virus purified from nasal fluid from an animal with ENT. The sequencing gaps remaining at either end of the virus were amplified by PCR from nasal tumor genomic DNA, taking advantage of the fact that the long terminal repeat (LTR) is duplicated in the provirus (product b in Fig. 1). ENT virus-specific U3 primers were designed by selecting regions of dif-

ference between the clones from the library and the published sequences of endogenous sheep retroviruses (1, 26) and used to amplify the region from U3 across U5 into *gag*. The LTR sequence obtained from these clones allowed the identification of a unique region in the ENT virus LTR, from which more efficient exogenous specific PCR primers, PL3 and PL4, were chosen. These primers, together with primers from within *pro* or *pol*, respectively, allowed amplification from tumor DNA of products c and d, which included the two remaining sequence gaps. Control reactions, with kidney DNA from the same animal as the template, verified that the PCRs amplified only exogenous ENT virus and not endogenous sequences. We have previously reported the presence of two restriction enzyme sites, *Pst*I and *Aat*II, within *gag* which are specific for ENT virus (3). These occur at bases 1691 to 1696 and 1757 to 1762, respectively, in the full-length sequence of ENT virus (Fig. 2). No JSRV or endogenous sequences that have these specific sites have been reported. Both these sites were present in all our clones spanning this region, which is further evidence that these clones were not derived from endogenous sequences. Furthermore, none of the clones contained the JSRV-specific *Sca*I site in *gag* (26).

The complete sequence of ovine ENT virus is shown in Fig. 2. The genome is 7,434 nt long. It has a genetic organization typical of type B and D retroviruses in that it has *gag*, *pro*, *pol*, and *env* genes (Fig. 3). No further open reading frames are present in ENT virus. Comparison with other retroviruses suggest that *pro* and *pol* are expressed as *gag-pro* and *gag-pro-pol* precursors generated by ribosomal frameshifting and that *env* transcripts are generated by splicing.

LTR. The LTR of ENT virus is 374 bases long and is bound by the inverted-repeat sequences GCAG (nt 122 to 125) and CTGC (nt 7172 to 7175) and flanked by the primer binding site (nt 127 to 142) and the polypurine tract (nt 7154 to 7170). U3 is 249 bp, R is 13 bp, and U5 is 112 bp. U3 is defined as the region between R and the polypurine tract that contains the initiation site for synthesis of the positive strand of viral DNA. Like JSRV, ENT virus uses tRNA^{Lys}_{1,2} as the primer for reverse transcriptase (RT) to initiate minus-strand synthesis. In the DNA form of retroviral LTRs, the U3 region contains *cis*-acting sequences necessary for viral replication and regulatory signals for retroviral transcription. Although the ENT virus LTR must interact with cellular factors, no recognized consensus sequences for transcription factor binding sites were identifiable. Experiments with expression constructs involving the LTR linked to a reporter gene must be performed before the functional regions of the ENT virus LTR can be dissected.

Leader sequence. The untranslated region between the primer binding site and the start of *gag* is likely to contain the splice donor site for generation of subgenomic (*env*) transcripts, as well as the packaging signal which specifies incorporation of genome RNA into virions. Alignment with JSRV suggests that the splice donor site is after base 189.

Viral proteins. (i) ***gag*.** ENT virus *gag* (nt 256 to 2094) encodes a 613-amino-acid (aa) polypeptide. The boundaries between the Gag proteins were assigned, as for JSRV (38), by alignment with MPMV, squirrel monkey retrovirus (SMRV), and mouse mammary tumor virus (MMTV) sequences. The ENT virus capsid protein (CA) is encoded by nt 1027 to 1677, and the nucleic acid binding protein is encoded by nt 1678 to 1995. There are two sequences of Cys-X₂-Cys-X₂-His-X₄-Cys, separated by 13 residues, in the nucleic acid binding protein. This is typical of a zinc finger domain and is thought to mediate nucleocapsid binding to the genomic RNA.

(ii) ***pro*.** The *pro* open reading frame (nt 1878 to 2858) encodes a 326-residue polypeptide. It is likely to be expressed

GCAAGGTATCAGCCATTTTGGTCTGATCCTCTCAACCCCATCTTTTGGCTCTCTTATTTTCTTAGCGGGATGCTCCGTTCTCTCTTATACAGGTGTAA
 100
 CTTTGTCCGTGCTGGCTGCGGCAGGTGGCGCCCAACGTTGGGCTCGACAGCTTTCTCTGCTACTCTTATTATTGAAACAAAGTGAATTTGAATG
 200
 IR < PBS U5 *gag*> M G Q T H S R Q L F V H M L S
 AATAAGTATTACATTAAGAGGAGTAGTAAGGTATATAGTTGAGAGTATAAATATGGGACAAACACATAGTCGTAGTTGTTGTACACATGTTGTCT
 300
 V M L K H R G I T V S K P K L I N F L S F I E E V C P W F P R E G T
 GTAATGTTAAACATAGGGGAATCACTGTCTCTAAACAAAATTAATCAATTTTCTTTCATTGAGGAAGTTTGTCTTGGTTCCCGAGAGAAGGTA
 400
 V N L E T W K K V G E Q I R T H Y T L H G P E K V P V E T L S F W
 CGGTAAATTTGGAAACATGAAAAAGGTAGGGGAACAAATTCGGACTCATTATACTTTACATGGCCCTGAGAAAGTTCTCTGTTGAAACTTTATCCTTTG
 500
 T L I R D C L D F D N D E L K R L G N L L K Q E E D P L H D P D S
 GACATGATTCGTGTTGTTGGATTTGATAATGATGAATGAAACGTTTAGGAAATTTATTAACAGGAAGAAGACCCCTCTCATGATCCTGACTCG
 600
 E N S Y D A P P P P H P S L H S S R P S D N D D L L S S T D E A E L
 GAAATAGTTATGATGCTCCTCCCTCCTCATCCGCTCACATTCGTCGCTCCCTCAGATAATGATGATTTACTTTTCATCTACAGATGAGGCTGAAT
 700
 D E E A A K Y H Q E D W G F L A Q E K G A L T S K D E L V E C F K
 TAGATGAAGAAGCTGCCAAATACCATCAAGAAGATTGGGGTTTTTTAGCACAGAAAAAGGAGCATTACATCTAAAGATGAATGGTTGAATGCTTTAA
 800
 N L T I A L Q N A G I K L P A N N D K F P S A P P L P P A Y A P S
 AAACCTCACTATTGCTTTACAAAACGAGGAATAAACTTCCTGCTAAATGATAAATTTCTTCTGCTCCGCTCTCTCTCTGCTTATGCCCTTTCT
 900
 V V A G L D P P P G P L P P S K N M S P L Q K A L R Q A Q R L G E V
 GTTGTAGCGGGTCTCGATCCTCCTCCGGGGCCCTTCCACCGCTCAAAAATATGCTCCACTACAAAAGGCATTGAGACAGGCACAGCGACTTGGTGAG
 1000
 V S D F S L A F P V F E N N N Q R Y Y E S L P F K Q L K E A G L T K I A
 TTGTTCTGATTTTCTCTGCTTCCCTGTTTTTGGAAAATAAATAACCGCTTATTATGAATCACTGCTTTTAAACAACATAAAAGAGTTAAAAATTCG
 1100
 C S Q Y G P T A P F T I A M I E N L G T Q A L P P N D W K Q T A R
 TTGCTCACAATACGGTCTACTGCTCCATTTCACATTGCTATGATAGAAAATTTGGGCACTCAGGCTTACCTCCAAATGATTGGAAGCAGACAGTAGG
 1200
 A C L S G G D Y L L W K S E F F E Q C A R I A D V N R Q Q G I Q T S
 GCATGCTTTTCAGGGGAGATATTTATTATGGAATCTGAATTTTTTGAACAATGCTCGTATCGCTGATGTTAACCGACAGCAAGGTATACAGACCT
 1300
 Y E M L I G E G P Y Q A T D T Q L N F L P G A Y A Q I S N A A R Q
 CCTATGAAATGTTGATGGTGAAGCCCTTACCAGGCTACTGATACTCAACTAATTTCTTGCCTGGTGCATATGCGCAAATATCAAATGCGGCTCGGCA
 1400
 A W K K L P S S S T K T E D L S K V R Q G P D E P Y Q D F V A R L
 GGCATGAAAAAATCTCTAGTTCCAGTACCAAGACAGAGGATCTTTCCAAAGTCCGACAGGGACCTGATGAACCTTACCAGGATTTCGTGGCAGCACT
 1500
 L D T I G K I M S D E K A G M V L A K Q L A F E N A N S A C Q A A L
 TTAGATACTATAGGTAAGATAATGTCAGATGAAAAGGCTGGGATGCTACTGGCAAAACAGTTGGCTTTTGAACCGCTAACTCTGCTTGTCAAGCTGCC
 1600
 R P Y R K K G D L S D F I R I C A D I G P S Y M Q G I A M A A A L
 TAAGACCTTATCGAAAAAGGAGATCTGCTGATTTTATTGCGATTTGTGCTGATTTGGACCTCATATATGCAAGGTATCGCTATGGCTGCAGCTTT
 1700
 Q G K S I K E V L F Q Q Q A R N K R G R Q R S G N S G C F V C G Q
 ACAAGAAAAAGCATAAAGAGGTACTCTTCCAACAACAAGCCCGAAACAAGAGGGGACGTCAAAGGTCAGGAAATTCGGTGTGTTTTGTTGTGGCCAG
 1800
 P G H R A A V C P Q K Q Q S P V N T P N L C P R C K K G K H W A R D
 CCTGGACATCGGGCTGCAGTATGCCCTCAAAAACAACAAGTCTGTTAACACTCCTAATTTGTGCCCGCGCTGTAAGAAAGGAAAGCATTGGGCACGAG
 1900
 L P V Q N R C S R Y P F T P G V G K L G E G P A P G P E T M L W G
 C R S K T D V Q G T P L P P V S G N W V R G Q P L A P K Q C Y *
 ATTGCCGTCCAAAACAGATGTTCAAGGTACCCCTTACCCCGGTGTCGGGAAACTGGGTGAGGGGCCAGCCCTGGCCCCGAAACAATGTTATGGGGC
 2000
 N T A G S K R T I A D L C R A T R G S A G L D L C A T S Y T V L T P
 AACACTGCAGGTTCCAAAAGAACCATTCGACAGCTCTGTCGAGCCACAAGAGGCAGCGGGGATTGGACCTCTGTGCCACCTCTACACAGTATTAACCT
 2100
 E M G V Q T L A T G V F G P L P P G T A G L L L L G R S S A S L K G
 CCGAGATGGGGTCCAAACCTTCCACAGGAGTGTGGGCTCTACCTCCAGGGACAGCTGGATTACTTTTGGGGCCAGCAGTGCCTTTAAAGGG
 2200
 I L I H P G V I D S D Y T G E I K I L A S A P N K I I V I N A G Q
 AATACTTATTCCTGGTGTGATTGACTCTGATTAACAGGAGAGATAAAAATATTAGCCTCCGCTCTCAACAAAATATTGTAATCAATGCAGGACAA
 2300
 R I A Q L L L V P L V I Q G K T I N R D R Q D K G F G S S D A Y W V
 CGTATAGCTCAACTCTTTTAGTCCCATAGTCATACAAGGAAAAACAATTAACCGAGACCGTCAAGATAAAGGTTTCGGGTCTCTGACGCCATTTGGG
 2400
 Q N V T E A R P E L E L R I N G K L F R G V L D T G A D I S V I S
 TACAAAATGTTACCGAGGCAGCAGCAACTTACGCTACGCATTAATGGCAAGCTTTTCCGTGAGTACTTGATACAGGGCCGATATTAGTGTATTTTC
 2500
 D K Y W P T T W P K Q M A I S T L Q G I G Q T T N P E Q S S S L L
 TGATAAATATTGGCTACTACATGGCCAAACAGATGGCTATTCTACTCTCCAGGGTATTGGCCAAACTACCAATCCAGAACAGAGTTTATCCCTCTTT
 2600
 T W K D K D G H T G Q F K P Y I L P H L P V N L W G R D I L S K M G
 ACTTGAAGGATAAAGATGGACATACAGGCCAATTTAAACCTTATATTGTCCTCCTTCCAGTAACTATGTTGGGGCGTGTATATTAAGCAAAATGG
 2700
 V Y L Y S P S P T V T D L M L D Q G L L P N Q G L G K Q H Q G I I
 GTGTTTATTATATAGTCCCTTCCCACTGTGACAGATTTGATGTTAGACCAAGGCTTACTTCCAAATCAAGGTTTAGGTAAACAACATCAAGGCATCAT
 2800
 L P L D L K S N Q D R K G L G C F S L *
 TTTACCCTTGATTTAAATCTAATCAAGATCGAAAAGGCTTGGGGTGTTCCTTAGGGACCTGATTTCTCTGTGACGCAGCCGATCCTATTGATT
 2900

FIG. 2. Complete nucleotide sequence of the ENTV genome together with the deduced amino acid sequences of the open reading frames. The limits of U3, R, and U5 are indicated by arrows. The inverted repeats (IR), primer binding site (PBS), and polypurine tract (PPT) are indicated by underlines. The consensus protease cleavage site for *env* SU and TM is indicated by V.

K S E E P V W V D Q W P L T Q E K L S A A Q Q L V Q E Q L R L G H
GGAAATCTGAGGAACCGGTGGGGTGCATCAGTGGCCCTAACACAGGAAAACTTTCTGCCGCACACAGCTGGTACAGGAACAGCTGAGACTTGGGCA 3000
I E P S T S A W N S P I F V I K K K S G K W R L L Q D L R K V N E
TATTGAACCTTCTACCTCTGCCTGGAATCCCCAATTTTGTATTAAAAAGAACTGGAAATGGAGATTATTACAAGATCTTCGCAAGGTAATGAA 3100
T M M H M G A L Q P G L P T P S A I P D K S Y I I V I D L K D C F Y
ACAATGATGCATATGGGAGCCCTACAACCTGGGTGGCCACTCCTTCTGTATACCTGATAAATCTTATATTATTGTTATAGATTTAAAAGATTGTTTT 3200
T I P L A P Q D C K R F A F S V P S V N F K E P M K R Y Q W K V L
ACACTATTCCACTTGCACCTCAAGATTGCAACAGATTGCTTTTTAGTGTGCCCTCTGTTAATTTCAAGGAACCTATGAAACGATATCAATGGAAGTCT 3300
P Q G M T N S P T L C Q K L V A T A I A P V R Q R F P Q I Y L V H
TCCACAAGGAATGACTAATAGCCCTACGCTGTGCAAAAATAGTTGCTACAGCAATAGCTCCTGTTCGACAGCGTTTTCCCTCAGATATATTGGTTCCAC 3400
Y M D D I L L A H T D E H L L Y Q A F S I L K Q H L S L N G L V I A
TATATGGATGATATATTACTAGCTCAGCTGATGAACATTTATTATATCAAGCTTTTTCTATTCTAAAACAGCATTTAAGTCTTAAACGGTCTTGTGATTG 3500
D E K I Q T H F P Y N Y L G F S L Y P R T Y N T Q L V K L Q T D H
CTGATGAAAAATCAAACGCATTTCCCTTATAATTATTAGTTCTCCCTGTATCCTCGTACCTATAACACCACTAGTAAAGCTGCAACCTGATCA 3600
L K T L N D F Q K L L G D I N W I R P Y L K L P T Y T L Q P L F D
TTTGAAAACCTCAAAATGATTTTCAAGAACTTTTAGGAGATATAAAGCTGATCCGCTCCTATTTTAAAGTTGCCCACTTTTAAACCCATTATTGAC 3700
I L K G D S D P A S P R T L S L E G Q T A L Q L I E E A I R R Q Q I
ATCCTTAAAGGTGACTCTGATCCTGCGTACCCCGAACACTTTCTTTAGAAGGACAAAACCTGCTTTTACAATTAATAGAAGAAGCTATCAGGCGACAACAGA 3800
T Y C D Y Q R S W G L Y I L P T P R A P T G V L Y Q D K P L R W I
TTACTTATTGTGATTATCAACGATCATGGGATTTATATATACTTCCCACTCCCGAGCACCTACAGGAGTACTCTATCAAGATAAACCTTTACCGTGGAT 3900
Y L S A T P T K H L L P Y Y E L V A K I V A K G R H E A I Q Y F G
ATACTTATCTGCTACTCCAACTAAACACTACTTCTTATTATGAATCTGTGCAAAAATGTAGCAAGGGACGACATGAGGCCATCCAATATTTTGGT 4000
M E P P F I C V P Y A L E Q Q D W L F Q F S D N W A I A F A N Y P G
ATGGAGCCTCCCTTTATCTGTCTTCTATGCTTTAGAACAAACAGGATTGGCTTTTTCAATTTTCAAGATAATGGGCCATAGCATTTGCGAATTACCCCG 4100
R I I H H Y P P D K L L Q F A S S S H A F I F P K I V R R Q P I P N
GACCAATATTATACCTTACCTGATAAATATTGCAATTTAGCTAGCTCTCATGCTTTTATTTTTCCAAAATAGTTCGCGGACAACCTATTCCCTAA 4200
A T L I F T D G S S N G T A A L I I N H Q T Y Y T Q T S F S S A Q
CGCTACCTTATATTACAGAGGGCTCTTCCAATGGGACTGCAGCCTTAATTATTAATCATCAAACTTATATACACAACAGTTCCTCTTCTGCTCAA 4300
V V E L F A V H Q A L L T V P T S F N L F T D S S Y V V G A L Q M I
GTTGGAATTTATGCACTCCCAAGCATTGCTAACAGTACCTACTTCTTCAATTTATTTACAGACAGCTCCACGTTGGTGGTGGTCCCTTACAAATGA 4400
E T V P I I G T T S P E V F N L F T L I Q Q V L H C R Q Y P C F P
TTGAGACTGTCCAATTATCGGCACCACTCTCCCTGAAGTTTTAACTTATTTACATTGATTCAACAGGTTCTACATTGCCCTCAATATCCCTGTCTTCT 4500
G H I R A H S T L P G A L V R G N H T A D A L T K Q V F F Q S A I
TGGACATATTCTGTCACACTCCACCCCTTCTGGTGGCTTGTACGAGGCAATCAGCTGCGGAGCGCTTACCAAAACAGTGTTTTTTCAATCAGCTATT 4600
D A A R K S H D L H H Q N S H S L R L Q F K I S R E A A R Q I V K S
GATGCAGCCCGAAAGTCCCATGATTTACATCAAAAATAGTCACTTCTTACGTTTGCAATTTAAAATTTCCCGTGAGGCTGCACGGCAAATTTGTAAT 4700
C S T C P Q F F V L P Q Y G V N P R R G L R P N H L W Q T D V T H I
CTTGTCTACTTGTCTCAATCTTGTCTTCCCAATATGGTGCAACCCCTGAGGTTACGCCCTAATCATCTTGGCAAAACAGCTTATCAGCAT 4800
P Q F G R L K Y V H V S I D T F S N F L M A S L H T G E S T R H C
TCCTCAATTTGGGCGTCAAAATATGTTCAATGTTTCAATCGACACCTTTTCCAATTTCTCATGGCTTCCCTTACACTGGAGAATCAACAGCTCACTGT 4900
I Q H L L F C F S T S G I P Q T L K T D N G P G Y T S R S F Q R F C
ATTCAACATTTGCTGTTTTGCTTTTCCACTTCAAGTATCCCAAAAACCTTAAAACAGATAATGGACCTGGCTATACCAGCGCTCTTTTCAACGCTTTT 5000
L S F Q I Q H K T G I P Y N P Q G Q G I V E R A H Q R L K H Q L L
GCCTTTCTTTCAAATTCAGCATAAAAACAGGAATTCGGTATAATCCTCAGGGACAAGGCATCGTGGAAAGAGCCCATCAACCTCTTAAACATCAATATT 5100
K Q K K G N E L Y S P S P H N V L N H A L Y V L N F L T L D A E G
AAAACAAAAAGGGAAATGAAGTATAGCCCTCACCACATAATGCTTAAATCATGCTCTTTACGTTTTAAATTTTTTAACTTTAGACGCAGAAGGC 5200
N S A A Q R F W G E R S S C K K P L V R W K D P L T N L W Y G P D P
AATTCAACAGCCAGCGTTTTTGGGGAGAACGATCCTCATGCAAAAACCCTTGTACGATGGAAAGATCCACTTACCAATCTGTGGTATGGCCCTGACC 5300
env M P K H R A G F R K G W Y A R Q R N S
V L I W G R G H V C V F P Q D A E A P R W I P E R L V R A A E E L
CTGATTAATATGGGACGAGGCATGTTTGTGTTTTTCCACAGGATGCCGAAGCACCGCGTGGATTCCGGAAGGCTGGTACGCGCGGACAGGAACT 5400
L T Y Q M Q R M T L N E T T N G L P T Q R Q V E A L M R H A W N E
P D I S N A T H D I E *
CCCTGACATCAAATGCAACGATGACATTGAAAGGAGACCAAGTGGACTACCTACCCAGAGACAAGTTGAAGCGCTGATCGCTCATGCTTGGAAATGA 5500
A H V Q P P A T P I K I L I M L L L L L Q R I Q N G A A A A F W A
GGCTCATGTACAACCCAGGACACCTTATAAATATTGATTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5600
Y I P D P P M I Q S L G W D R E I V P V Y V N D T S L L G G K S D I
TATATCCCAGCCCGCTATGATTCAATCCTTAGGATGGGATAGAGAAAATAGTACCTGTATATGTCAATGATACTAGCTGTAGGGGGAAAATCAGATA 5700
H I S P Q Q A N I S F Y G L T T Q Y P M C F S Y Q S Q H P H C I Q
TTCACATCTCTCCTCAGCAAGCAATATTCTTTTACGGTCTCACTACTCAATACCTATGTGCTTTCTTATCAATCCCGAGCATCTCATTGTATACA 5800

FIG. 2—Continued.

```

V S A D I S Y P R V T I S G I D E K T G K R S Y R N G T G P L D I
GGTGTCCAGTGATATATCTATCCCTCCGAGTGACTATTTTCAGGTATTGATGAAAAAACTGGAAAAAGATCATACCCTAATGGGACCCGACCTCTCGACATT
5900
P F C D K N L S I S I G I D T P W T L C R A R I A S V Y N I N N A N
CCGTTTGTGACAAAAATTAAGTATCAGCATAGGAATAGACACTCCTTGGACTTTATGTCGAGCACGAATCGCATCAGTATATAACATTAACAATGGCA
6000
T T L L W D W A P G G T P D F P E Y R G Q H P P I L S V N T A P I
ACACCACCTCTGTGGGACTGGGCACCTGGAGGAACACCTGATTTCCCTGAATATCGAGGACAGCATCCACCCTCTCTCTGTAAACACTGCTCCCTAT
6100
F Q T E L W K L L A A F G H G N S L Y L Q P N I S G S K Y G D V G
ATTTCAAACGAACGTGGAAACTTCTGGCTGCTTTTGGTCATGGCAACAGTCTATATCTACAGCCCAACATTAGTGGGAAGCAAATATGGTGATGTGGGA
6200
V T G F L Y P R A C V P Y P F M L I Q G H M E I T L S L N I Y H L N
GTTACAGGATTTTATATCCCGAGCTTGTGTTCTTATCCTTTATGTTGATACAGGGTCATATGGAATAACACTATCTTTAAACATCTATCATTTAA
6300
C S N C I L T N C I R G V A K G E Q V I I V K Q P A F V M L P V E
ATTTGTTCAATTGTATACTACTAATTCATAGAGGTGTGGCCAAAGGAGAACAAGTTATAATAGTAAAAACAACCTGCTTTTGTAAATGCTGCCTGTTGA
6400
I T E E W Y D E T A L E L L Q R I N T A L S R P K R G L S L I I L
AATAACTGAAGAATGGTATGATGAAACTGCCTTAGAATTTGTTACAACGTATCAATACAGCGCTTAGCCGCTCTAAAAGAGGTTTAAAGCCTAATTATATTA
6500
G I V S L I T L I A T A V T A S V S L A Q S I Q A A H T V D S L S Y
GGTATCGTATCTTAAATCACCTCATAGCGACTGCTGTACCGCTTCTGTATCTTTAGCGCAATCCATTCAAGCTGCTCATACTGTGGATTCTTGTTCAT
6600
N V T K V M G T Q E D I D K K I E D R L S A L Y D V V R V L G E Q
ACAATGTTACTAAAGTAAATGGGAACTCAAGAAGATATAGATAAAAAAATAGAAGATAGATTATCAGCTTTATATGATGTGGTAAGAGTTTTGGGGGAACA
6700
V Q S I N F R M K I Q C H A N Y K W I C V T K K P Y N T S D Y P W
GGTTCAAAGCATTAATTTTCGTATGAAGATCAATGTCATGCTAATATAAATGGATTGTGTCAAAAAAGCCCTACAATACCTCTGATTATCCGTGG
6800
D K V E K H L Q G I W F N T N V S L D L L Q L H N E I L D I E N S P
GATAAGGTAGAAAAACATCTGCAAGGAATTTGGTTTAACTAATGTTTCCCTGGATCTTTTACAATTCGATAAATACTCGACATTGAAAATTCAC
6900
K A T L N I A D T V N N F L Q N L P S N F P S L H S L W Q S I I V
CAAAAGCTACTTTGAATATAGCTGATCTGCAATAATTTTACAAAACTCTTCTCTAACTTCCCTAGCCTCCATTCATTGGCAGAGCATTATGTT
7000
I A I I V F I V I V V I L L V P C L T R G L I K D F L Q M R I E L
TATAGCTATAATGTTGTTTATGTAATGTTGTAATCTTTTGGTTCCTTGCCTCACTCGAGGTTTAAATTAAGATTTCCTACAAATGAGAATGAAATTA
7100
I H M K Y R Y M L Q Y Q K L M D F V K K R E G S C G G Q P A E G *
ATACATATGAAATACAGATACATGTTGCAATACCAAAAACCTTATGGATTTTGTGAAAAAAGAGAGGGAGCTGCGGGGGACAACCTGCGGAGGTTAAG
7200
TCCTGGGAGCTCCTTGGCAAAATGCCAAGGCTTAGGCAAGTACCTAAGCTCCCTGTCCCGCCACCCTCAGGAAGTCTTAAAGAGCTTTTGGAGCTCAAAATG
7300
TGTTTGGTTTGTCAACATGGCTTAAAAAGCAGGAAATCTGATTATGTAAGAAACCAGTAATTATGTAAGTAAATGATCTCAAGTTACTTAACTTGCTTA
7400
TATATACTGCTGCACAATAAAGCAAGGTATCAGC
TATA Poly-A< 7434
    
```

FIG. 2—Continued.

as a *gag-pro* fusion polypeptide generated by ribosomal frameshifting, as has been shown for MMTV (17). The exact site and relative efficiency of this frameshifting event remain to be elucidated. The protein encoded by *pro* is thought to comprise two domains, a pseudoprotease (or protease-like) domain which exhibits a dUTPase activity, as has been suggested for other retroviruses by both computer analysis (22) and functional studies (9), and an active protease (Pr). The sequence Leu-Asp-Thr-Gly, indicative of the core amino acid sequence of cellular aspartyl protease, is encoded by nt 2469 to 2480.

(iii) *pol*. The *pol* open reading frame (nt 2825 to 5437) encodes a peptide of 870 aa giving a predicted molecular mass of 99 kDa. *pol* is likely to be translated as a fused *gag-pro-pol* polypeptide, again generated by ribosomal frameshifting. *pol* encodes the RT and integrase domains. The putative RT active

site is the Tyr-Met-Asp-Asp sequence encoded by nt 3401 to 3412 in ENTV.

(iv) *env*. The *env* open reading frame (nt 5346 to 7196), at its start, overlaps *pol*, and the *env* transcript is likely to be generated by splicing. Like JSRV, ENTV contains two possible ATG start codons 7 aa apart, but the second is thought to be the initiation codon used since it is preceded by a predicted splice acceptor sequence. Retroviral *env* open reading frames encode a polypeptide precursor which is cleaved twice, first following the peptide signal sequence and second following an internal proteolytic cleavage site, which generates the surface (SU) and transmembrane (TM) envelope proteins. In ENTV, *env* encodes a polypeptide precursor of 617 aa (predicted molecular mass, 68 kDa). A hydrophobic region (encoded by nt 5526 to 5573) is thought to be the peptide signal sequence, and cleav-

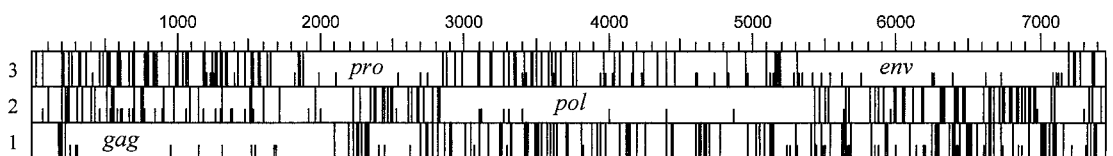


FIG. 3. Genetic organization of ENTV. Stop codons in the three reading frames are represented by long vertical bars; methionine residues are represented by short vertical bars.

TABLE 1. Percentage nucleotide identity and amino acid identity and similarity between ENTV and JSRV.

Region	Nucleic acid similarity (%)	Amino acid similarity (%)	Amino acid identity (%)
<i>gag</i>	89.7	96.7	95.0
<i>pro</i>	95.2	97.2	96.3
<i>pol</i>	92.3	98.0	96.9
<i>env</i>	85.5	95.2	89.0
SU	87.9	96.3	93.4
TM	81.9	93.2	81.9
LTR	75.7		
U5	82.9		
R	100		
U3	72.6		

age probably occurs cotranslationally at the carboxyl-terminal end of this sequence. A consensus proteolytic cleavage site RPKR (encoded by nt 6468 to 6479) generates a SU protein of 302 aa and a TM protein of 239 aa. A 21-aa hydrophobic region (encoded by nt 6994 to 7046) is likely to be the trans-membrane domain that anchors the TM protein in the membrane. This gives a predicted cytoplasmic domain of 50 aa and an extracellular domain of 182 aa.

The predicted molecular mass of the *env* proteins from the sequence data alone are 34 kDa for SU and 27 kDa for TM. However, the actual molecular masses will be larger due to glycosylation. The motif Asn-X-Ser/Thr, which can serve as a site for N-linked glycosylation, is found nine times in SU and twice in TM.

Comparison with JSRV. As shown in Table 1, the genes of ENTV and JSRV are very similar, with an overall amino acid similarity of greater than 95%. The biggest differences occur at the 3' end of *env* and in the U3 LTR region. In addition, the open reading frame ORFX postulated for JSRV contains two stop codons in ENTV and therefore does not appear to be an open reading frame in ENTV.

The 3' end of *env* is the most divergent of the coding regions of JSRV and ENTV. The viruses are 96% identical in SU and 93% identical in TM at the amino acid level. The TM protein can be further subdivided into external, transmembrane, and intracytoplasmic domains, which exhibit 98, 81, and 75% amino acid similarity, respectively (86, 80, and 63% nucleotide identity). Sequencing of six independent PCR-derived clones and two cDNA clones across this site confirmed that there is nucleotide sequence divergence and not a frameshift mutation at the 3' end of *env*. Many of the amino acid substitutions are

conservative, suggesting that there are some functional restraints on mutation of this protein.

Alignment of the LTR of ENTV with that of JSRV shows that they are quite similar (75.7% similarity), although JSRV has an insertion of 21 bp in U3 compared to ENTV. Several regions contain a cluster of base differences. These regions are also the most divergent from SERV sequences and for this reason were chosen for PCR primers. Interestingly, ENTV does not appear to contain the same regulatory sequences as JSRV. For example, the NF-1 site (TGGCA₄GCCA) seen in JSRV (38) (Fig. 4, bases 44 to 55) is interrupted by the insertion of a T in ENTV (TGGCA₄TGCCA). Also, the 19-bp tandem repeat identified in JSRV (38) is shorter in ENTV (15 bp). Finally, the C/EBP sequence (Fig. 4, bases 211 to 218) is not found in ENTV.

Comparison with other retroviruses. Construction of a phylogenetic tree based on comparison of *gag* genes of a range of different retroviruses (Fig. 5A) shows that ENTV and JSRV group most closely with MPMV. However, if *env* sequences are compared (Fig. 5B), ENTV and JSRV are most similar to MMTV. (The tree shown is for SU, but comparison of *env* TM regions generates a very similar tree.) These results confirm that ENTV has a type B envelope and a type D capsid.

Construction of a phylogenetic tree based on LTR sequences (data not shown) shows that ENTV is most closely related to JSRV and SERVs. These have only approximately 43% identity to the type D retroviruses examined (simian SRV-1 type D retrovirus [SRV-1] and MPMV) and less than 33% identity to any of the other retroviruses listed in Fig. 5. Although this degree of identity is in the same range as that for comparison between the *gag* and *env* genes compared above, the LTR sequences from diverse retroviruses do not generate a reliable alignment and do not aid the identification of functional regions of the LTR.

DISCUSSION

This is the first report of the full sequence of ENTV. Determination of the sequence has enabled comparisons to be made between ENTV and other viruses. This confirms that ENTV is a chimeric typeB/D retrovirus and that it is distinct from the related virus JSRV and from the endogenous retroviral sequences found in the sheep genome. Work currently under way suggests that ovine ENTV is also distinct from caprine ENTV (25a). It is hoped that similarities and differences between members of this family of viruses may give an insight into the mechanism of cellular transformation and factors determining cellular tropism.

Although ENTV and JSRV are distinct viruses associated

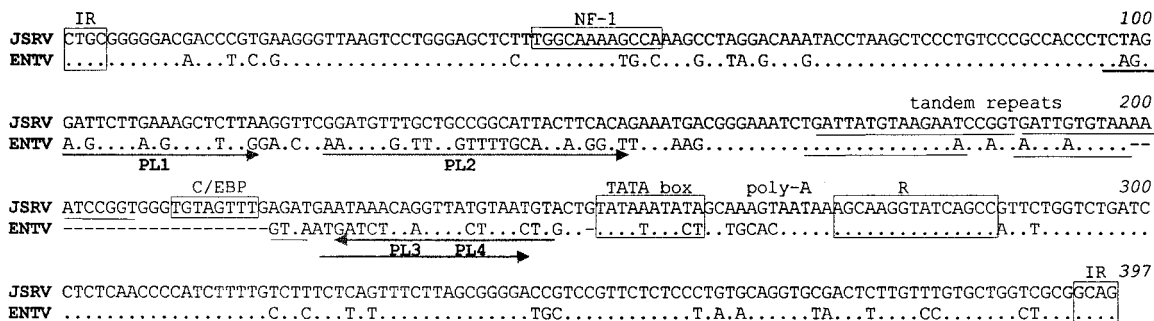


FIG. 4. Alignment of the LTR sequences of ENTV and JSRV. The JSRV sequence shown is that of York et al. (38). Sequences of the LTR primers are underlined. Gaps in alignment are shown by dashes; dots represents identity.

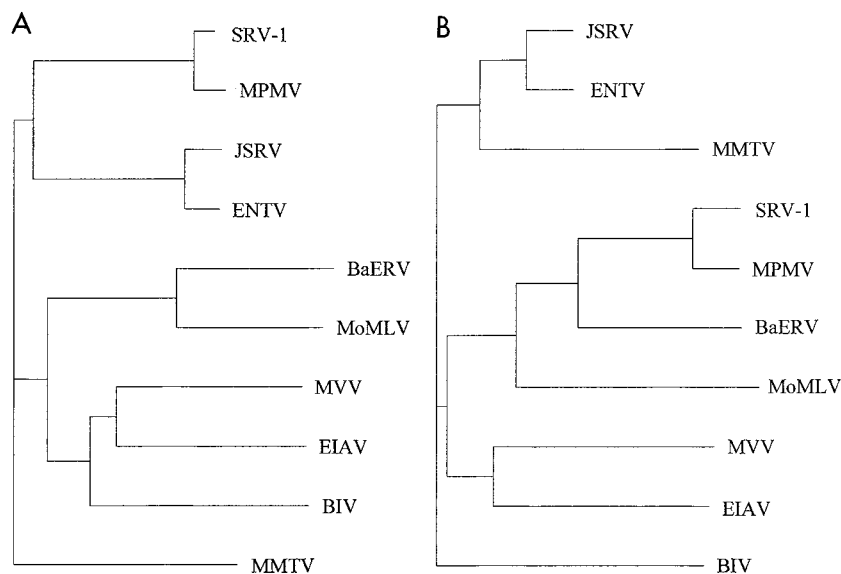


FIG. 5. Phylogenetic trees of (A) *gag* and (B) *env* SU nucleotide sequences of ENTV, baboon endogenous retrovirus (BaERV) (18), bovine immunodeficiency virus (BIV) (11), equine infectious anemia virus (EIAV) (19), JSRV (38), MMTV (24), simian SRV-1 type D retrovirus (SRV-1) (27), MPMV (33), maedi-visna virus (MVV) (28), and Moloney murine leukemia virus (MoMLV) (32). The sequences were aligned by using GCG Eclustalw, and unrooted trees were constructed by the neighbor-joining method of Saitou and Nei (30) (Clustree). Horizontal branch lengths are proportional to evolutionary distances as estimated by percent divergence.

with discrete pathologic findings, translation and alignment of the open reading frames suggests a very high degree of homology between them (Table 1). The similarity between the coding regions of these viruses is in agreement with the evidence that antisera against CA proteins of ENTV and JSRV are cross-reactive. Antisera against other proteins have not been tested, but the degree of sequence similarity between JSRV and ENTV genes suggests that the proteins will all be antigenically cross-reactive.

Results for the few regions of SERVs that have been sequenced suggest that there is also a high degree of similarity between any putative expressed endogenous sequences and ENTV or JSRV. Although these JSRV and ENTV related endogenous sequences have not been linked to any pathologic findings, they may play an important role in allowing these viruses to persist in the host and eventually induce tumours. By reverse transcription-PCR, SERV transcripts have been detected in all sheep tissues tested, but protein has not been detected, probably due to the limited sensitivity of detection methods for proteins compared to nucleic acids. If these SERV sequences are expressed during ontogeny, it is likely that sheep would become immunotolerant to them and therefore to the related exogenous viral proteins, as has been hypothesized previously (25, 38). The surprisingly small number of sequence variations observed in our clones from the population of ENTVs in a single case of ENT supports the hypothesis that this virus is not subject to immune pressure. The earlier sequence data from two regions of ENTV *gag* (3, 29) also contains few changes from the ENTV sequence presented in this paper.

A putative open reading frame named ORFX was described in JSRV (29), and it has been suggested that this may encode an oncogenic product. In ENTV, the region homologous to ORFX contains two stop codons and therefore is not an open reading frame. Since it is likely that these two viruses have the same oncogenic mechanism, we therefore argue against ORFX playing any role in cellular transformation.

Apart from ORFX, the main differences between ENTV

and JSRV occur in the LTR and the *env* TM subunit, suggesting that one or both of these regions must be important in determining viral tropism. The role of the cytoplasmic domain of *env* is thought to be important in envelope processing (21) and in ensuring the incorporation of the envelope glycoproteins into the budding virus (8, 12). It is the SU region of the retroviral envelope that interacts with the cell surface receptor and therefore is largely responsible for directing cell tropism at the surface binding level. Therefore, it is surprising to find that it is the TM and intracytoplasmic domains of *env* which differ the most between these two viruses. The change in the 3' end of *env* could be of little significance compared to that of U3 and may be a by-product of a single recombinational event in the 3' end of the ancestral type B/D retrovirus, which led to the divergence of JSRV and ENTV. However, it has been shown that alterations in the TM cytoplasmic domain can alter the pathogenicity of retroviruses (20, 39), and so a role for this domain in affecting the pathogenicity of ENTV and JSRV cannot be ruled out.

While a number of transcription factor binding sites are identifiable in JSRV, they could not be shown in the ENTV LTR. The differences in the LTRs of these two viruses are likely to affect the replication potential and therefore the pathogenicity of these viruses and may reflect the presence of different transcription factors in the cells that they infect. Sequences in the LTR that regulate transcription in specific cell types may also contribute to the ability of the retrovirus to transform these same cell types, as has been shown for murine leukemia virus (34). Thus, an understanding of transcriptional control of ENTV may also contribute to elucidating the mechanism of viral oncogenesis. We are currently characterizing the LTR function of these viruses.

In conclusion, we have presented, for the first time, the complete sequence of ENTV. Analysis of the sequence places this virus as another member of a group of B/D-type retroviruses, which also includes JSRV and SERVs. The disease pathology associated with ENTV has some similarities to that associated with JSRV in that both are tumors derived from

secretory epithelial cells. However, the diseases are clearly independent and affect different target organs. A virus species is defined as "a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche" (International Committee on the Taxonomy of Viruses). Thus, it can be argued that JSRV and ENTV are distinct viral species.

The main sequence differences between JSRV and ENTV reside in ORFX, the U3 LTR, and the 3' end of *env*. Whether it is these differences and/or more subtle sequence differences that direct the tropism of these viruses requires further investigation.

The availability of the sequence of ENTV is a significant step that opens the way for the development of ENTV-specific detection techniques, which are vital to the further study of the distribution and pathogenesis of this virus and to the development of strategies to eradicate ENT.

ACKNOWLEDGMENTS

This work was supported by CEC contract AIR3-CT94-084, BBSRC grant RO6361, SOAFED grant ROAME MRI/027/95, and Spanish grant CICYT AGF96-0535-C02-01.

We thank I. D. Bennett for the sequencing and P. Harrison for critical discussion of the manuscript.

REFERENCES

- Bai, J., R. Y. Zhu, K. Stedman, C. Cousens, J. Carlson, J. M. Sharp, and J. C. DeMartini. 1996. Unique long terminal repeat U3 sequences distinguish exogenous Jaagsiekte sheep retroviruses associated with ovine pulmonary carcinoma from endogenous loci in the sheep genome. *J. Virol.* **70**: 3159–3168.
- Chomczynski, P., and N. Sacchi. 1987. Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156–159.
- Cousens, C., E. Minguijón, M. García, L. M. Ferrer, R. G. Dalziel, M. Palmarini, M. De las Heras, and J. M. Sharp. 1996. PCR-based detection and partial characterization of a retrovirus associated with contagious intranasal tumors of sheep and goats. *J. Virol.* **70**:7580–7583.
- De las Heras, M., J. A. García de Jalon, and J. M. Sharp. 1991. Pathology of enzootic intranasal tumour in thirty-eight goats. *Vet. Pathol.* **28**:474.
- De las Heras, M., J. M. Sharp, J. A. García de Jalon, and P. Dewar. 1991. Enzootic nasal tumour of goats: demonstration of a type D-related retrovirus in nasal fluids and tumours. *J. Gen. Virol.* **72**:2533–2535.
- De las Heras, M., J. M. Sharp, L. M. Ferrer, J. A. García de Jalon, and L. M. Cebrian. 1993. Evidence for a type D-like retrovirus in enzootic nasal tumour of sheep. *Vet. Rec.* **132**:441.
- De las Heras, M., E. Minguijón, L. M. Ferrer, A. Ortín, P. Dewar, L. M. Cebrian, Z. Pascual, L. García, J. A. García de Jalón, and J. M. Sharp. 1998. Naturally occurring enzootic nasal tumours of sheep in Spain: pathology and associated retrovirus. *Eur. J. Vet. Pathol.* **4**:1–5.
- Dubay, J. W., S. J. Roberts, B. Brody, and E. Hunter. 1992. Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity. *J. Virol.* **66**:4748–4765.
- Elder, J. H., D. L. Lerner, C. S. Hasselkus-Light, D. R. Fontenot, E. Hunter, P. A. Luciw, R. C. Montelaro, and T. R. Phillips. 1992. Distinct subsets of retroviruses encode dUTPase. *J. Virol.* **66**:1791–1794.
- Fan, H. 1994. Retroviruses and their role in cancer, p. 313–362. *In* J. A. Levy (ed.), *The Retroviridae*, vol. 3. Academic Press, Inc., New York, N.Y.
- Garvey, K. J., M. S. Oberste, J. E. Elser, M. J. Braun, and M. A. Gonda. 1990. Nucleotide sequence and genome organization of biologically active proviruses of the bovine immunodeficiency-like virus. *Virology* **175**:391–409.
- Gebhardt, A., J. V. Bosch, A. Ziemiecki, and R. R. Friis. 1984. Rous sarcoma virus p19 and gp35 can be chemically cross-linked to high molecular weight complexes. An insight into virus assembly. *J. Mol. Biol.* **174**:297–317.
- Genetics Computer Group. 1994. Program manual for the Wisconsin package, version 8. GCG, Inc., Madison, Wis.
- Hecht, S. J., J. O. Carlson, and J. C. DeMartini. 1994. Analysis of a type D retroviral capsid gene expressed in ovine pulmonary carcinoma and present in both affected and unaffected sheep genomes. *Virology* **202**:480–484.
- Hecht, J. S., J. M. Sharp, and J. C. DeMartini. 1996. Retroviral aetiopathogenesis of ovine pulmonary carcinoma: a critical appraisal. *Br. Vet. J.* **152**: 395–409.
- Herring, A. J., J. M. Sharp, F. M. Scott, and K. W. Angus. 1983. Further evidence for a retrovirus as the aetiological agent of sheep pulmonary adenomatosis (Jaagsiekte). *Vet. Microbiol.* **8**:237–249.
- Jacks, T., K. Townsley, H. E. Varnus, and J. Majors. 1987. Two efficient ribosomal frameshifting events are required for synthesis of MMTV gag-related polyproteins. *Proc. Natl. Acad. Sci. USA* **84**:4298–4302.
- Kato, S., K. Matsuo, N. Nishimura, N. Takahashi, and T. Takano. 1987. The entire nucleotide sequence of baboon endogenous virus DNA: A chimeric genome structure of murine type C and type D retrovirus. *Jpn. J. Genet.* **62**:127–137.
- Kawakami, T., L. Sherman, J. Dahlberg, A. Gazit, A. Yaniv, S. R. Tronick, and S. A. Aaronson. 1987. Nucleotide sequence analysis of equine infectious anemia virus proviral DNA. *Virology* **158**:300–312.
- Kodama, T., D. P. Wooley, Y. M. Naidu, H. W. Kestler, M. D. Daniel, Y. Li, and R. C. Desrosiers. 1989. Significance of premature stop codons in *env* of simian immunodeficiency virus. *J. Virol.* **63**:4709–4714.
- Kowalski, M., J. Potz, L. Basiripour, T. Dorfman, W. C. Goh, E. Terwilliger, A. Dayton, C. Rosen, W. Haseltine, and J. Sodroski. 1987. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. *Science* **237**:1351–1355.
- McGeoch, D. J. 1990. Protein sequence comparisons show that the "pseudoproteases encoded by the poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase family. *Nucleic Acids Res.* **18**:4105–4100.
- McKinnon, A. O., J. Thorsen, M. A. Hayes, and C. R. Misener. 1982. Enzootic nasal adenocarcinoma of sheep in Canada. *Can. Vet. J.* **23**:88–94.
- Moore, R., M. Dixon, R. E. Smith, G. Peters, and C. Dickson. 1987. Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift suppression events are required for translation of *gag* and *pol*. *J. Virol.* **61**:480–490.
- Ortín, A., E. Minguijón, P. Dewar, M. García, L. M. Ferrer, M. Palmarini, L. Gonzalez, J. M. Sharp, and M. De las Heras. 1998. Lack of a specific immune response against a recombinant capsid protein of Jaagsiekte sheep retrovirus in sheep and goats naturally affected by enzootic nasal tumour sheep pulmonary adenomatosis. *Vet. Immunol. Immunopathol.* **61**:229–237.
- Ortín, A. Unpublished data.
- Palmarini, M., C. Cousens, R. G. Dalziel, J. Bai, J. C. DeMartini, and J. M. Sharp. 1996. The exogenous form of jaagsiekte retrovirus (JSRV) is specifically associated with a contagious lung cancer of sheep. *J. Virol.* **70**:1618–1623.
- Power, M. D., P. A. Marx, M. L. Bryant, M. B. Gardner, P. J. Barr, and P. A. Luciw. 1986. Nucleotide sequence of SRV-1, a type D simian acquired immune deficiency syndrome retrovirus. *Science* **231**:1567–1572.
- Querat, G., G. Audoly, P. Sonigo, and R. Vigne. 1990. Nucleotide sequence analysis of SA-OMVV, a visna-related ovine lentivirus: phylogenetic history of lentiviruses. *Virology* **175**:434–447.
- Rosati, S., J. Kwang, F. Tolari, and J. Keen. 1996. Characterisation of enzootic nasal tumor virus capsid antigen. *Vet. Microbiol.* **53**:261–269.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
- Sharp, J. M., and A. J. Herring. 1983. Sheep pulmonary adenomatosis: demonstration of a protein which cross-reacts with the major core proteins of Mason-Pfizer monkey virus and mouse mammary tumour virus. *J. Gen. Virol.* **64**:2323–2327.
- Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe. 1981. Nucleotide sequence of Moloney murine leukaemia virus. *Nature* **293**:543–548.
- Sonigo, P., C. S. Barker, E. Hunter, and S. Wain-Hobson. 1986. Nucleotide sequence of Mason-Pfizer monkey virus: an immunosuppressive D-type retrovirus. *Cell* **45**:375–385.
- Tupper, J. C., H. Chen, E. F. Hays, G. C. Bristol, and F. K. Yoshimura. 1992. Contributions to transcriptional activity and to viral leukemogenicity made by sequences within and downstream of the MCF13 murine leukemia virus enhancer. *J. Virol.* **66**:7080–7088.
- Vitelozzi, G., L. Mughetti, M. Palmarini, M. T. Mandara, L. Mechelli, J. M. Sharp, and I. Manocchio. 1993. Enzootic intranasal tumour of goats in Italy. *J. Vet. Med. Ser. B* **40**:459–468.
- Wu, Q., M. Chen, M. Buchwald, and R. A. Phillips. 1995. A simple, rapid method for isolation of high quality genomic DNA from animal tissues. *Nucleic Acids Res.* **23**:5087–5088.
- Yonemichi, H., Y. Ohgi, K. Fujimoto, M. Okado, T. Onuma, and T. Mikami. 1978. Intranasal tumour of the ethmoid olfactory mucosa in sheep. *Am. J. Vet. Res.* **39**:1599–1606.
- York, D. F., R. Vigne, D. W. Verwoerd, and G. Querat. 1992. Nucleotide sequence of the Jaagsiekte retrovirus, an exogenous and endogenous type D and B retrovirus of sheep and goats. *J. Virol.* **66**:4930–4939.
- Zhao, Y., L. Zhu, C. A. Benedict, D. Chen, W. F. Anderson, and P. M. Cannon. 1998. Functional domains in the retroviral transmembrane protein. *J. Virol.* **72**:5392–5398.