Arch Virol (1994) [Suppl] 9: 441–448

Archives of Virology

Lelystad virus belongs to a new virus family, comprising lactate dehydrogenase-elevating virus, equine arteritis virus, and simian hemorrhagic fever virus

J. J. M. Meulenberg, M. M. Hulst, E. J. de Meijer, P. L. J. M. Moonen, A. den Besten, E. P. de Kluyver, G. Wensvoort, and **R. J. M. Moormann**

Central Veterinary Institute, Department of Virology, Lelystad, The Netherlands

Summary. Lelystad virus (LV) is an enveloped positive-stranded RNA virus, which causes abortions and respiratory disease in pigs. The complete nucleotide sequence of the genome of LV has been determined. This sequence is 15.1 kb in length and contains a poly(A) tail at the 3′ end. Open reading frames that might encode the viral replicases (ORFs 1a and 1b), membrane-associated proteins (ORFs 2 to 6) and the nucleocapsid protein (ORF7) have been identified. Sequence comparisons have indicated that LV is distantly related to the coronaviruses and toroviruses and closely related to lactate dehydrogenase-elevating virus (LDV) and equine arteritis virus (EAV). A 3′ nested set of six subgenomic RNAs is produced in LV-infected alveolar lung macrophages. These subgenomic RNAs contain a leader sequence, which is derived from the 5′ end of the viral genome. Altogether, these data show that LV is closely related evolutionarily to LDV and EAV, both members of a recently proposed family of positive-stranded RNA viruses, the Arteriviridae.

Introduction

A new pig disease, causing reproductive failures in sows and respiratory problems in piglets, was first observed in 1987 in the United States [1]. In subsequent years the disease spread through the United States to Canada. In late 1990 it appeared in Germany, after which it spread rapidly through Western Europe [2]. As the disease spread, it acquired more and more names, three of which are most prominent: swine infertility and respiratory syndrome (SIRS), porcine reproductive respiratory syndrome (PRRS), porcine epidemic abortion and respiratory syndrome (PEARS).

The causal agent of the disease, Lelystad virus (LV), was first isolated by Wensvoort et al. [3]. In experimentally induced infections, LV causes reproductive failure in sows, resulting in abortions and mummified, stillborn and weak piglets [3]. In addition, respiratory disease was observed in fattening pigs and in piglets. In the United States, the SIRS virus, which is antigenically and structurally closely related to LV, was identified as the causal agent of the disease [4].

LV is a small, enveloped, single-stranded RNA virus that replicates in vitro only in primary cultures of porcine alveolar macrophages. Titres in these macrophage cultures reach a maximum of $10^{6.5}$ $TCID_{50}$/ml. The buoyant density of LV ranges from 1.14 g/ml on a sucrose gradient to 1.19 g/ml in a caesium-chloride gradient. In ultra-thin sections of LV-infected macrophages, LV virions appear as 45 to 55 nm large spherical particles that consist of a 30 to 35 nm large nucleocapsid which is surrounded by a lipid bilayer membrane [5].

In this report we will present recent data concerning the organization and expression of the genome of Lelystad virus.

Genome organization

Recently, the nucleotide sequence of the genome of LV has been determined [6]. RNA was isolated from LV-infected alveolar macrophages and this was used to construct a cDNA library. This library was screened with a radioactive-labeled probe synthesized from LV genomic RNA. A set of overlapping LV-specific cDNA clones was isolated and a consecutive sequence of 15088 was obtained. A poly (A) sequence was present at the 3′ end of the genome. Eight open reading frames (ORFs) that might encode virus-specific proteins were identified (Fig. 1). ORF1a and ORF1b comprise about 80% of the viral genome and are predicted to encode the viral RNA polymerase. Their amino acid sequences contain elements conserved in RNA polymerases of the torovirus Berne virus (BEV) [7], equine arteritis virus (EAV) [8], lactate dehydrogenase-elevating virus (LDV) [9], and coronaviruses [10–12]. A putative serine protease domain and two papain-like protease domains were identified in the coding region of ORF1a (Fig. 2). The four characteristic domains, identified in ORF1b of the viruses mentioned above were also present in 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; (II) a cysteine- and histidine-rich Zinc finger domain; (III) a nucleoside triphosphate binding or helicase motif; and (IV) a conserved domain of unknown function.

ORF1a and ORF1b overlap over a small region of 16 nucleotides and occupy different reading frames. We assume that the expression of

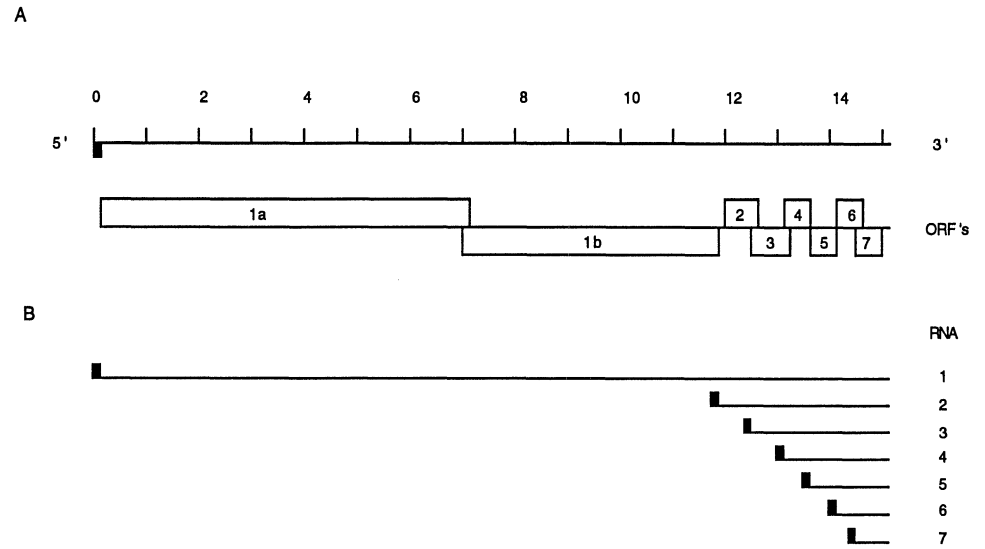

Fig. 1. **A** Organization of the LV genome. The ORFs, identified in the nucleotide sequence are shown. **B** shows the subgenomic set of RNAs, encoding the ORFs. The leader sequence is indicated by a solid box

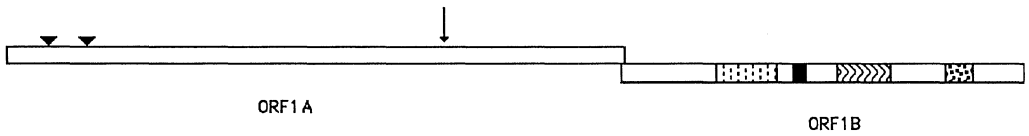

Fig. 2. Conserved domains in the polymerase encoding regions, ORF1a and ORF1b, of Lelystad virus. Two putative papain protease domains and a putative serine protease domain in ORF1a are indicated with black triangles and an arrow, respectively. The four conserved domains in ORF1b are indicated by shaded boxes. From left to right they are: polymerase motif, zinc finger domain, nucleoside triphosphate binding domain and a conserved domain of unknown function

ORF1b involves ribosomal frameshifting. 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 [10, 13], were identified in the overlapping region of ORF1a and ORF1b of LV [6].

The calculated sizes of the gene products of ORFs 2 to 6 range from 13.8–30.6 kDa and they show features reminiscent of membrane proteins [6]. They all contain putative N-linked glycosylation sites and N- and C-terminal hydrophobic sequences that may function as a signal sequence and a membrane anchor, respectively. The polypeptide encoded by ORF7 was extremely basic and the degree of identity with nucleo-

capsid proteins of other viruses (see below) suggests a similar function for this polypeptide.

Sequence comparisons of LV, LDV, and EAV

Comparison of the amino acid sequences encoded by the ORFs identified in the genome of LV with those of other viruses indicate that LV is distantly related to coronaviruses and toroviruses and closely related to LDV and EAV. Apart from the conserved regions in the polymerase encoding regions, mentioned earlier, no extensive identity between the amino acid sequences encoded by the ORFs of LV and coronaviruses or toroviruses was identified, using the FASTA sequence comparison program [14]. In contrast, the RNA polymerase-encoding regions (ORF1a and ORF1b) and the ORFs encoding the putative envelope proteins (ORFs 2 to 6) and the nucleocapsid protein (ORF7) of LV and LDV share a high percentage (29 to 67%) identical amino acids (Fig. 3). The amino acid sequences encoded by ORFs 1b to 7 of LV and LDV can be aligned without the introduction of large gaps. This is not the case for ORF1a of these viruses. Although the N-terminal amino acids (1–500) and the C-terminal amino acids (1 100–2 380) of ORF1a of LV and LDV are highly identical (46 and 40%, respectively), the sequence of the amino acids between were not significantly identical (as determined with FASTA). Comparison of the amino acid sequences encoded by ORF1a and ORF1b of LV and EAV also indicated that ORF1b is more conserved than is ORF1a. The amino acid sequences encoded by ORF1b of LV and EAV were 36% identical over the complete coding region. However, the amino acid sequence encoded by ORF1a of EAV is about 600 residues shorter than the amino acid sequence encoded by ORF1a of LV and is only 25% identical to the C-terminal portion of the amino acid sequence encoded by ORF1a of LV. Further experiments must be done to establish whether these differences in ORF1a also result in different functions of the RNA replicases of these viruses. Comparison of the amino acid sequences of ORFs 2 to 7 of LV and EAV showed that the amino acid sequences encoded only by ORFs 6 and 7 of these viruses share some amino acid identity (23 and 20%, respectively; Fig. 3).

In summary, the analysis of the amino acid sequences indicates that LV is more similar to LDV than to EAV. It is conjectured that such relationships indicate evolutionary development.

Subgenomic RNAs

Analogous to RNA synthesis during replication of EAV and LDV, multiple subgenomic RNAs are synthesized in LV-infected alveolar macrophages. A set of oligonucleotides was used in Northern blot

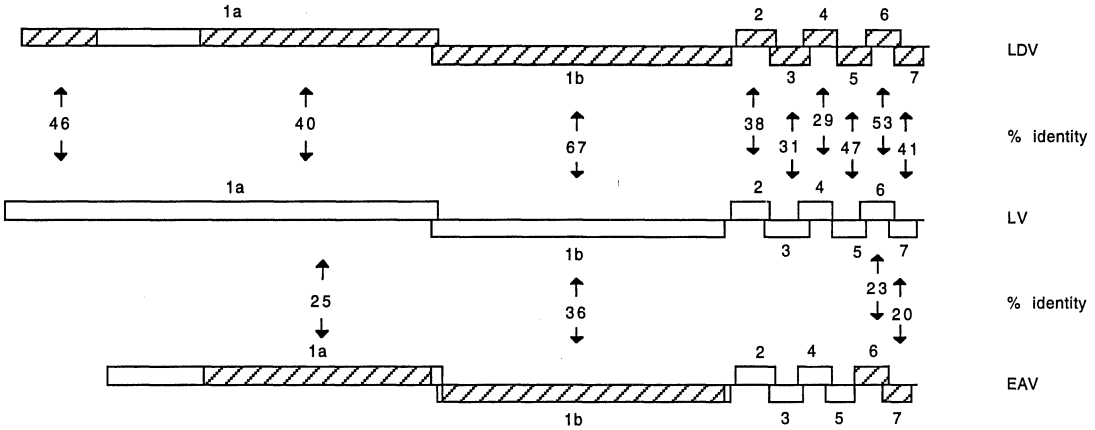

Fig. 3. Comparison of the viral proteins of LV, EAV and LDV. The percentage identity between the amino acid sequences encoded by ORFs of LV, LDV ([9], M. Brinton, pers. comm.) and EAV [8] are indicated. The boxes represent the open reading frames, identified in the nucleotide sequence of the genome of these viruses. The regions of the ORFs of LDV and EAV, which have identity with the ORFs of LV are shaded

hybridization analyses to identify and characterize LV subgenomic RNAs. These LV-specific oligonucleotides were located in the unique part of the various ORFs and at the extreme 5′ and 3′ end of the LV genome (Fig. 4A). We were able to identify six subgenomic RNAs by northern analyses of RNA isolated from LV-infected alveolar macrophages using oligonucleotide 25U101R, which is complementary to the 5′ noncoding region of the viral genome. This oligonucleotide hybridized to the 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. Hence, these six subgenomic RNAs possess a common leader sequence, which is derived from the 5′ end of the viral genome. The hybridization patterns of the other oligonucleotides demonstrated a correlation between ORFs 1a to 7 and RNAs 1 to 7 (Fig. 4B). Oligonucleotides specific for ORF1a and ORF1b (5U94R and 41U101R, respectively) hybridized only to the largest RNA, RNA1, oligoncleotide 60R219, specific for ORF2, hybridized to the two largest RNAs, RNA1 and RNA2, etc. This indicated that these subgenomic RNAs, together with RNA1 form a 3′ co-terminal nested set (Figs. 1 and 4B). Besides the seven RNAs, we observed two additional bands at about 1.5 and 4.3 kb. They appear to reflect aspecific binding of the probes to the ribosomal 18S and 28S RNA.

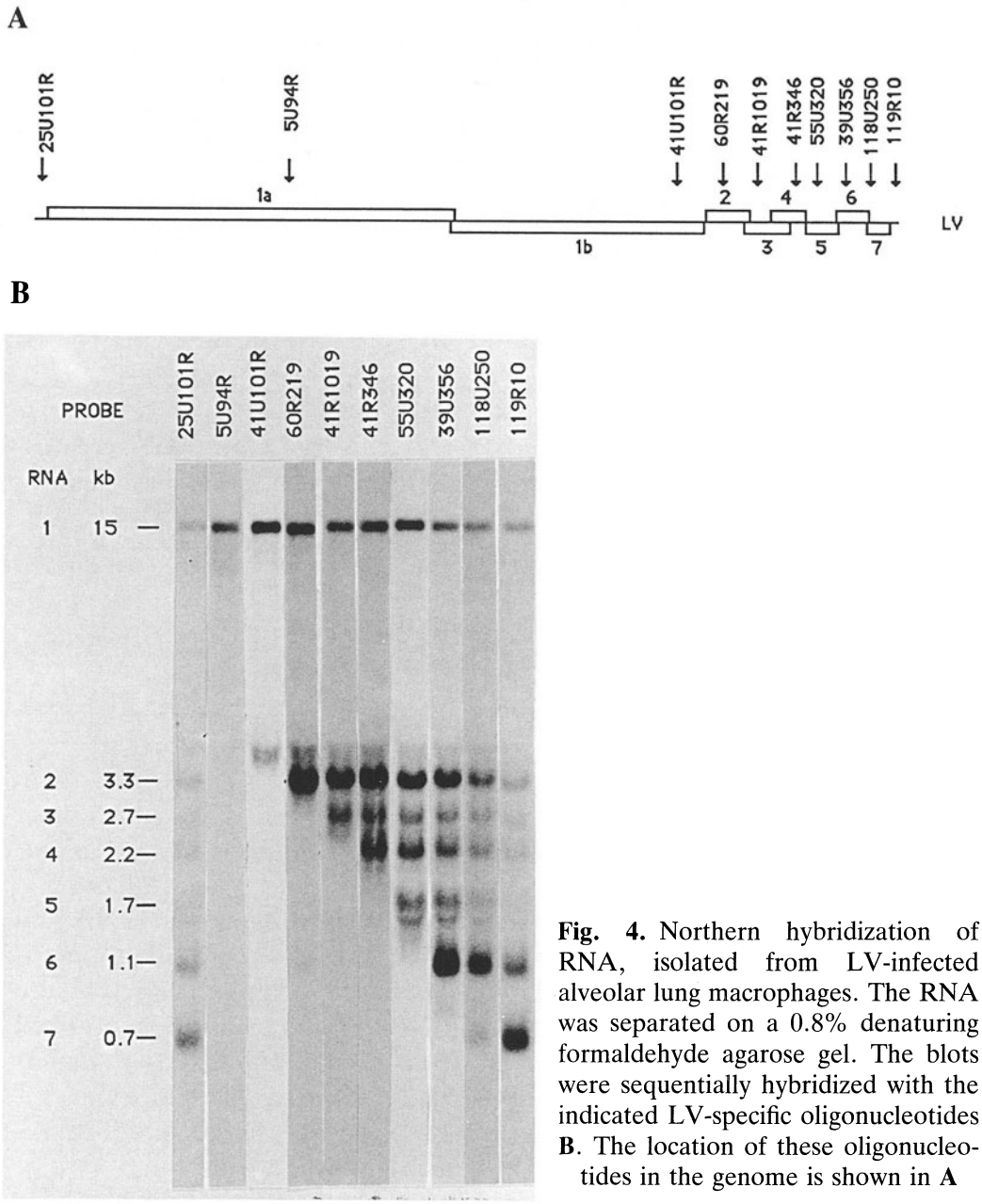

Fig. 4. Northern hybridization of RNA, isolated from LV-infected alveolar lung macrophages. The RNA was separated on a 0.8% denaturing formaldehyde agarose gel. The blots were sequentially hybridized with the indicated LV-specific oligonucleotides **B**. The location of these oligonucleotides in the genome is shown in **A**

A new family of positive-strand RNA viruses

The genomic data presented here indicate that LV is closely related to EAV and LDV and distantly related to coronaviruses and toroviruses. The size and organization of the genome of LV, LDV and EAV is similar, they all produce a coterminal nested set of 6–7 subgenomic RNAs and the ORFs of these viruses share extensive sequence homology. Preliminary sequence data indicate that simian hemorrhagic virus

(SHFV) is also closely related to LV, LDV, and EAV (E. Godeny, pers. comm.). Besides the similarities, mentioned above, these four viruses have other structural and biological properties in common [6, 15].

LDV and EAV have been classified in the family *Togaviridae* and SHFV in the family *Flaviviridae*. Recently, however, it has been suggested that LDV, EAV, SHFV, together with LV should be grouped in a new virus family, called the *Arteriviridae* [15]. We agree that these four viruses share enough characteristics and are sufficiently different from the togaviruses or flaviruses to warrant inclusion in a new virus family. However, of these viruses only EAV causes arteritis, whereas the other three viruses cause diseases with different clinical symptoms. We therefore feel that an appropriate name for this new group of viruses still has to be found. We are in favor of a name, related to the macrophage tropism of these viruses in their natural hosts.

Acknowledgements

We thank Drs. P. Plagemann and M. Brinton for providing additional sequence data. Part of this work was supported by Boehringer Ingelheim, Germany, and the Produktschap voor Vee en Vlees (PVV), the Netherlands.

References

1. Hill H (1990) Overview and history of mystery swine disease (swine infertility respiratory syndrom). In: Proceedings Mystery Swine Disease Committee Meeting, October 6, 1990, Denver, Colorado. Livestock Conservation Institute, Madison, pp 29–31
2. Paton DJ, Brown IH, Edwards S, Wensvoort G (1991) Blue ear disease of pigs. Vet Rec 128: 617
3. Wensvoort G, Terpstra C, Pol JMA, Ter Laak EA, Bloemraad RA, de Kluyver EP, Kragten C, van Buiten L, den Besten A, Wagenaar F, Broekhuijsen JM, Moonen PLJM, Zetstra T, de Boer EA, Tibben HJ, de Jong MF, van 't Veld P, Groenland GJR, van Gennep JA, Voets MTh, Verheijden JHM, Braamskamp J (1991) Mystery swine disease in the Netherlands: the isolation of Lelystad virus. Vet Q 13: 121–130
4. Collins JE, Benfield DA, Christianson WT, Harris L, Hennings J, Shaw DP, Goyal SM, McCullough S, Morrison RB, Joo HS, Gorcyca D, Chladek D (1992) Isolation of swine infertility and respiratory syndrome virus (Isolate ATCC VR 2332) in North America and experimental reproduction of the disease in gnotobiotic pigs. J Vet Diagn Invest 4: 117–126
5. Pol JAM, Wagenaar F (1992) Morphogenesis of Lelysted virus in porcine lung alveolar macrophages. Am Ass Swine Practit 4: 29
6. Meulenberg JJM, Hulst MM, de Meijer EJ, Moonen PLJM, den Besten A, de Kluyver EP, Wensvoort G, Moornmann RJM (1992) Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. Virology 192: 62–72
7. Snijder EJ, den Boon JA, Bredenbeek PJ, Horzinek MC, Rijnbrand R, Spaan WJM (1990) 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
8. den Boon JA, Snijder EJ, Chirnside ED, de Vries AAF, Horzinek MC, Spaan WJM (1991) Equine arteritis virus is not a togavirus but belongs to the coronavirus superfamily. J Virol 65: 2910–2920
9. Kuo L, Harty JT, Erickson L, Palmer GA, Plagemann PGWW (1991) A nested set of eight RNAs is formed in macrophages infected with lactate dehydrogenase-elevating virus. J Virol 65: 5118–5123
10. Boursnell MEG, Brown TDK, Foulds IJ, Green PF, Tomley FM, Binns MM (1987) Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J Gen Virol 68: 57–77
11. Brierly I, Boursnell MEG, Binns MM, Bilimoria B, Blok VC, Brown TDK, Inglis SC (1987) An efficient ribosomal frameshifting signal in the polymerase-encoding region of the coronavirus IBV. EMBO J 6: 3779–3785
12. Lee H, Shieh C, Gorbalenya AE, Koonin EV, Monica NL, Tuler N, Bagdzhadzhyan A, Lai MMC (1991) The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology 180: 567–582
13. Brierley I, Diggard P, Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57: 537–547
14. Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85: 2444–2448
15. Plagemann PGW, 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

Authors' address: Dr. J. J. M. Meulenberg, Central Veterinary Institute, Department of Virology, Houtribweg 39, 8221 RA Lelystad, The Netherlands.