Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

VIROLOGY 192, 62–72 (1993)

Lelystad Virus, the Causative Agent of Porcine Epidemic Abortion and Respiratory Syndrome (PEARS), Is Related to LDV and EAV

JANNEKE J. M. MEULENBERG, MARCEL M. HULST, EMILE J. DE MEIJER, PETER L. J. M. MOONEN, ANNELIEN DEN BESTEN, ERIC P. DE KLUYVER, GERT WENSVOORT, AND ROB J. M. MOORMANN[1]

Central Veterinary Institute, Department of Virology, Houtribweg 39, 8221RA Lelystad, The Netherlands

Received June 18 1992; accepted August 24, 1992

The genome of Lelystad virus (LV), the causative agent of porcine epidemic abortion and respiratory syndrome (previously known as mystery swine disease), was shown to be a polyadenylated RNA molecule. The nucleotide sequence of the LV genome was determined from a set of overlapping cDNA clones. A consecutive sequence of 15,088 nucleotides was obtained. Eight open reading frames (ORFs) that might encode virus-specific proteins were identified. ORF1a and ORF1b are predicted to encode the viral RNA polymerase because the amino acid sequence contains sequence elements that are conserved in RNA polymerases of the torovirus Berne virus (BEV), equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV), the coronaviruses, and other positive-strand RNA viruses. A heptanucleotide slippery sequence (UUUAAAC) and a putative pseudoknot structure, which are both required for efficient ribosomal frameshifting during translation of the RNA polymerase ORF1b of BEV, EAV, and the coronaviruses, were identified in the overlapping region of ORF1a and ORF1b of LV. ORFs 2 to 6 probably encode viral membrane-associated proteins, whereas ORF7 is predicted to encode the nucleocapsid protein. Comparison of the amino acid sequences of the ORFs identified in the genome of LV, LDV, and EAV indicated that LV and LDV are more closely related than LV and EAV. A 3′ nested set of six subgenomic RNAs was detected in LV-infected cells. These subgenomic RNAs contain a common leader sequence that is derived from the 5′ end of the genomic RNA and that is joined to the 3′ terminal body sequence. Our results indicate that LV is closely related evolutionarily to LDV and EAV, both members of a recently proposed family of positive-strand RNA viruses, the Arteriviridae.

INTRODUCTION

Mystery swine disease (MSD), which causes abortions and respiratory disease in pigs and which was first detected in 1987 in North America (Hill, 1990), emerged in 1990 in Europe (Paton *et al.*, 1991; Plana *et al.*, 1991). This disease has acquired several other names, such as swine infertility and respiratory syndrome (SIRS), which is most commonly used in the United States, and porcine reproductive and respiratory syndrome (PRRS) and porcine epidemic abortion and respiratory syndrome (PEARS), which are most commonly used in Europe. PEARS mainly affects sows and their piglets in breeding herds. It causes anorexia and respiratory distress in pigs of all ages and abortions and stillbirths in pregnant sows (Loula, 1990).

Recently, we isolated Lelystad virus (LV) and demonstrated that infection with this virus caused the disease (Terpstra *et al.*, 1991; Wensvoort *et al.*, 1991). Lelystad virus appears to replicate *in vitro* only in primary cultures of porcine alveolar lung macrophages. The highest titer measured in the supernatant of these cell cultures was $10^{5.5}$ $TCID_{50} \cdot ml^{-1}$ (Wensvoort *et al.*, 1992). Electron microscopic studies showed that LV is 50 to 65 nm in size and that it is a spherical enveloped virus containing a nucleocapsid of 30 to 35 nm (Wensvoort *et al.*, 1992).

To unravel the genomic organization and replication strategy of LV, we analyzed virus-specific RNA in infected cells and determined the cDNA sequence of the genome of LV. In this paper, it is shown that LV contains a polyadenylated RNA genome of 15.1 kb and produces a 3′ coterminal nested set of subgenomic RNAs in infected cells. The nucleotide sequence and the organization of the genome indicate that LV is a member of a recently proposed new family of viruses, the Arteriviridae, consisting of lactate dehydrogenase-elevating virus (LDV), equine arteritis virus (EAV) and simmian hemorrhagic fever virus (SHFV) (Plagemann and Moennig, 1991).

MATERIALS AND METHODS

Isolation of RNA and DNA from LV-infected macrophage cultures

Primary cultures of porcine alveolar lung macrophages were grown in 75-cm^2 bottles, containing 25 ml RPMI 1640 growth medium (Flow). The initial cell concentration was 8×10^5 cells $\cdot ml^{-1}$. After 16 hr incu-

[1] To whom reprint requests should be addressed.

0042-6822/93 $5.00

bation at 37°, the medium was discarded and 3 ml virus suspension, with a titer of 10^5 $TCID_{50} \cdot ml^{-1}$, was added. After 2 hr, 25 ml fresh medium was added and the cells were incubated for 24 hr at 37°.

Cytoplasmic RNA was isolated from the infected cells as described by Favaloro *et al.* (1980). The RNA was purified by centrifugation through a 5.7-*M* CsCl cushion (Setzer *et al.*, 1980) and treated with RNase-free DNase (Pharmacia). RNA was analyzed in either an 0.8% neutral agarose gel, as described by Moormann and Hulst (1988), or a denaturing agarose formaldehyde gel, according to Sambrook *et al.* (1989).

For the isolation of nuclear DNA, the infected cells were lysed and the cell nuclei were pelleted as described by Favaloro *et al.* (1980). The nuclei were treated with proteinase K, extracted with phenol–chloroform, and precipitated with alcohol, as described by Sambrook *et al.* (1989).

Construction of a cDNA library

Intracellular RNA of LV-infected porcine alveolar lung macrophages (10 μg) was incubated with 10 m*M* methylmercury hydroxide for 10 min at room temperature. For reverse-transcription, the denatured RNA was incubated at 42° with 50 m*M* Tris–HCl, pH 7.8, 10 m*M* $MgCl_2$, 70 m*M* KCl, 0.5 m*M* dATP, dCTP, dGTP, and dTTP, 0.6 μg calf thymus oligonucleotide primers pd(*N*)6 (Pharmacia), and 300 units of Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories) in a total volume of 100 μl. After 1 hr, EDTA was added to a final concentration of 20 m*M*. The reaction mixture was extracted with phenol–chloroform, passed through a Sephadex G-50 column, and precipitated with ethanol.

To synthesize the second cDNA strand, we used DNA polymerase I (Boehringer) and RNase H (Pharmacia) (Gübler and Hoffman, 1983). To generate blunt ends, double-stranded cDNA was incubated with T4 DNA polymerase (Pharmacia) and Klenow polymerase (Pharmacia) in a reaction mixture that contained 0.05 m*M* deoxynucleotide–triphosphates. Subsequently, cDNA was fractionated in a 0.8% agarose gel (Moormann and Hulst, 1988). Fragments of 0.5 to 4 kb were electroeluted, ligated into the *Sma*I site of pGEM-4Z (Promega), and used for transformation of *Escherichia coli* strain DH5α (Hanahan, 1983). Colony filters were hybridized with a ^{32}P-labeled single-stranded cDNA probe. Reverse transcriptase was used to synthesize this probe from LV RNA (fractionated in a neutral agarose gel) (Moormann *et al.*, 1990). The single-stranded cDNA probe was incubated with excess cytoplasmic RNA from mock-infected alveolar lung macrophages and subsequently added to the hybridization solution.

The relationship between the LV cDNA clones was determined by restriction enzyme analysis and Southern blot hybridization with nick-translated cDNA probes (Sambrook *et al.*, 1989). The specificity of a few cDNA clones was confirmed. These clones were used as probes in Northern analysis of RNA extracted from LV-infected and mock-infected macrophages. Once the sequence of the cDNA clones was obtained, specific oligonucleotides were synthesized and used in a second screening of the cDNA library to isolate overlapping cDNA clones spanning the gaps between different cDNA groups. T4 kinase (Pharmacia) or the 3′ end-labeling kit (containing terminal transferase) of Boehringer was used to label these oligonucleotides with ^{32}P.

To obtain the 3′ end of the viral genome, we constructed a second cDNA library, using oligo$(dT)_{12-18}$ and a 3′ LV-specific oligonucleotide that was complementary to the minus-strand viral genome (39U183R) as primers in the first-strand reaction. The reaction conditions for first- and second-strand synthesis were identical to those described above. This library was screened with virus-specific oligonucleotide probes.

DNA sequencing

Most (>95%) of the cDNA sequence was determined with an Automated Laser Fluorescent DNA sequencer from Pharmacia LKB. The AutoRead sequencing kit (Pharmacia) was used essentially according to Procedures C and D as described in the Autoread sequencing kit protocol. The sequence was determined on double-stranded DNA. Fluorescent primers were prepared with FluorePrime (Pharmacia). The remaining part of the sequence was determined on double-stranded DNA using oligonucleotide primers, the T7 polymerase sequencing kit of Pharmacia, and [α-^{35}S]dATP (Amersham). The sequence analysis programs PCGENE (Intelligenetics, Inc., Mountain View, CA), the sequence analysis software package of the University of Wisconsin Genetics Computer Group (GCG; Devereux *et al.*, 1984), and FASTA (Pearson and Lipman, 1988) were used to assemble and analyze sequence data.

Nucleotide sequence accession number

The nucleotide sequence data reported in this paper have been deposited with the EMBL/GenBank Data Libraries under Accession No. M96262.

Northern Hybridizations

The total cytoplasmic RNA from LV- and mock-infected porcine alveolar lung macrophages was separated on a neutral agarose gel or a denaturing formalde-

hyde gel, as described above. After electrophoresis, the RNA was transferred to Hybond N membrane (Amersham). Virus-specific oligonucleotides (20-mers) were γ-^{32}P-labeled using T4 kinase and hybridized overnight in an aqueous solution of 6× SSC, 0.5% sodium pyrophosphate, 1× Denhardt's, 100 $\mu g \cdot ml^{-1}$ salmon sperm DNA, and 20 $\mu g \cdot ml^{-1}$ tRNA at 42° to the Northern blots. The Northern blots were washed extensively in 0.5× SSC, 0.5% sodium pyrophosphate at 42°, prior to autoradiography at −70°, using an intensifying screen. The Northern blots were used several times. The blots were washed in 0.1% SDS (heated to 90°) for 60 min to remove the probes. The removal of the probes was checked by autoradiography.

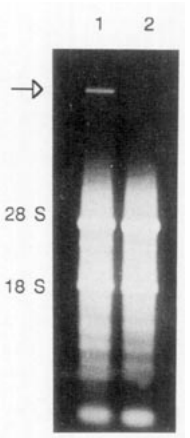


FIG. 1. *Neutral agarose gel electrophoresis of RNA isolated from* LV-infected (lane 1) and mock-infected (lane 2) alveolar lung macrophages. The LV genomic RNA is indicated with an arrow.

Oligonucleotides

The sequence of the oligonucleotides, which were used in Northern hybridization experiments or for the construction or screening of the cDNA library, and their position in the genome of LV are: 25U101R (5′ GGATACCTGTTGGGCTAGAAC 3′; 86–105), 5U94R (CGGCCCAGAACCAGTTGTGC; 4515–4534), 41U103R (CGGATGAATGTATACAACGGC; 11,375–11,394), 60R219 (CGACGAGAGACCATCTCATC; 12,098–12,117), 41R 1019 (GTTTGAGGTTGTCGTACCCG; 12,681–12,700), 41R346 (GGGTATGTTGGGTCACATGG; 13,355–13,374), 55U320 (CCGTAGACGCTGCAGAGTAC; 13,808–13,827), 39U356 (GCGACACGGTTGGTGGATTG; 14,269–14,288), 39U183R (CGTCTGGTAACCGAGCATAC; 14,453–14,472), 118U250 (CAGCCAGGGGAAAATGTGGC; 14,746–14,765), 119R64R (TCGCGTGACTTCTACATCC; 14,935–14,954), 119R10 (TAGGTGACTCAGAGGCCACA; 15,006–15,025).

RESULTS

Identification, cloning, and sequencing of the LV genome

To identify the LV genome, we isolated DNA and RNA from LV-infected and uninfected porcine alveolar lung macrophages. The analysis of the isolated nucleic acid fractions demonstrated that infected cells contained an RNA band that was absent from mock-infected cells (Fig. 1). This RNA migrated slightly slower than a preparation of hog cholera virus RNA of 12.3 kb (Moormann and Hulst, 1988) in a 0.8% neutral agarose gel (data not shown) and was estimated to be 14 to 15 kb in length. The analysis of nuclear DNA, isolated from LV-infected cells, in an ethidium–bromide stained agarose gel did not reveal any LV-specific DNA bands.

To analyze the LV genome in more detail, we determined its cDNA sequence. A random primed cDNA library was constructed from RNA of LV-infected alveolar lung macrophages, and LV-specific cDNA clones were isolated as described (see Materials and Methods). The inserts of positive clones were analyzed by restriction enzyme digestion, nick-translated, and used to search for overlaps between the various cDNA clones. The clones were found not to be contiguous but formed three separate blocks. cDNA clones in the region of the two remaining gaps were obtained using LV-specific oligonucleotide probes, complementary to the cDNA sequences at both sides of the gaps. The specificity of the cDNA clones was confirmed by hybridizing a set of five clones, located in different regions of the viral genome, to Northern blots carrying RNA of LV-infected macrophages. Remarkably, cDNA clones 4, 12, and 20 hybridized only with the viral RNA band of 14 to 15 kb, whereas cDNA clones 39 and 41 hybridized not only to this viral RNA but to a panel of four or five virus-specific RNAs of lower molecular weight as well (data not shown). The latter clones were located at the 3′ end of the viral genome, suggesting that LV RNA transcription results in the production of a nested set of subgenomic RNAs (see below).

Figure 2 shows the positions of all cDNA clones used to obtain the nucleotide sequence of the viral genome. The nucleotide sequence was determined on at least two independent cDNA clones and was 15,088 nucleotides in length. This sequence has been submitted to the EMBL/GenBank Database (Accession No. M96262) and will not be duplicated in this publication. We assume that the 5′ ultimate nucleotide is very close to the 5′ end of the LV genomic RNA. In addition to clones 25 and 155, four additional cDNA clones were isolated from the 5′ end of the viral genome (results not shown). The sequences of these clones were only two or three nucleotides shorter than

FIG. 2. Organization of the LV genome. The cDNA clones, which have been used for the determination of the nucleotide sequence, are indicated in (A). The parts of the clones, which were sequenced, are filled in (black bars). The ORFs, identified in the nucleotide sequence are shown. In (B) the subgenomic set of RNAs, encoding the ORFs, is shown. The leader sequence is indicated by a solid box.

the ultimate 5′ nucleotide. However, we cannot exclude that these cDNA clones terminate around the same nucleotide because of a strong-stop in the RNA which prematurely terminates the reverse-transcription reaction.

Since the sequence data indicated a close relationship among LV, EAV, and LDV (see below), we expected LV to be a positive-strand RNA virus, containing a 3′ poly(A) tail. Hence, to clone the 3′ end of the viral genome, we constructed a second cDNA library, using oligo(dT) and primer 39U183R in the reverse transcriptase reaction. Primer 39U183R is complementary to LV minus-strand RNA, which is likely present in a preparation of RNA isolated from LV-infected cells. This library was screened with virus-specific probes (nick-translated cDNA clone 119 and oligonucleotide 119R64R), resulting in the isolation of five additional cDNA clones (e.g., cDNA clone 151, Fig. 2). Sequencing of these cDNA clones revealed that LV contains a 3′ poly(A) tail. The length of the poly(A) tail varied among the selected cDNA clones, but its maximum length was 20 nucleotides.

Eight open reading frames (ORFs) that might encode virus-specific proteins were identified in the nucleotide sequence of LV (Fig. 2). Most of the sequence (approximately 80%) encodes two very large ORFs, ORF1a and ORF1b (Table 1). The C-terminus of ORF1a overlaps the putative N-terminus of ORF1b over a small distance of 16 nucleotides. We assume that ORF1a and ORF1b encode the viral polymerase and that ORF1b is expressed by ribosomal frameshifting (see below). The remaining ORFs, ORFs 2 to 7, are located at the 3′ part of the viral genome and partially overlap. These ORFs encode small polypeptides of 128 to 265 amino acid residues (Table 1).

Similarities between ORF1a and ORF1b and the RNA polymerases of BEV, EAV, LDV, and the coronaviruses

Molecular characterization of polymerase encoding regions of the coronaviruses mouse hepatitis virus (MHV-A59; Bredenbeek *et al.*, 1990; and MHV-JHM; Lee *et al.*, 1991) and avian infectious bronchitis virus (IBV; Boursnell *et al.*, 1987; and Brierley *et al.*, 1987), of the torovirus Berne virus (BEV; Snijder *et al.*, 1990a), and of EAV (den Boon *et al.*, 1991a) has indicated that the RNA polymerase is expressed through two overlapping ORFs, ORF1a and ORF1b. An RNA sequence element, containing a so-called slippery sequence and a sequence that may form an RNA pseudoknot structure in the ORF1a/ORF1b overlap region, was shown to be essential for the expression of ORF1b via a mechanism of ribosomal frameshifting (Brierley *et al.*, 1989). The former elements were also identified in the overlap region of LV ORF1a/ORF1b by a computer program designed to predict RNA secondary structure, including pseudoknotting, as described by Abrahams *et al.* (1990). The heptanucleotide slippery sequence UUUAAAC is located just upstream of the UAG stop codon of ORF1a (Fig. 3A). A potential stem–loop structure was identified downstream of this slippery sequence. In addition, a region encompassing nucleotides 7453 to 7460 may participate in forming a pseudoknot structure (Fig. 3B).

Further analysis of the predicted amino acid sequence of ORF1a of LV revealed that it has other features in common with the amino acid sequence of ORF1a of BEV, EAV, and the coronaviruses: hydrophobic regions (amino acids 1106 to 1196, 1226 to 1266, 1516 to 1646, and 1856 to 2036) and cysteine-rich domains (amino acids 262 to 290, 472 to 525, 1203 to 1284, and 2010 to 2108) were identified, and a putative serine protease consensus sequence was present between amino acids 1724 and 1830. The catalytic triad of this protease is likely formed by the His, Asp, and Ser residues at positions 1732, 1757, and 1810, respectively (Fig. 4A). As can be observed in Fig. 4B, the coronaviruses MHV and IBV are predicted to contain a 3C like cysteine protease. Sequence comparison and secondary structure predictions have indicated that the trypsin-like serine proteases and 3C-like cysteine protease likely belong to the same protease superfamily (Bazan *et al.*, 1988; and Gorbalenya *et al.*,

TABLE 1

CHARACTERISTICS OF THE ORFs OF LELYSTAD VIRUS

ORF	Nucleotides (first–last)	No. of amino acids	Calculated size of the unmodified product (kDa)	Glycosylation sites
ORF1a	212–7399	2396	260.0	3
ORF1b	7384–11772	1463	161.8	3
ORF2	11786–12532	249	28.4	2
ORF3	12394–13188	265	30.6	7
ORF4	12936–13484	183	20.0	4
ORF5	13484–14086	201	22.4	2
ORF6	14077–14595	173	18.9	2
ORF7	14588–14971	128	13.8	1

1989). In addition, a papain-like protease domain was detected in the N-terminal part of ORF1a of IBV (Boursnell *et al.*, 1987), MHV (Lee *et al.*, 1991), and EAV (den Boon *et al.*, 1991a), whereas two putative papain-like protease domains were identified in ORF1a of LV. The sequence around the Cys at positions 76 (GCCWL) and 276 (GKCWL) of LV ORF1a has similarity with the sequence surrounding the putative catalytic Cys of the protease domains of cellular and viral papain-like proteases (Lee *et al.*, 1991). The amino acid sequence of ORF1a contains several His residues (e.g., His at positions 146, 157, and 237 for the first domain and His at positions 345 and 432 for the second domain), which may function as the active site residue.

Except for these domains, no significant overall identity was observed between the amino acid sequences of ORF1a of LV and ORF1a of coronaviruses. However, the sequences of ORF1a of LV and EAV shared 25% identical residues over a length of 1283 residues (Fig. 5B). Although the amino acid sequence encoded by ORF1a of EAV was about 600 residues shorter at the N-terminus than that encoded by ORF1a of LV, the sequences could be aligned from amino acid 1118 of LV ORF1a (Fig. 5A, panel 1). The identity of LV ORF1a (amino acids 1 to 540 and 1006 to 1350) with polypeptides encoded by the ORFs identified in the sequences of cDNA clones of LDV, i.e., cDNA clone 4-37 (P. Plagemann, personal communication) and cDNA

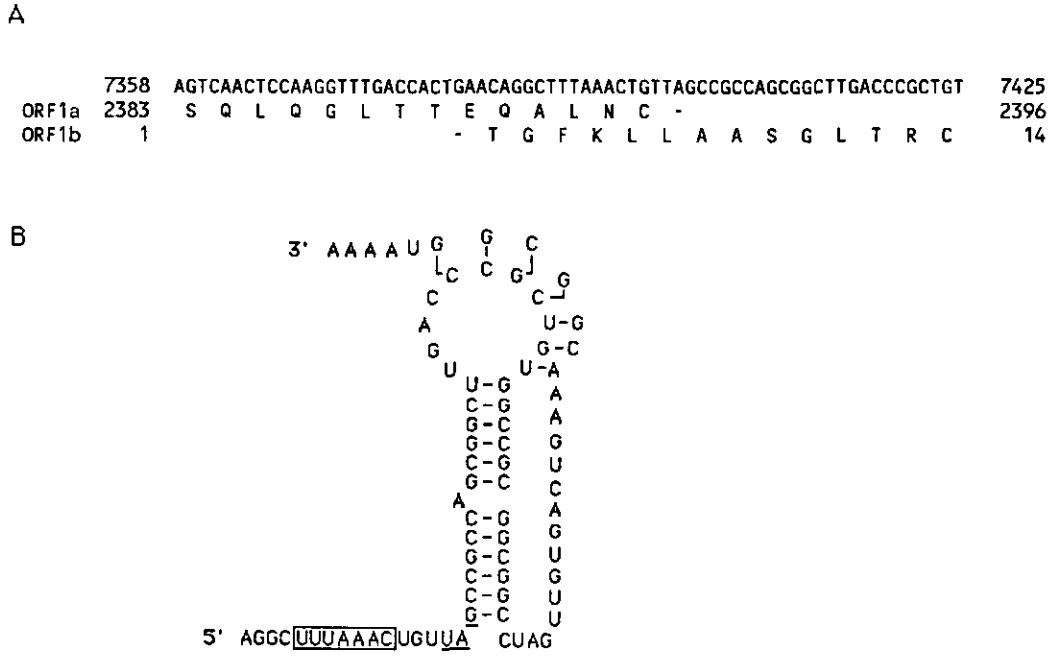


FIG. 3. Sequence of the overlapping region between ORF1a and ORF1b (A) and putative tertiary RNA structure (pseudoknot; B), as predicted by the computer program described by Abrahams *et al.* (1990). The heptanucleotide slippery sequence is boxed. The UAG stop codon of ORF1a is underlined.

```
A       TRP     QWVVSAAHCYKS - 37 - NNDIMLIKLKS - 77 - KDSCQGDSGGPVV-CSG-----KLQGIVSWGSGC
        HChV    GGISSVDHVTCG - 21 - MTDECETGVKT - 50 - LKNLKGWSGCPIF-EASSGR---VVGRVKVGKNE
        EAV     VVVLTASHVVGR - 19 - NGDFAEAVTTQ - 40 - AWTTSGDSGSAVV-QGDA-----VVGVHTGSNTSGVAYVTTPSGK
        LV      RTVVTAAHVL-N - 19 - NGDY--AWSHA - 40 - CFTNCGDSGSPVISESGD-----LIGIHTGSNKLGSGLVTTPEGE
                          ^             ^                 ^

B       PV1     NVAILPTHASPG - 38 - FRDIRQHIPTQ - 46 - FPTRAGQCGG-VITCTG------------KVIGMH-VGGNG
        CPMV    ACKHFFTHIKTK - 43 - CWDLFCWDPDK - 56 - APTIPEDCGS-LVIAHIGGKH--------KIVGVHVAGIQG
                          ^             ^                 ^
        EAV     VVVLTASHVVGR - 19 - NGDFAEAVTTQ - 40 - AWTTSGDSGSAVVQGD-------------AVVGYHT-GSNTSGVAY
        LV      RTVVTAAHVL-N - 19 - NGDY--AWSHA - 40 - CFTNCGDSGSPVISESG------------DLIGIHT-GSNKLGSGL
                          ^             ^                 ^
        MHV     DKVYCPRHVICS -          97          - GSFLCGSCGSVGYVLTGDSVRFVYMHQLELSTGCHT-GTDFSGNFY
        IBV     DTIYCPRHVLGK -         101          - ASFLAGACGSVGFNIEKGVVNFFYMHHLELPNALHT-GTDLMEGFY
                          ^                               ^

C       LV    372 KSPIALGKNKFKEL -  7 - CLEADLASCDRSTPAIVR - 31 - QDGAFTKRGGLSSGDPVTSVSNTVYS - 37 - PMLVYSDDLVLY  516
        LDV   264 GTPICLGKNKFTPL -  7 - CLEADLASCDRSTPAIIR - 31 - NSGCFDKRGGLSSGDPVTSISNTVYS - 38 - PLLVYSDDVVFY  409
        EAV   371 GSPIYLGKSKFDPI -  6 - CLETDLESCDRSTPALVR - 31 - GSVAFTKRGGLSSGDPITSISNTIYS - 38 - RVYIYSDDVVLT  515
        MHV   572 GVPVVIGTTKFYGG - 16 - LMGWDYPKCDRAMPNILR - 39 - G-CYYVKPGGTSSGDATTAFANSVFN - 58 - SMMILSDDGVVC  743
        IBV   578 NASVVIGTTKFYGG - 16 - LMGWDYPKCDRAMPNLLR - 39 - G-GIYVKPGGTSSGDATTAYANSVFN - 58 - SLMILSDDGVVC  749
        BEV   509 GGFCLIGVSKY--G - 18 - VFGSDYTKCDRTFPLSFR - 35 - G-MLLNKPGGTSSGDATTAHSNTFYN - 51 - FLNFLSDDSFIF  679
                    * *               *  *** *  *                * ** ****     *              ***
```

FIG. 4. Sequence analysis of conserved domains in the polymerase encoding region of LV. Comparison of the putative trypsin-like serine protease domain of LV with (A) these protease domains of trypsin (TRP), hog cholera virus strain Brescia (HChV; Moormann *et al.*, 1990), and EAV and (B) the 3C-like cysteine protease motifs of picornavirus (PV1) cowpea mosaic virus (CPMV) and the coronaviruses mouse hepatitis virus (MHV) and infectious bronchitis virus (IBV). The alignments were based on sequence data and comparisons presented by den Boon *et al.* (1991). The putative catalytic residues are indicated with an arrowhead. (C) Alignment of the polymerase motif of ORF1b of LV, LDV (Kuo *et al.*, 1991), EAV (den Boon *et al.*, 1991), MHV (Lee *et al.*, 1991), IBV (Boursnell *et al.*, 1984), and BEV (Snijder *et al.*, 1990). Identical amino acids are indicated with asterisks.

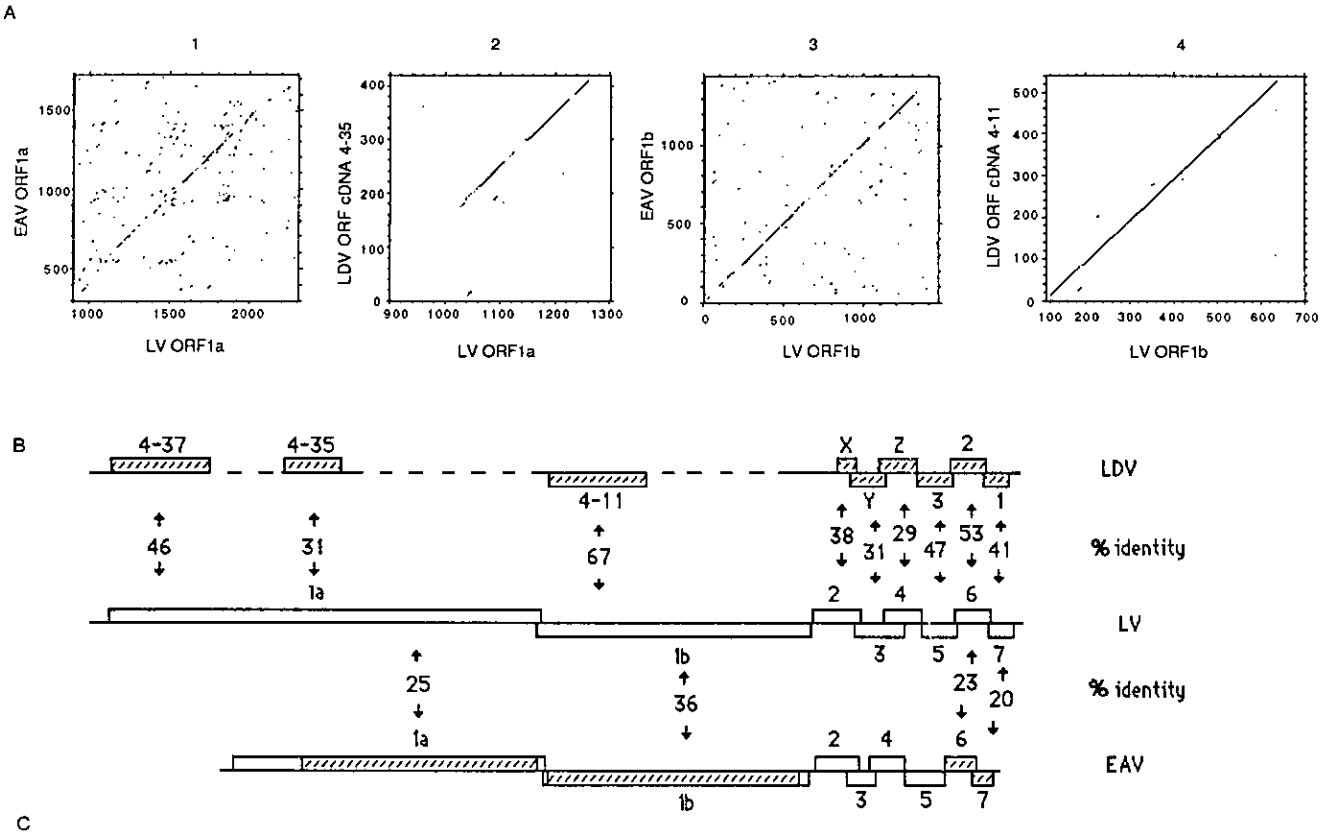
clone 4-35 (Kuo *et al.*, 1991), was 46 and 31%, respectively (Figs. 5A, panel 2, and 5B).

Comparison of the predicted amino acid sequence of ORF1b of LV and the polypeptides encoded by ORF1b of BEV, LDV, EAV and the coronaviruses showed that ORF1b is more conserved than ORF1a. The four characteristic domains identified in ORF1b of these viruses were also present in LV ORF1b at the same relative position. They were: (I) the polymerase motif containing the core sequence S/GDD, identified in the RNA polymerase of all positive-strand RNA viruses (Poch *et al.*, 1989; amino acids 372 to 575); (II) a putative cysteine- and histidine-rich zinc finger domain (amino acids 646 to 688); (III) a nucleoside triphosphate binding and helicase motif (amino acids 792 to 1015); and (IV) a conserved domain of unknown function (amino acids 1217 to 1308). An alignment of the polymerase motif of LV, LDV, EAV, MHV, IBV, and BEV is shown in Fig. 4C. The similarity between the polypeptides encoded by ORF1b of LV and EAV was not restricted to these four conserved domains. The amino acid sequence encoded by ORF1b of both viruses could be aligned without the introduction of large gaps, and they shared 36% identical amino acids over a length of 1305 residues (Figs. 5A, panel 3, and 5B). Furthermore, the amino acid sequence encoded by LV ORF1b had 67% identical residues with the amino acid sequence of a fragment of LDV ORF1b encoded by cDNA clone 4-11 (Kuo *et al.*, 1991; Figs. 5A, panel 4, and 5B). The identity between ORF1b of LV and ORF1b of EAV in this specific region was lower (45%).

ORFs 2 to 7 encode putative viral structural proteins

The amino acid sequences encoded by ORFs 2 to 6 all show features reminiscent of membrane-associated (envelope) proteins. They all contain putative N-linked glycosylation sites (Table 1), and N- and C-terminal hydrophobic sequences that may function as a signal sequence and a membrane anchor, respectively, were identified. The hydropathicity plot of the product of ORF6 had a very remarkable pattern (Fig. 6). Three hydrophobic, putative membrane-spanning fragments were identified in the 90 N-terminal amino acids. A similar hydropathicity profile was observed for ORF6 of EAV (den Boon *et al.*, 1991a), ORF2 or VpX of LDV (Godeny *et al.*, 1990; Kuo *et al.*, 1992), the M protein of MHV (Boursnell, 1984) and IBV (Rottier *et al.*, 1986), and the E protein of BEV (den Boon *et al.*, 1991b) (Fig. 6). These findings suggest a similar topology and membrane-associated function.

The amino acid sequences of ORFs 2 to 6 shared 29 to 53% identical residues with the amino acid sequences of ORFs identified in the nucleotide sequence of LDV (Godeny *et al.*, 1990; Kuo *et al.*, 1992; and P. Plagemann, personal communication), whereas the



C

```
LV  ORF6  MGGLDDFCNDPIAAQKLVLAFSITYTPIMIYALKVSRGRLLGLLHILIFLNCSFTFGYMTYVHFQSTNRVALTLGAVVALLWGVYSFTESWKFITSRCRL  100
LDV ORF2  MGGLE-FCDQTSWYQILI-AFSLTYTPIAIYSLKVFRGTLAGIVNIFIFINCCVSFVYLMY-HHSVINTVALSLGAVIALVWGIYTLVKIVWWMVLRCRL   97
EAV ORF6  MGAIDSFCGDGILGE--YLDYFILSVFLLLLLTRYVASGLVYVLTALFY-----SFVLAAYIWFVIVGR-AFSTAYAFVLL-AAFLLLVMRMIVGMMPRL   91
          **... **..    .    ..    .*. .  .  . *  .. ...    .*   *     ..  *. . . . .*. . . .      .  . **

LV  ORF6  CCLGRRYILAPAHHVESAAGLHSISASGNRAYAVRKPGLTSVNGTLVPGLRSLVLGGKRAVKRGVVNLVKY-GR                          173
LDV ORF2  CFLGRSYILAPPSHVDTSDGRQSLTTSSTTAFVVRKPGSILVNGQLVPDFQRLVLGGKKAVSKGAVNLLKYVSK                          171
EAV ORF6  RSIFNHRQLVVADFVDTPSGPVPIPRS-TTQVVVRGNGYTAVGNKLVDGVKTITSAGRLFSKRTAAT--AYKLQ                          162
            . .   *. .  *....*  ... * . . .** * * *.. ** . . . . .*.  ......  *  .

LV  ORF7  MAGK-NQSQKKKKSTAPMGNGQPVNQLCQLLGAM-IKSQRQQPRGGQAKKKKPEKPHFPLAAEDDIRHHLTQTERSLCLQSIQTAFNQGAGTASLSSSGK   98
LDV ORF1  M------SQNKKKSGQNKGANQ---QLNQLINAL-LRNAGQN-KGKGQKKKKQPKLHFPMAGPSDLRHVMTPNEVQMCRSSLVTLFNQGGGQCTLVDSGG   89
EAV ORF7  MASRRSRPQAASFRNGRRRQPTSYNDLLRMFGQMRVRKPPAQPTQAI------------IAEPGDLRHDLNQQERATLSSNVQRFFMIGHGSLT-ADAGG   87
          *      .*  .  .  .  .    .* ..... .... ..                  .*. .*.** ... *      ..   *  * *  .  ..*

LV  ORF7  VSFQVEFMLFVAHTVRLIRVTSTSASQGAS                                                                        128
LDV ORF1  INFTVSIMLPTHAIVRLI----NASANSSA                                                                        115
EAV ORF7  LTYTVSWV-PTKQIQR------KVAPPAGP                                                                        110
          ... *... *.  . *      . .. ...
```

FIG. 5. Identity between the viral proteins of LV, EAV, and LDV. (A) Dot matrix comparison of the amino acid sequences of ORF1a of LV and EAV (panel 1), ORF1a of LV and LDV (panel 2), ORF1b of LV and EAV (panel 3), and ORF1b of LV and LDV (panel 4), using the GCG program DOTPLOT at a stringency of 13 and a window of 21. Numbers of the amino acid residues are indicated at the axis. (B) Percentage of identity between the amino acid sequences encoded by ORFs of LV, LDV (Kuo *et al.*, 1991, 1992, and Godeny *et al.*, 1990), and EAV (den Boon *et al.*, 1991). The boxes represent the open reading frames, identified in the nucleotide sequence of the genome of these viruses. The sequence of cDNA clone 4-37 was kindly provided by P. Plagemann. Several names are currently used to refer to the structural proteins of LDV. The two ORFs, located at the 3' terminal part of the genome, are designated ORF2 and ORF1 by Kuo *et al.* (1992), whereas others refer to VpX and Vp1, respectively (Godeny *et al.*, 1990). The sequence of the ORFs, designated X, Y, and Z in this figure, was obtained from P. Plagemann and has not yet been published. The dashed line indicates the regions of the LDV genome which have not yet been sequenced. The regions of the ORFs of LDV and EAV, which have identity with the ORFs of LV, are shaded. (C) Alignment of the amino acid sequences of ORFs 6 and 7 of LV with corresponding sequences of LDV and EAV. Asterisks indicate an identical amino acid residue, and a period indicates a conservative substitution.

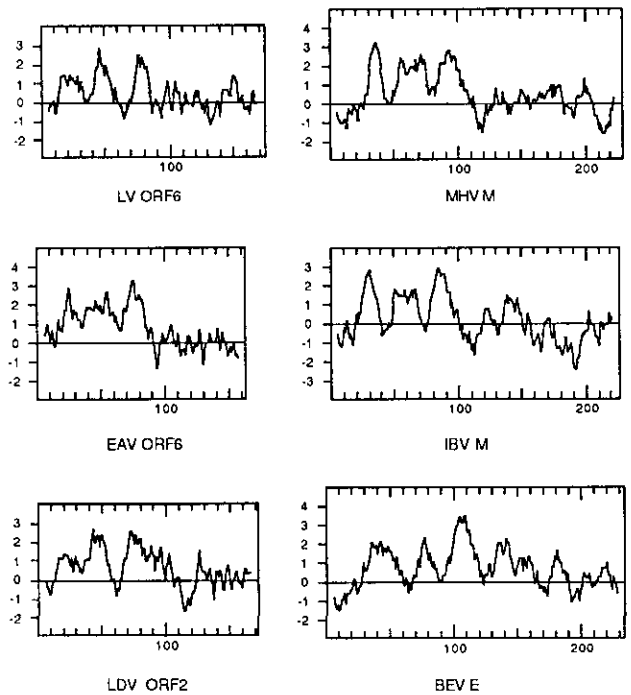


FIG. 6. Hydropathicity analysis of LV ORF6, EAV ORF6, LDV ORF2 (VPX), MHV M, IBV M, and BEV E, according to the method of Kyte and Doolittle (1982). The vertical scale is the average hydropathicity for a frame of seven amino acids. Hydrophobic sequences appear above the base line and hydrophilic sequences appear below the base line. The numbers of the amino acid residues are indicated on the horizontal axis.

amino acid sequence of ORF6 of LV shared a significant percentage of identical residues (26%) with ORF6 of EAV only (den Boon *et al.*, 1991a) (Figs. 5B and 5C).

In contrast to the structural proteins, likely encoded by ORFs 2 to 6, the polypeptide encoded by ORF7 was extremely basic; 26% of the N-terminal half of this polypeptide consisted of Arg, Lys, and His residues. The amino acid sequence encoded by ORF7 shared 20 and 41% identical residues with the polypeptide encoded by ORF7 of EAV and ORF1 or Vp1 of LDV, respectively (Figs. 5B and 5C). Godeny *et al.* (1990) have reported that Vp1 is the nucleocapsid protein of LDV. The basic character and the identity with Vp1 suggest a similar function for the product of LV ORF7.

Identification of a 3′ nested set of subgenomic RNAs

In Northern hybridization experiments, using cDNA clones located at the 3′ terminal part of the LV genome as probes, RNAs of smaller size than the genomic RNA were also detected in LV-infected cells. This finding indicated that the genome of LV is expressed in a manner similar to the genomes of BEV, EAV, LDV, and coronaviruses, i.e., via the production of subgenomic RNAs, which together form a 3′ coterminal nested set (de Vries *et al.*, 1990; Kuo *et al.*, 1991, 1992; Snijder *et al.*, 1990b; Spaan *et al.*, 1988; van Berlo *et al.*, 1986). To analyze these subgenomic RNAs in more detail, a set of LV-specific oligonucleotides located in the unique part of the various ORFs and at the extreme 5′ and 3′ ends of the LV genome were synthesized (Materials and Methods and Fig. 7A). These oligonucleotides were labeled and hybridized to Northern blots of RNA isolated from LV-infected and mock-infected alveolar lung macrophages. Hybridization bands were observed in the analysis of RNA from LV-infected cells (Fig. 7B), but not in the analysis of RNA from mock-infected cells (data not shown). Figure 7B shows that oligonucleotide 25U101R, which is complementary to the 5′ end of the viral genome, hybridized to viral genomic RNA of about 15 kb (RNA1) and to six smaller RNAs of 3.3 (RNA2), 2.7 (RNA3), 2.2 (RNA4), 1.7 (RNA5), 1.1 (RNA6), and 0.7 (RNA7) kb. Thus, these six subgeno-

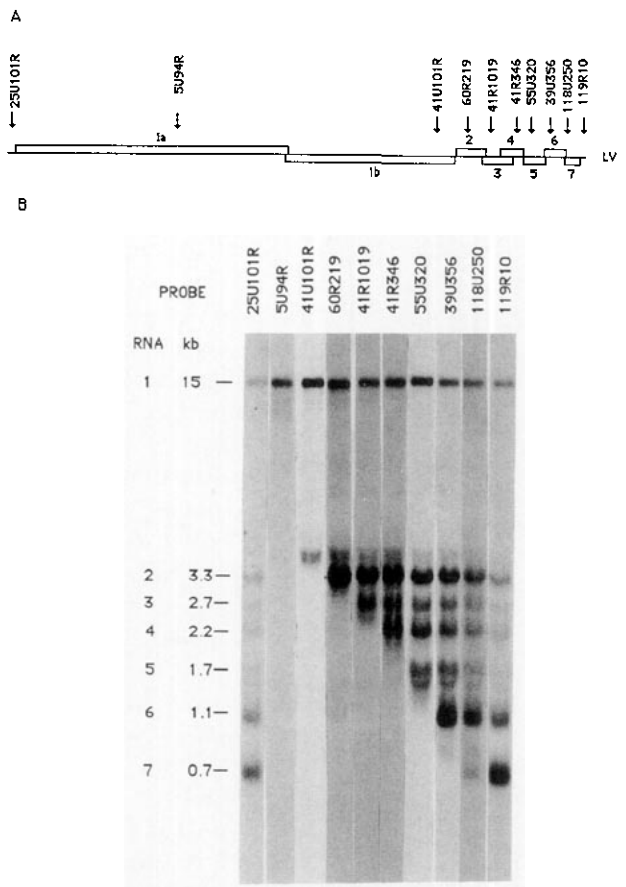


FIG. 7. Northern hybridization of RNA, isolated from LV-infected alveolar lung macrophages. The RNA was separated on a denaturing formaldehyde agarose gel. The blots were sequentially hybridized with the indicated LV-specific oligonucleotides (B). The location of these oligonucleotides in the genome is shown in (A).

mic RNAs likely possess a common leader sequence, which is derived from the 5′ end of the viral genome. The hybridization patterns of the other oligonucleotides indicated that these subgenomic RNAs form a 3′ coterminal nested set (Figs. 2 and 7B). Besides the seven RNAs, two additional bands were observed to varying extents at about 1.5 and 4.3 kb. These bands comigrated with the 18S and 28S ribosomal RNA bands, as detected by ethidium bromide staining of the gel. For this reason they might reflect aspecific binding of the probes to ribosomal RNA which is present in high concentrations in the total cytoplasmic RNA preparations. Alternatively, the bulk amount of ribosomal RNA might cause retardation of LV-specific RNA in the gel, resulting in the 4.3- and 1.5-kb bands observed.

DISCUSSION

We have shown that the genome of LV, the causative agent of porcine epidemic abortion and respiratory syndrome, is a polyadenylated RNA molecule. Because LV is closely related to several positive-strand RNA viruses, such as EAV and LDV, we presume that its genomic RNA also has a positive polarity. Although investigators have shown in *in vitro* transfection experiments that the genomic RNA of the latter viruses is infectious (Brinton-Darnell and Plagemann, 1975; van der Zeijst *et al.*, 1975), it is still unknown whether this is true to for LV. The nucleotide sequence of the LV genome, which was determined from a set of overlapping cDNA clones, consists of 15,088 bases. At the 5′ end, the viral genome may be a few bases longer than the sequence described in this paper. In the nucleotide sequence, ORFs have been identified that might encode the viral RNA polymerase (ORF1a and ORF1b), membrane-associated proteins (ORFs 2 to 6), and the nucleocapsid protein (ORF7). The order of the genes, i.e., 5′–gene encoding the RNA polymerase–genes encoding the membrane-associated proteins–gene encoding the nucleocapsid protein–3′, is the same in BEV (Snijder *et al.*, 1990a,b), EAV (den Boon *et al.*, 1991a), LDV (Plagemann and Moennig, 1991), and the coronaviruses (Spaan *et al.*, 1988). Furthermore, the genomes of LV, EAV, LDV, and the coronaviruses appear to be similarly expressed. During the replication of LV RNA, a 3′ coterminal nested set of six subgenomic RNAs is formed, which contain an identical leader sequence derived from the noncoding 5′ end of the genomic RNA. Studies on the expression of subgenomic mRNAs of coronaviruses have shown that only the unique region, i.e., the region of the mRNA that is not present in the smaller subgenomic mRNAs, is translated (Spaan *et al.*, 1988). Sequence analysis of subgenomic cDNA clones has indicated that the so-called junction sequence, the site in the subgenomic mRNA where the leader is fused to the body mRNA, is conserved (Spaan *et al.*, 1988). De Vries *et al.* have shown that the leader of EAV consists of 208 nucleotides and is fused to the body sequence of mRNAs 6 and 7 at the sequence 5′–UCAAC–3′ (den Boon *et al.*, 1991a; de Vries *et al.*, 1990). The junction site for mRNA7 of LDV appeared to be the sequence 5′–AAACC–3′, whereas 5′–UAACC–3′ was identified for mRNAs 5, 6, and 8 (Plagemann and Moennig, 1991). The latter sequence was also identified in the nucleotide sequence of LV, immediately upstream of ORF1a. ORFs 4 to 7 are preceded by a similar motif, 5′–(U/A)(U/C)AACC–3′, which may represent the junction sequence of subgenomic RNAs 4 to 7. The segments 5′–UAAACC–3′ and 5′–UUGACC–3′ were identified as putative junction sequences of RNA2 and RNA3, respectively. Experiments are in progress to determine the length of the LV leader sequence exactly and to identify the junction sites.

The organization of the LV genome, the conserved domains in the RNA polymerase gene, and the expression of the structural proteins via a 3′ nested set of subgenomic RNAs are findings that imply that LV, the coronaviruses, and toroviruses are evolutionarily related. However, the sequence data presented here indicate that LV, EAV, and LDV are even closer related. The size of the genome of these viruses (13 to 15 kb) is much smaller than that of coronaviruses and torovirus BEV (27 to 30 kb; Snijder *et al.*, 1990b; Spaan *et al.*, 1988). Especially the gene, encoding the viral polymerase of coronaviruses and BEV (20 to 22 kb; Boursnell *et al.*, 1987; Lee *et al.*, 1991; Snijder *et al.*, 1990a) is much longer than that of LV, EAV, or LDV (10 to 11 kb; den Boon *et al.*, 1991a; Plagemann and Moennig, 1991). The genomes of LV, EAV, and LDV (although the latter is only partially sequenced) are similarly organized and these viruses also synthesize a 3′ nested set of subgenomic RNAs in infected cells (de Vries *et al.*, 1990; Kuo *et al.*, 1991, 1992; van Berlo *et al.*, 1986). The percentage of identical amino acids between the sequences of the viral RNA polymerase, the putative membrane-associated proteins, and the putative nucleocapsid protein of LV and LDV (29 to 67%) and LV and EAV (20 to 36%) indicated that LV is more closely related to LDV than it is to EAV.

LV, EAV, and LDV also share other properties. First, macrophages are most likely the target cells of all three viruses in their respective hosts. This fact has been established most conclusively for LDV (Stueckemann *et al.*, 1982a). It is still unknown whether LV and EAV infect other cells as well *in vivo*. LV and LDV grow *in vitro* only on primary cultures of macrophages while EAV is exceptional in that it replicates in many different

cell lines (McCollum *et al.*, 1971). Second, LV, LDV, and EAV all seem to establish asymptomatic persistent infections in their natural hosts (G. Wensvoort, unpublished results; Stueckemann *et al.*, 1982b; Timoney *et al.*, 1986). Third LV, LDV, and EAV have similar morphological features. The size of the virion particles, as determined in electron microscopic studies, is similar. The spherical virions of 50 to 65 nm contain a nucleocapsid core with a diameter of 30 to 35 nm (Brinton-Darnell and Plagemann, 1975; Hyllseth, 1973; Wensvoort *et al.*, 1992), and the nucleocapsid shows a cubical rather than a helical symmetry.

LV, LDV, and EAV appear to be closely related to each other and distantly related to coronaviruses and toroviruses. Snijder *et al.* (1990a) and den Boon *et al.* (1991a) proposed a coronavirus-like superfamily, comprising coronaviruses and toroviruses, based on similarity in genome organization, genome expression, and amino acid sequence. On the basis of the same criteria, EAV, initially classified as a togavirus because of its morphological resemblance to this family (Westaway *et al.*, 1985), was also proposed to be a member of the coronavirus-like superfamily (den Boon *et al.*, 1991a). Recently, Plagemann and Moennig (1991) suggested classifying LDV and simian hemorrhagic fever virus (SHFV) together with EAV in a separate and new family of viruses, the Arteriviridae. Although the SHFV genome has not yet been sequenced, several properties of the virus—the size and shape of the virion, the size and polyadenylation of the genome, the nature of the structural proteins, and its propensity for macrophages—indicate that SHFV resembles LV, EAV, and LDV (Plagemann and Moennig, 1991). We conclude that LV, because of its nucleotide sequence, its genomic organization, and its replication strategy, represents a new member of the proposed Arteriviridae family.

ACKNOWLEDGMENTS

We thank P. Terpstra for his assistance in the sequence analysis, E. ten Dam for analyzing the pseudoknot structure, V. Thatcher for editorial assistance, P. Plagemann, E. Snijder, and W. Spaan for fruitful discussions, and A. Gielkens for critical reading of the manuscript. Part of this work was supported by Boehringer Ingelheim, Germany, and the Produktschap voor Vee en Vlees (PVV), the Netherlands.

REFERENCES

ABRAHAMS, J. P., VAN DEN BERG, M., VAN BATENBURG, E., and PLEIJ, C. (1990). Prediction of RNA secondary structure, including pseudoknotting, by computer simulation. *Nucleic Acids Res.* **18,** 3035–3044.

BAZAN, J. F., and FLETTERICK, R. J. (1988). Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. *Proc. Natl. Acad. Sci. USA* **85,** 7872–7876.

BOURSNELL, M. E. G., BROWN, T. D. K., and BINNS, M. M. (1984). Sequence of the membrane protein gene from avian coronavirus IBV. *Virus Res.* **1,** 303–314.

BOURSNELL, M. E. G., BROWN, T. D. K., FOULDS, I. J., GREEN, P. F., TOMLEY, F. M., and BINNS, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* **68,** 57–77.

BREDENBEEK, P. J., PACHUK, C. J., NOTEN, A. F. H., CHARITÉ, J., LUYTJES, W., WEISS, S. R., and SPAAN, W. J. M. (1990). The primary structure and expression of the second open reading frame of the polymerase gene of coronavirus MHV-A59. *Nucleic Acids Res.* **18,** 1825–1832.

BRIERLEY, I., BOURSNELL, M. E. G., BINNS, M. M., BILIMORIA, B., BLOK, V. C., BROWN, T. D. K., and INGLIS, S. C. (1987). An efficient ribosomal frameshifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* **6,** 3779–3785.

BRIERLEY, I., DIGGARD, P., and INGLIS, S. C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal, requirement for an RNA pseudoknot. *Cell* **57,** 537–547.

BRINTON-DARNELL, M., and PLAGEMANN, P. G. W. (1975). Structure and chemical physical characteristics of lactate dehydrogenase-elevating virus and its RNA. *J. Virol.* **16,** 420–433.

DEN BOON, J. A., SNIJDER, E. J., CHIRNSIDE, E. D., DE VRIES, A. A. F., HORZINEK, M. C., and SPAAN, W. J. M. (1991a). Equine arteritis virus is not a togavirus but belongs to the coronavirus superfamily. *J. Virol.* **65,** 2910–2920.

DEN BOON, J. A., SNIJDER, E. J., LOCKER, J. K., HORZINEK, M. C., and ROTTIER, P. J. M. (1991b). Another triple-spanning envelope protein among intracellularly budding RNA viruses: The torovirus E protein. *Virology* **182,** 655–663.

DEVEREUX, J., HAEBERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387–395.

DE VRIES, A. A. F., CHIRNSIDE, E. D., BREDENBEEK, P. J., GRAVESTEIN, L. A., HORZINEK, M. C., and SPAAN, W. J. M. (1990). All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res.* **18,** 3241–3247.

FAVALORO, J., TREISMAN, R., and KAMEN, R. (1980). *In* "Methods in Enzymology" (L. Grossman and K. Moldave, Eds.), Vol. 65, pp. 718–749. Academic Press, New York.

GORBALENYA, A. E., DONCHENKO, A. P., KOONIN, E. V., and BLINOV, V. M. (1989). Cysteine proteases of positive-strand RNA viruses and chymotrypsin-like serine proteases: A distinct protein superfamily with a common structural fold. *FEBS Lett.* **243,** 103–114.

GODENY, E. K., SPEICHER, D. W., and BRINTON, M. A. (1990). Map location of lactate dehydrogenase-elevating virus (LDV) capsid protein (Vp1) gene. *Virology* **177,** 768–771.

GÜBLER, U., and HOFFMAN, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25,** 263–269.

HANAHAN, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166,** 577–580.

HILL, H. (1990). Overview and history of mystery swine disease (Swine Infertility Respiratory Syndrome). *In* "Proceedings of the Mystery Swine Disease Committee Meeting", October 6, 1990, Denver, CO; and Livestock Conservation Institute, October 29–31, 1990, Madison, WI.

HYLLSETH, B. (1973). Structural proteins of equine arteritis virus. *Arch. Gesamte Virusforsch.* **40,** 177–188.

KUO L., CHEN, Z., ROWLAND, R. R. R., FAABERG, K. S., and PLAGEMANN, P. G. W. (1992). Lactate dehydrogenase-elevating virus (LDV): Subgenomic mRNAs, mRNA leader and comparison of 3′ terminal sequences of two LDV isolates. *Virus Res.* **23,** 55–72.

KUO, L., HARTY, J. T., ERICKSON, L., PALMER, G. A., and PLAGEMANN, P. G. W. (1991). A nested set of eight RNAs is formed in macrophages infected with lactate dehydrogenase-elevating virus. *J. Virol.* **65,** 5118–5123.

KYTE, J., and DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132.

LEE, H., SHIEH, C., GORBALENYA, A. E., KOONIN, E. V., MONICA, N. L., TULER, J., BAGDZHADZHYAN, A., and LAI, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* **180,** 567–582.

LOULA, T. (1990). Clinical presentation of mystery pig disease in the breeding herd and suckling piglets. *In* "Proceedings of the Mystery Swine Disease Committee Meeting", October 6, 1990, Denver, CO; and Livestock Conservation Institute, October 29–31, 1990, Madison, WI.

MCCOLLUM, W. H., PRICKET, M. E., and BRYANS, J. T. (1971). Temporal distribution of equine arteritis virus in respiratory mucosa tissues and body fluids of horses infected by inhalation. *Res. Vet. Sci.* **12,** 459–464.

MOORMANN, R. J. M., and HULST, M. M. (1988). Hog cholera virus: Identification and characterization of the viral RNA and the virus-specific RNA synthesized in infected swine kidney cells. *Virus. Res.* **11,** 281–291.

MOORMANN, R. J. M., WARMERDAM, P. A. M., VAN DER MEER, B., SCHAAPER, W. M. M., WENSVOORT, G., and HULST, M. M. (1990). Molecular cloning and nucleotide sequence of hog cholera virus strain Brescia and mapping of the genomic region encoding envelope protein E1. *Virology* **177,** 184–198.

PATON, D. J., BROWN, I. H., EDWARDS, S., and WENSVOORT, G. (1991). Blue ear disease of pigs. *Vet. Rec.* **128,** 617.

PEARSON, W. R., and LIPMAN, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.

PLAGEMANN, P. G. W., and MOENNIG, V. (1991). Lactate dehydrogenase-elevating virus, equine arteritis virus, and simian hemorrhagic fever virus: A new group of positive-strand RNA viruses. *Adv. Virus Res.* **41,** 99–192.

PLANA, J., VAYREDA, M., and VILARASSA, J. L. (1991). Porcine epidemic abortion and respiratory syndrome (mystery pig disease), isolation in Spain of the causative agent and experimental reproduction of the disease. *In* "Proceedings of the Porcine Reproductive and Respiratory Syndrome Seminar", November 4–5, 1991, Brussels, Belgium, European Commission VI/B/II.2 (pvet/en/0207-2), 7.

POCH, O., SAUVAGET, I., DELARUE, M., and TORDO, N. (1989). Identification of four conserved motifs among the RNA dependent polymerase encoding elements. *EMBO J.* **8,** 3867–3874.

ROTTIER, P. J. M., WELLING, G. W., WELLING-WESTER, S., NIESTERS, H. G. M., LENSTRA, J. A., and VAN DER ZEIJST, B. A. M. (1986). Predicted membrane topology of the coronavirus protein E1. *Biochemistry* **25,** 1335–1339.

SAMBROOK, J., FRITSCH, E. F., and MANIATIS, T. (1989). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

SETZER, D. R., MCGROGAN, M., NUNBERG, J. H., and SCHIMKE, R. T. (1980). Size heterogeneity in the 3'-end of the dehydrofylate reductase messenger RNAs in mouse cells. *Cell* **22,** 361–370.

SNIJDER, E. J., DEN BOON, J. A., BREDENBEEK, P. J., HORZINEK, M. C., RIJNBRAND, R., and SPAAN, W. J. M. (1990a). The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionary related. *Nucleic Acids Res.* **18,** 4535–4542.

SNIJDER, E. J., HORZINEK, M. C., and SPAAN, W. J. M. (1990b). A 3'-coterminal nested set of independently transcribed mRNAs is generated during Berne virus replication. *J. Virol.* **64,** 331–338.

SPAAN, W. J. M., CAVANAGH, D., and HORZINEK, M. C. (1988). Coronaviruses: Structure and genome expression. *J. Gen. Virol.* **69,** 2939–2952.

STUECKEMANN, J. A., HOLTH, M., SWART, W. J., KOWAICHYK, K., SMITH, M. S., WOLSTENHOLME, A. J., CAFRUNY, W. A., and PLAGEMANN, P. G. W. (1982a). Replication of lactate dehydrogenase-elevating virus in macrophages. 2. Mechanism of persistent infection in mice and cell culture. *J. Gen. Virol.* **59,** 263–272.

STUECKEMANN, J. A., RITZI, D. M., HOLTH, M., SMITH, M. S., SWART, W. J., CAFRUNY, W. A., and PLAGEMANN, P. G. W. (1982b). Replication of lactate dehydrogenase-elevating virus in macrophages. 1. Evidence for cytocidal replication. *J. Gen. Virol.* **59,** 245–262.

TERPSTRA, C., WENSVOORT, G., and POL, J. M. A. (1991). Experimental reproduction of porcine epidemic abortion and respiratory syndrome (mystery swine disease) by infection with Lelystad virus: Koch's postulates fulfilled. *Vet. Q.* **13,** 131–136.

TIMONEY, P. J., MCCOLLUM, W. H., ROBERTS, A. W., and MURPHY, T. W. (1986). Demonstration of the carrier state in naturally acquired equine arteritis virus infection in the stallion. *Res. Vet. Sci.* **41,** 279–280.

VAN BERLO, M. F., ROTTIER, P. J. M., HORZINEK, M. C., and VAN DER ZEIJST, B. A. M. (1986). Intracellular equine arteritis virus (EAV)-specific RNAs contain common sequences. *Virology* **152,** 492–496.

VAN DER ZEIJST, B. A. M., HORZINEK, M. C., and MOENNIG, V. (1975). The genome of equine arteritis virus. *Virology* **68,** 418–425.

WENSVOORT, G., DE KLUYVER, E. P., POL, J. M. A., WAGENAAR, F., MOORMANN, R. J. M., HULST, M. H., BLOEMRAAD, R., DEN BESTEN, A., ZETSTRA, T., and TERPSTRA, C. (1992). Lelystad virus, the cause of porcine epidemic abortion and respiratory syndrome: A review of mystery swine disease research at Lelystad. *Vet. Microbiol.,* in press.

WENSVOORT, G., TERPSTRA, C., POL, J. M. A., TER LAAK, E. A., BLOEMRAAD, M., DE KLUYVER, E. P., KRAGTEN, C., VAN BUITEN, L., DEN BESTEN, A., WAGENAAR, F., BROEKHUIJSEN, J. M., MOONEN, P. L. J. M., ZETSTRA, T., DE BOER, E. A., TIBBEN, H. J., DE JONG, M. F., VAN 'T VELD, P., GROENLAND, G. J. R., VAN GENNEP, J. A., VOETS, M. TH., VERHEIJDEN, J. H. M., and BRAAMSKAMP, J. (1991). Mystery swine disease in the Netherlands: The isolation of Lelystad virus. *Vet. Q.* **13,** 121–130.

WESTAWAY, E. G., BRINTON, M. A., GAIDAMOVICH, S. Y., HORZINEK, M. C., IGARASHI, A., KAARIANEN, L., LVOV, D. K., PORTERFIELD, J. S., RUSSEL, P. K., and TRENT, D. W. (1985). Togaviridae. *Intervirology* **24,** 125–139.