Arch Virol (2000) 145: 2297–2323

Archives of Virology

© Springer-Verlag 2000
Printed in Austria

Full-length sequence of a Canadian porcine reproductive and respiratory syndrome virus (PRRSV) isolate*

S. Wootton[1], **D. Yoo**[1], and **D. Rogan**[2]

[1]Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada
[2]Vetrepharm Research, London, Ontario, Canada

Accepted May 22, 2000

Summary. Presently, one of the most economically important pathogens affecting swine is the porcine reproductive and respiratory syndrome virus (PRRSV). This virus is prevalent in herds throughout the world and continues to pose a significant threat as newer and more virulent disease phenotypes emerge. In this report we describe the full-length nucleotide sequence of a Canadian PRRSV isolate, designated PA8. A consecutive sequence of 15,411 nucleotides was obtained from a set of overlapping cDNA clones. In order to determine the extent of genetic variation among isolates recovered from swine in Canada and the US, as well as to understand the molecular mechanisms governing the evolution of PRRSV, the full-length sequence of PA8 was compared with that of two US isolates, VR2332 and 16244B. The genomic sequence of PA8 shared 98.2% and 99.2% identity with 16244B and VR2332, respectively. The untranslated regions (UTR) at the 5′ and 3′ ends of the genome were very well conserved. Notable exceptions include an eight nucleotide difference at the 5′ end of the 5′ UTR of VR2332 relative to PA8 and 16244B and a two nucleotide difference in the 3′ UTR of PA8 relative to VR2332 and 16244B. In contrast to PA8 and VR2332, 16244B possessed two nucleotide differences within the RNA pseudoknot structure of the ribosomal frameshift region between open reading frame (ORF)1a and ORF1b. Amino acid differences were distributed throughout the genome, however they appeared to be most extensive in Nsp1β and ORF5 of the nonstructural and structural coding regions, respectively, suggesting that the evolutionary pressure to conserve these viral genes is somewhat lower.

*The nucleotide sequence reported in this manuscript was deposited into GenBank database with the accession number AF176348.

Introduction

Porcine reproductive and respiratory syndrome virus (PRRSV) has been one of the most economically important pathogens of swine since its emergence in the late 1980's. PRRSV is a member of the family *Arteriviridae* in the order *Nidovirales*, together with equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV) and simian hemorrhagic fever virus (SHFV) [4, 45]. The virus is comprised of a 15 kilobase (kb) single-stranded positive sense RNA genome that is encapsulated within an enveloped, icosahedral-shaped nucleocapsid [5]. The replicase gene covers the 5′-terminal four-fifths of the genome and is composed of two overlapping open reading frames, ORF1a and ORF1b, which are expressed from the genomic RNA. ORF1b is expressed by means of a ribosomal frameshift mechanism [3]. With the exception of the newly characterized ORF2a [46], the structural genes encoded by ORF2 to 7 are expressed from the 5′ end of a co-terminal nested set of functionally monocistronic subgenomic mRNAs [28].

PRRSV infection is characterized by mild to severe reproductive failure in breeding swine and respiratory disease in young pigs. The disease syndrome was first recognized in the United States in 1987 [22] and shortly thereafter it emerged in Europe [34] where the causative agent was originally isolated [53]. Since their isolation, the prototype North American (VR2332) [5] and European [Lelystad virus (LV)] [53] strains of PRRSV have been extensively characterized. Despite the fact that North American and European strains of PRRSV do not manifest appreciable differences in disease phenotype [17], they are antigenically [31, 54, 55] and genotypically disparate [27, 32, 1]. In light of these differences, it has been proposed that two distinct species of PRRSV exist. Even within these distinct subgroups, however, there is precedence for strain divergence. Substantial sequence variation among North American isolates has been documented [2, 21, 30]. Such variation could be due to the innate errors of RNA polymerase or RNA recombination [51, 21]. Moreover, significant variation in the pathogenicity of American PRRSV isolates has been described [18, 26]. One such example is from a recent report describing a strain of PRRSV that causes severe encephalitis in neonatal pigs, a condition not normally associated with PRRSV infection [37]. Interestingly, this phenotype was particularly prevalent in herds that were using, or had recently used, a commercial modified-live vaccine.

In order to study the evolution of PRRSV within its natural host, it is essential to determine the full-length sequence of a collection of PRRSV isolates. The objective of this research, therefore, was to clone and sequence the full-length genome of a Canadian PRRSV isolate, PA8, and to compare it with PRRSV isolates, ATCC strain VR2332 [5, 32] and Michelle strain 16244B [1]. With the exception of the 5′UTR, the three North American isolates were identical with respect to their coding capacity. Nucleotide changes appeared to be tolerated in both of the UTRs, particularly in the 5′ end of the 5′ UTR where VR2332 differed from PA8 and 16244B by 8 nucleotides. Overall, PA8 differed from VR2332 and

16244B by a total of 51 and 104 amino acids, respectively, 74% of which were nonconservative in both cases.

Materials and methods

Virus and cells

PA8 was isolated during an outbreak in 1995 in Nederwood, Alberta, from a piglet displaying clinical signs indicative of PRRSV infection. Although pigs on this farm were not actively vaccinated against PRRSV, the farm had previously purchased vaccinated pigs and these were housed in the same area with the unvaccinated pigs. In September 1995, when pigs on the farm were experiencing severe reproductive failure and respiratory problems, serum samples were collected and submitted to the Animal Disease Research Institute (ADRI) in Lethbridge, Alberta, for diagnosis. The serum samples were used to inoculate Marc-145 cells [23] for virus isolation. The PA8 isolate was propagated in Marc-145 cells in Dulbecco's modified Eagle's medium supplemented with 5% fetal bovine serum (CanSera, Mississauga, ON), and a passage four stock was used in this study.

Enzymes and DNA cloning

Restriction enzymes and other molecular biological reagents were purchased from Amersham Pharmacia Biotech (Baie d'Urfe, PQ), New England Biolabs (Mississauga, ON), and GibcoBRL (Burlington, ON). Manipulation of DNA, including cDNA cloning, was performed according to standard procedures [41].

Preparation of viral RNA

Marc-145 cells were infected at a multiplicity of infection 10 and incubated for two days. Concomitant with the appearance of 80% cytopathic effects, cells were scraped into the medium and pelleted at 1500×**g** for 10 min in a bench top centrifuge (Beckman Model TJ-6, Palo Alto, CA). The clarified supernatant was removed and stored at 4 °C while the cell pellet was treated to one round of freeze-thaw at −80 °C. The virus-containing supernatant that was separated prior to freezing was added back to the pellet, and cellular debris was removed once again by centrifugation at 1500×**g** for 10 min. The clarified supernatant was layered onto a 50% (w/v) sucrose cushion and centrifuged at 25,000 rpm for 2 h in a SW28 rotor (Beckman Model XL-90). The virus-containing interface was removed and pelleted at 35,000 rpm for 2 h in TE buffer (10 mM Tris-HCl, pH 7.4, 1 mM Na_2EDTA, pH 8.0). Viral RNA was extracted using an equal volume of phenol, then phenol-chloroform and finally chloroform. RNA was treated with RQ DNase (Promega, Madison, WI) at 37 °C for 30 min to remove contaminating cellular DNA.

Reverse transcription and polymerase chain reaction (PCR) amplification

Oligonucleotide primers used to amplify the PA8 genome were designed based on sequence information available on GenBank for PRRSV strains VR2332 and LV. Table 1 illustrates the list of primers together with their sequence composition and relative nucleotide position with respect to the PA8 genome. For first strand cDNA synthesis, the viral RNA equivalent to 1/3 of a 100 mm tissue culture dish was incubated at 90 °C for 10 min with 0.5 μg of oligonucleotide. After chilling, 1 mM each of dCTP, dGTP, dTTP, and dATP, 10 mM DTT, 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$ and 200 units Superscript II reverse transcriptase (GibcoBRL) were added. The reaction was incubated for 1 h at 42 °C followed by heat inactivation at 70 °C for 15 min. One tenth the volume of the first-stand cDNA reaction

Table 1. Oligonucleotide primer sequences utilized to generate cDNA clones encompassing the full-length genome of the PA8 virus

Fragment	Primer pair	Primer sequence[a]	Location[d]
5′end	5′end-Fwd	5′-ATGCATGCTAATACGACTCACTATAG CGCCCGGGCAGGTGTTG-3′[b]	0–17
	5′end-Rv	5′-GCGGATCCAACTCCCTTAACGG-3′	1181–1203
3-3B	3-Fwd	5′-CTAAACGGACCTATCGTCG-3′	1150–1168
	3B-Rv	5′-AGGTGTCGATTACGCGTGGC-3′	2172–2191
3B-D	3B-Fwd	5′-GTTTGACCTGTACCTCCGTGG-3′	2103–2123
	D-Rv	5′-CTGCTTGATGACACGGACG-3′	3194–3212
D-4	D-Fwd	5′-GCATGAAGCTGAGGAAACC-3′	3135–3153
	4-RV	5′-ATGGAACAGCGGAAACCTTGACC-3′	4719–4741
4–5	4-Fwd	5′-CTGTATCTTGGCTGGAGCTTACGTGC-3′	4329–4354
	5-Rv	5′-GCATGTCCCATCATTCTCCACAGG-3′	6270–6293
4B-7	4B-Fwd	5′-CTTTGTGCCTTGCTTGCTGCC-3′	6193–6213
	7-Rv	5′-CTTTGGCAGTCAGTTCGC-3′	7613–7630
7–9	7-Fwd	5′-GAGTTCAATGGGAAGCTGC-3′	7550–7568
	9-Rv	5′-AGTTGTGTGCGACCTTGG-3′	10960–10977
9-ORF2	9-Fwd	5′-CATTCGATGTGGTTACATTGCATTTGCCC-3′	10630–10658
	ORF2-RV	5′-CCAACCGGCGATGGTGAAGC-3′	12190–12209
ORF2-ORF6	ORF2-Fwd	5′-CGGATCCATGAAATGGGGTCCATGCA-3′	12073–12091
	ORF6-RV	5′-GTCTGCTTGCCGTTGTTA-3′	14897–14914
ORF6-poly(dT)	ORF6-Fwd	5′-CGGATCCCAGCGGAACAATGGGGT-3′	14365–14381
	Poly(dT)	5′TTCTAGAATTCAGCGGCCGC$(T)_{30}N_{-1}$N-3′[c]	poly(A) tail

[a]Primers were designed based on sequence information available for VR2332 and LV

[b]Underlined nucleotides indicate nonviral sequence representing, in the order from 5′ to 3′, an *Sph* I site, a T7 RNA polymerase promoter sequence and a single G for initiation of transcription

[c]N_{-1} = G, A or C and N = G, A, C, or T

[d]Location corresponds to position within the PA8 genome

was used for second strand synthesis. PCR was performed for a total of 30 cycles using the hot start method in the presence of 0.3 μg each of forward and reverse primer and final concentrations of 1 mM dNTPs, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH 8.8), 2 mM $MgSO_4$, 0.1% Triton X-100 and 2 units of Vent Polymerase (NEB). The amplification cycle was as follows: denaturation at 95 °C for 30 sec, annealing at 55 °C to 65 °C for 30 sec and extension at 72 °C for 1 min per kb pair. PCR products were gel purified with GeneClean II (Bio101, Vista, CA). Purified products were phosphorylated with T4 polynucleotide kinase (NEB) according to manufacturer's instructions and cloned into the *Sma* I site of pGEM3zf(+) (Promega).

Determination of the 5′and 3′ end of the genome

cDNA clones of the 5′ end of the viral genome were obtained by reverse transcription with the primer 5′ end-RV (5′-GCGGATCCAACTCCCTTAACGG-3′) and PCR amplification with the sense primer 5′ end-Fwd (5′-ATGCATGCTAATACGACTCACTATAGCGCCCGGGCA-GGTGTTG-3′). The forward primer contains in the order from 5′ to 3′ an *Sph* I site, a T7 RNA polymerase promoter sequence, a single G for initiation of transcription and nucleotides 1 to

17 of the previously determined 5′ terminal sequence for either LV [29] or 16244B [1]. The purified PCR product was phosphorylated and cloned into the *Sma* I site of pGEM3zf(+).

cDNA clones comprising the 3′ end and poly(A) tail of the PA8 genome were obtained by reverse transcription of viral RNA with the marathon cDNA synthesis primer 5′TTCTAGAATTCAGCGGCCGC$(T)_{30}N_{-1}$N-3′ from Clontech (Palo Alto, CA), where N_{-1} = G, A or C and N = G, A, C, or T. The primer contains in the order 5′ to 3′ an *Eco* RI site followed by a *Not* I site, the former of which was used in cloning the cDNA. Reverse transcription was followed by PCR amplification with the marathon cDNA synthesis primer and ORF6-Fwd (5′CGGATCCCAGCGGAACAATGGGGT 3′), a sense primer corresponding to nucleotides 14897–14914 of the PA8 genome containing a *Bam* HI site. cDNAs were cloned into the *Eco* RI-*Bam* HI site of pGEM3zf(+).

Poly (A) test

In order to examine the number of poly(A)s, as well as to increase the number of templated A's obtained within a cDNA clone of the 3′ end of the genome, the protocol of Salles et al. [40] was performed with minor modifications. Briefly, approximately 1 μg of genomic RNA was heated to 70 °C for 5 minutes in the presence of 0.5 μg/μl phosphorylated oligo (dT) [$p(dT)_{12\text{-}18}$] (Amersham Pharmacia Biotech). A reaction mixture comprising 10 mM Tris-HCl (pH 8.3), 15 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM dNTPs, 0.5 mM ATP and 30 Weiss units of T4 DNA ligase (Pharmacia) was preheated to 42 °C and added to the RNA-oligo (dT) mixture in a final volume of 20 μl. The reaction was incubated at 42 °C for 30 min and cooled to 12 °C for 10 min. 200 ng of anchor primer (5′-GCGAGCTCCGCGGCCGCGTTTTTTTTTTTT-3′) was added and the reaction was incubated for 2 h at 12 °C to allow for ligation of the anchor primer to the extreme 3′ end of the poly(A) tail. 200 units of Superscript II reverse transcriptase were added and the reaction was incubated at 42 °C for 1 h followed by heat-inactivation at 70 °C for 30 min. A 2 μl aliquot was removed and added to a PCR mixture containing 300 ng each of ORF7-Fwd primer (5′-CGGATCCCCTTGTCAAATATGCCAA-3′) and either anchor or 3′UTR-Rv primer (5′-TTAATTTCGGCCGCATGGTTCT-3′) in a final volume of 100 μl. PCR products from two separate reactions were cloned and sequenced two times with universal primers in order to determine the exact number of poly(A)s within the clones.

Determination and analysis of nucleotide sequence

Individual cDNA clones were sequenced in both directions with universal as well as virus-specific primers using the Big Dye Dideoxy cycle sequencing kit (Perkin-Elmer, Norwalk, CT). Automated sequencing reactions were completed using a PE 9600 Thermocycler and a 377 Sequencer (Perkin-Elmer) at the Guelph Molecular Super Center, University of Guelph. For sequence analysis, the software package of the University of Wisconsin Genetics Computer group (GCG) was used [9]. The complete genomic sequence for the PA8 strain of PRRSV detailed in this report has been deposited into GenBank database with the accession numbers AF176348, AF066066 and AF066068.

Results and discussion

Genome structure

The entire genomic RNA of the PA8 strain of PRRSV was cloned using 10 sets of primer pairs by RT-PCR with a high fidelity polymerase possessing proofreading activity (Table 1). The amplified fragments were cloned into pGEM3zf(+) and sequenced in both directions (Fig. 1). When the sequence in both directions did

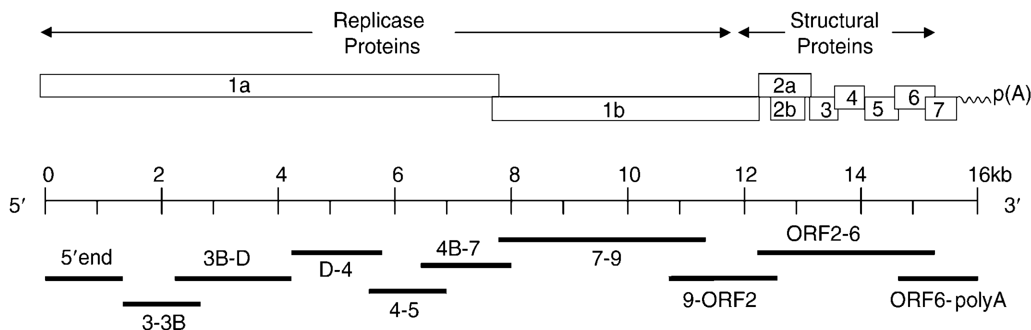

Fig. 1. Genome organization of the PRRSV isolate PA8. The 5′ four-fifths of the genome contains two large overlapping open reading frames (ORF1a and ORF1b). A ribosomal frameshift region is located at the junction between ORF1a and ORF1b. The 3′one-fifth of the genome contains six open reading frames, two of which (ORF4 and ORF5) do not overlap. Additionally, a second open reading frame (ORF2b) is located entirely within ORF2a. The relative length and position of the cDNA clones which were used in the sequence analysis of the PA8 genome are depicted by solid black bars and are named according to the primer pairs used for amplification as indicated in Table 1

not conform, samples were re-sequenced. The sequence data were assembled into one contiguous sequence of 15, 411 nucleotides, excluding the poly(A) tail. The nucleotide sequence and genome organization of PA8 were similar to that of other arteriviruses (Table 2). PA8 possessed a 5′UTR of 189 nucleotides and a 3′UTR of 151 nucleotides. Two large overlapping open reading frames, ORF1a and ORF1b, of 7509 and 4389 nucleotides, respectively, followed the 5′UTR. Six additional overlapping ORFs were identified immediately downstream of ORF1b. These ORFs were able to encode the viral structural proteins of PRRSV for the minor membrane glycoproteins GP2a, GP3, GP4, a major membrane glycoprotein GP5, a membrane associated protein M, and the nucleocapsid protein N. A small internal open reading frame capable of encoding a polypeptide of 73 amino acids was identified within the ORF2a coding region and was designated ORF2b (Fig. 2). The homolog of ORF2b is found in all other members of the arterivirus family and was recently shown, with the use of an EAV infectious clone, to be an essential component for virus replication [46]. The organization of this region of the genome differs significantly among members of the arterivirus family. Moreover, even the North American and European species of PRRSV differ considerably within this region. The upstream leader body junction, presumably utilized for the production of subgenomic mRNA2, is 20 nucleotides upstream of ORF2a in PA8 while it is located 38 nucleotides upstream of ORF2a in LV. Furthermore, there is only one nucleotide separating ORF1b from ORF2a in PA8, whereas there are 10 in the case of LV.

PA8 represents the first Canadian PRRSV isolate for which the full-length sequence has been determined. During the course of this study, the full-length genomic sequence of two American PRRSV isolates, VR2332 and 16244B, became available. The full-length sequence of the prototype European isolate, LV was

Table 2. The genome organization and the predicted polypeptides encoded by the PA8 viral RNA

mRNA	Locus/ ORF	Cleavage product	Nucleotide position	Size in nucleotides	Amino acid position	Size in amino acid	Motif reference
	Total length [excluding poly(A)]		15411				
	5′UTR		1–189	189			
1	ORF1a			7509		2503	
		Nsp1α	189–687	498	1–166	166	papainlike cysteine protease (PCPα)[a]
		Nsp1β	688–1335	648	167–382	216	papainlike cysteine protease (PCPβ)[a]
		Nsp2	1336–4278	2943	383–1363	981	chymotrypsin-like cysteine protease (CP)[b]
		Nsp3	4279–5616	1338	1364–1809	446	hydrophobic domains
		Nsp4	5617–6228	612	1810–2013	204	serine protease (SP)[c]
		Nsp5	6229–6738	510	2014–2183	170	hydrophobic domains
		Nsp6	6739–6786	48	2184–2199	16	
		Nsp7	6787–7563	777	2200–2458	259	
		Nsp8	7564–7698	135	2459–2503	45	
1	ORF1ab			11880		3960	
		Nsp9[h]	7564–9617	2055	2459–3143	685[i]	RdRp (SDD[d])
		Nsp10	9618–10940	1323	3144–3584	441	Zinc-finger/Helicase[e]
		Nsp11	10941–11609	669	3585–3807	223	Coronavirus-like[f]
		Nsp12	11610–12068	459	3808–3960	153	
2	ORF2a		12073–12843	771		256	
2	ORF2b		12078–12299	222		73	
3	ORF3		12696–13460	765		254	
4	ORF4		13241–13777	537		178	
5	ORF5		13788–14390	603		200	
6	ORF6		14375–14899	525		174	
7	ORF7		14889–15260	372		123	Nucleolar localization signal[g]
	3′UTR		15261–15411	151			
	poly(A)		15412–15484	72			

References for protein motifs: [a]Dougherty and Semler [11];
[b]Snijder et al. [43];
[c]Snijder et al. [44];
[d]Brierley et al. [3];
[e]Koonin et al. [24];
[f]Kadare and Haenni [20];
[g]Rowland et al. [39]
[h]Amino acid numbering for ORF1b corresponds to the ORF1ab fusion protein
[i]The amino terminus for Nsp9 is predicted to be the same as that of Nsp8

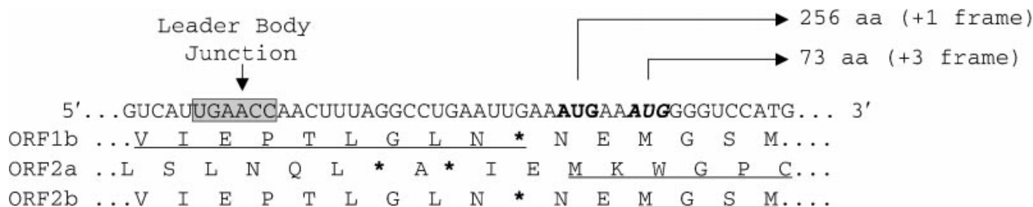

Fig. 2. Organization of the biscistronic mRNA2 predicted to encode the ORF2a and ORF2b protein products. The boxed sequence, UGAACC, represents the only upstream element capable of functioning as the leader-body junction during the transcription of mRNA2. Initiation of translation from the first AUG (bold) would result in the production of a 256 amino acid polypeptide (ORF2a), while initiation of translation from the second AUG (bold-italics) in the +3 frame would produce a 73 amino acid polypeptide (ORF2b). Open reading frames are underlined and translational stops are indicated by asterisks

Table 3. Nucleotide identity of the full-length genomic sequences of North American and European type PRRSV

	PA8	VR2332	16244B	LV[a]
PA8	–	99.2[b]	98.2	60.5
VR-2332	120[c]	–	98.4	60.3
16244B	287	238	–	60.5
LV[d]	5960	5990	5960	–

[a]Sequence for PA8 is presented in this study, VR2332 was reported by Nelson et al. [32], 16244B was reported by Allende et al. [1] and LV was reported by Meulenberg et al. [28]

[b]Percentages of nucleotide indentity are presented above the diagonal

[c]Absolute distances expressed as number of polymorphic sites are shown below the diagonal

[d]Absolute distances between LV and VR2332/16244B/PA8 are expressed as the relative number of polymorphic sites within the 15088 nucleotide sequence of LV

determined previously [28]. Therefore, it was of interest to compare the genomic sequence of PA8 with that of the three other full-length sequences available in order to evaluate the relative evolutionary pressure on viral genes to either conserve or change their amino acid sequence. Table 3 illustrates the relative nucleotide homology shared between these four strains of PRRSV. The three North American strains of PRRSV shared over 98% sequence identity with each other and only 60% sequence identity with LV, which suggests that the evolutionary pressures and constraints within North America differ from those in Europe.

Comparison of the 5′UTR

The 5′UTRs of VR2332, 16244B and LV were 96.3%, 98.9% and 44% identical to the 5′UTR of PA8, respectively (Fig. 3A). The 211 nucleotide 5′UTR of LV was considerably longer than that of VR2332 at 190, and PA8 and 16244B at

A

```
   PA8 : -CGCCCGGGC AGGTGTTGGC TCTATGCCT- TGGCATTTGT ATTGTCAGGA :48
VR2332 : TATGACGTAT AGGTGTTGGC TCTATGCCT- TGGCATTTGT ATTGTCAGGA :49
16244B : -CGCCCGGGC AGGTGTTGGC TCTATGCCT- CGGCATTTGT ATTGTCAGGA :48
    LV : -GGGTATTCC CCCTACATAC ACGACACTTC TAGTGTTTGT GTACCTTGGA :49

   PA8 : GCTGTGACCA TTGGCACAGC CCAAAACTTG CTGCACAGAA ACACCCTTCT :98
VR2332 : GCTGTGACCA TTGGCACAGC CCAAAACTTG CTGCACAGAA ACACCCTTCT :99
16244B : GCTGTGACCA TTGGCACAGC CCAAAACTTG CTGCACAGAA ACACCCTTCT :98
    LV : GGCGTGGGTA CAGCCCCGCC CCACCCCTTG GCCCCTGTTC TAGCCCAACA :99

   PA8 : GTGATAGCCT CCTTCAGGGG AGCTTAGG-- ----GTTTGT CCCTAGCACC :142
VR2332 : GTGATAGCCT CCTTCAGGGG AGCTTAGG-- ----GTTTGT CCCTAGCACC :143
16244B : GTGACAGCCT CCTTCAGGGG AGCTTAGG-- ----GTTTGT CCCTAGCACC :142
    LV : GGTATCCTTC TCTCTCGGGG CGAGTGCGCC GCCTGCTGCT CCCTTGCAGC :149

   PA8 : T--------- -------TGC TTCCGGAGTT GCACTGCTTT ACGGTCTCTC :176
VR2332 : T--------- -------TGC TTCCGGAGTT GCACTGCTTT ACGGTCTCTC :177
16244B : T--------- -------TGC TTCCGGAGTT GCACTGCTTT ACGGTCTCTC :176
    LV : GGGAAGGACC TCCCGAGTAT TTCCGGAGAG CACCTGCTTT ACGGGATCTC :199

   PA8 : CACCCCTTTA ACC :189
VR2332 : CACCCCTTTA ACC :190
16244B : CACCCCTTTA ACC :189
    LV : CACCC-TTTA ACC :211
```

Fig. 3 *(continued)*

B

```
PA8    : TGGGCTGGCA TTCTTGAGGC ATCTCAGTGT TTAAATTGGA AGAATGTGTG :15310
VR2332 : TGGGCTGGCA TTCTTGAGGC ATCTCAGTGT TTGAATTGGA AGAATGTGTG :15311
16244B : TGGGCTGGCA TTCTTGAGGC ATCTCAGTGT TTGAATTGGA AGAATGTGTG :15310
LV     : ---------- ---------- ---------- --------TT TGACAGTCAG :14986

PA8    : GTGAATGGCA CTGATTGACA TTGTGCCTCT AAGTCACCTA TTCAATTAGG :15360
VR2332 : GTGAATGGCA CTGATTGACA TTGTGCCTCT AAGTCACCTA TTCAATTAGG :15361
16244B : GTGAATGGCA CTGATTGACA TTGTGCCTCT AAGTCACCTA TTCAATTAGG :15360
LV     : GTGAATGGCC GCGAT-GGCG TGTGGCCTCT GAGTCACCTA TTCAATTAGG :15035

PA8    : GCGACCGTGT GGGGGTGAGA TTTAATT--G GCGAGAACCA TGCGGCCGAA :15408
VR2332 : GCGACCGTGT GGGGGTGAGA TTTAATT--G GCGAGAACCA TGCGGCCGAA :15409
16244B : GCGACCGTGT GGGGGTGAGA TTTAATT--G GCGAGAACCA TGCGGCCGAA :15408
LV     : GCGATCACAT GGGGGTCATA CTTAATTCAG GCAGGAACCA TGTGACCGAA :15085

PA8    : ACT :15411
VR2332 : ATT :15412
16244B : ATT :15411
LV     : ATT :15088
```

Fig. 3. Alignment of the 5′ UTR (**A**) and 3′ UTR (**B**) of PRRSV isolates PA8, VR2332, 16244B and LV. Deletions are indicated by dashes. Nucleotides that are completely conserved are shaded. Numbers on the right indicate nucleotide position from the 5′ terminus of the genome. The six nucleotide junction sequence at the 3′ end of the 5′ UTR is boxed

189 nucleotides. The 16244B 5′UTR differed from PA8 by only 2 nucleotides, both of which were T to C mutations conceivably due to pyrimidine wobble. The 7 nucleotides that differed between VR2332 and both PA8 and 16244B were located at the 5′-most end of the 5′UTR. Interestingly, this region, deemed the "5′-leader start" motif, was found to be one of three conserved domains shared between North American and European strains of PRRSV [33]. We attempted to amplify the 5′ end of PA8 using primers specific for the 5′ end of either VR2332 or 16244B. The 16244B specific primer consistently produced a product while, despite our efforts, this same region of PA8 could not be amplified using the VR2332 specific primer. This suggested that the 5′ end of PA8 was more similar to 16244B. The leader-body junction sequence (UUAACC) as well as the string of 8-11-9 absolutely conserved nucleotides, described by Oleksiewicz et al. [33], were identical among all four strains.

Comparison of 3′UTR

The 3′UTR of all three North American isolates was 151 nucleotides long, which differs considerably from LV, whose 3′UTR is only 114 nucleotides in length (Fig. 3B). The 3′ UTR of PA8 differed from both VR2332 and 16244B by two nucleotides, one of which was located 2 nucleotides upstream of the beginning of the poly(A) tail region. Length aside, the 3′ UTR displayed a higher level of homology within the aligned regions than did the 5′ UTR. However, in both the 3′ UTR and 5′ UTR, the degree of conservation was highest towards the 3′ end which suggests that these regions may encode important *cis*-acting elements. In order to determine the approximate number of poly(A)s and in doing so, to increase the length of the templated poly (A) tail within a cDNA clone of the 3′ end of the genome, the Poly(A) Test (PAT) described by Salles et al. [40] was employed. PAT involves first saturating the poly(A) with phosphorylated oligonucleotides $(dT)_{12-18}$ and then ligating these oligonucleotides together at 42 °C. This process generates a nearly full-length complementary copy of the poly(A) tail, leaving only a short 3′ overhang of oligo(A) as a result of the instability of short A/dT hybrids at 42 °C [40]. An oligo(dT)-anchor primer was then added in excess with respect to oligo$(dT)_{12-18}$ and the temperature lowered to 12 °C to favour the hybridization and ligation of the oligo(dT)-anchor to the extreme 3′ end of the poly(A) tail. After ligation, the sample was reverse transcribed and the poly(dT)-anchor stretch that had annealed to the extreme 3′ end of the poly(A) tail served as the reverse primer. For determination of the PA8 poly(A) tail length, an ORF7 specific forward primer (ORF7-Fwd) was used. A second reaction using the same forward primer in combination with a reverse primer (3′UTR-RV) specific for the extreme 3′ end of the 3′UTR was performed to illustrate the molecular weight contribution of the poly(A) tail region (Fig. 4A). Results from the PAT experiment are shown in Fig. 4B. The product migrating at 520 bp was amplified using the ORF7-Fwd and the 3′UTR-RV primer-pair (lane 2) while the sample migrating at 592 bp was amplified using the ORF7-Fwd and the poly(dT)-anchor primer-pair (lane 3). The difference in size between the two

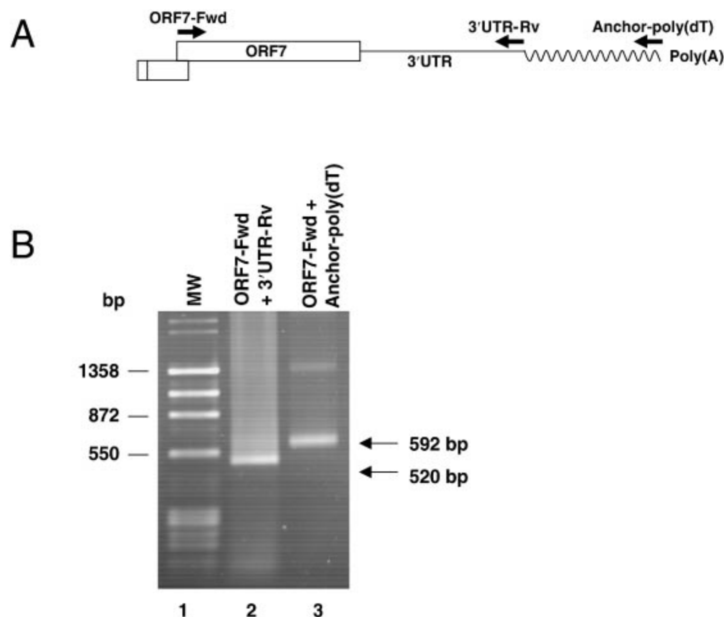

Fig. 4. Determination of the length of the PA8 poly(A) tail. **A** Diagram of the 3′ end of the PA8 genome depicting ORF7, the 3′ UTR and the poly(A) tail. The name and location of primers used are indicated by arrows. **B** Agarose gel electrophoresis of PCR products. *1* Molecular weight markers; *2* product amplified using ORF7-Fwd and 3′ UTR-Rv primers; *3* product amplified using ORF7-Fwd and oligo(dT)-anchor primers. The difference in size between the two PCR products is due to the length of the poly(A) tail

products represents the length of the poly(A) tail. The 592 bp product (lane 3) was cloned two times from two separate reactions. Each of these clones was sequenced twice and the length of the poly(A) tail was determined to be 72, thereby extending the total number of templated nucleotides. Although this length does not necessarily reflect that of a virus in a natural pig infection, it does provide some information about poly(A) tail length during infection. Furthermore, this information may be important to consider when constructing an infectious cDNA clone of PRRSV.

Comparison of ORF1a

ORF1a of PA8 encodes a 2503 amino acid polyprotein that is predicted to be cleaved at 8 sites, yielding a number of processing intermediates and 9 end products (Fig. 5) [52]. These include Nsp1α, Nsp1β [8], and Nsp2 to Nsp8 [52, 42] (Table 2). The N-terminal Nsp1α and Nsp1β each contain a papainlike cysteine protease domain which serves to cleave the Nsp1α/1β and Nsp1β/2 junctions, respectively (Fig. 5) [8]. In PA8, the C76-H147 and C270-H339 residues, which likely form the catalytic dyads of PCPα and PCPβ respectively, are conserved as are the predicted cleavage sites. Based on the size of the Nsp1α cleavage product [8], and amino acid homology between LDV [12] and PRRSV, the cleavage site for Nsp1α/1β was predicted to be located between amino acids 163–166 [8]. In

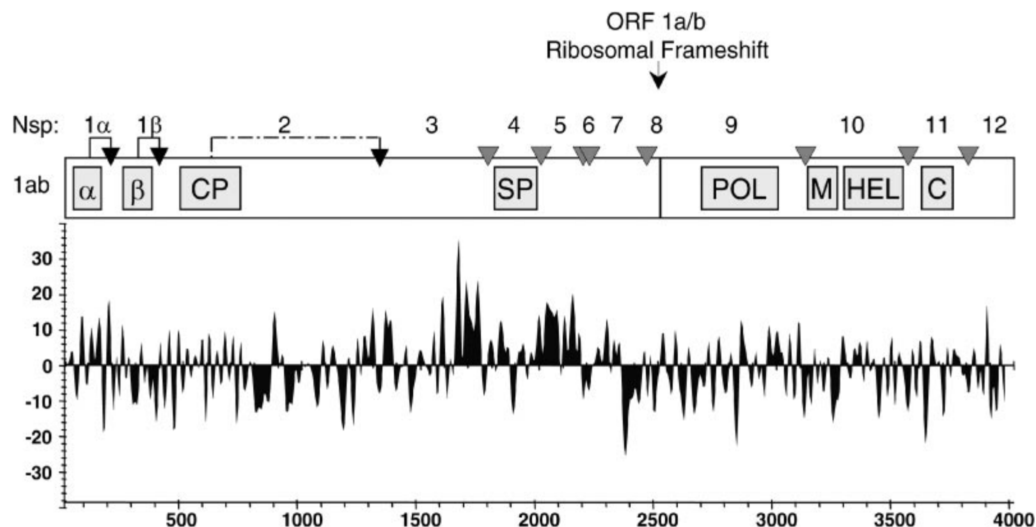

Fig. 5. Putative processing scheme of the PA8 ORF1ab replicase protein. The nonstructural protein products (Nsp1α to Nsp12) are shown above the schematic diagram and conserved domains are indicated by boxes: α/β papainlike cysteine protease; *CP* putative cysteine protease; *SP* putative serine protease; *POL* putative RNA-dependent RNA polymerase; *M* putative metal-binding domain; *HEL* putative RNA helicase; *C* conserved nidovirus domain. The putative cleavage site for the Nsp1α papainlike cysteine protease is shown by a black arrow as is the cleavage site for the Nsp1β papainlike cysteine protease. The hatched arrow and grey arrowheads indicate potential cleavage sites for the putative CP and SP, respectively. The hydropathy plot generated for ORF1ab is shown below the schematic diagram. Values above the axis indicate hydrophobicity and values below indicate hydrophilicity. Amino acids are plotted along the y-axis

[1], the Nsp1α/1β cleavage site was predicted to be Q166/R167. The Nsp1β/2 cleavage site appears to be between Y382/G383 [8].

The Nsp1α region was highly conserved among the three North American strains examined. PA8 and VR2332 shared 99.4% nucleotide (Fig. 6A) and 100% amino acid identity (Fig. 7) while PA8 and 16244B shared 98.8% nucleotide and 98.8% amino acid identity. Both of the amino acid differences in 16244B were conservative in nature (Fig. 7). PA8 Nsp1β was the most variable of the ORF1a processing products, sharing 98.9% and 96.8% nucleotide homology with VR2332 and 16244B, respectively (Fig. 6A). VR2332 and 16244B differed from PA8 in this region by 5 and 10 amino acids, respectively (Fig. 7). In both cases, 80% of these changes were nonconservative.

Based on homology with EAV, Nsp2 is predicted to encode a cysteine protease (CP) [43], the functional domain of which appears to be located between amino acids 435 and 506 in the amino terminus. A proline-rich region located between amino acids 801 and 950 was also identified. The Nsp2 protein of arteriviruses appears to be highly variable. This is particularly true for the Nsp2 of LV which possesses three large, centrally located deletions of 14, 78 and 31 amino acids relative to North American strains of PRRSV. The LDV and EAV Nsp2 proteins are only 733 and 401 residues long, respectively [7, 12], while the North American

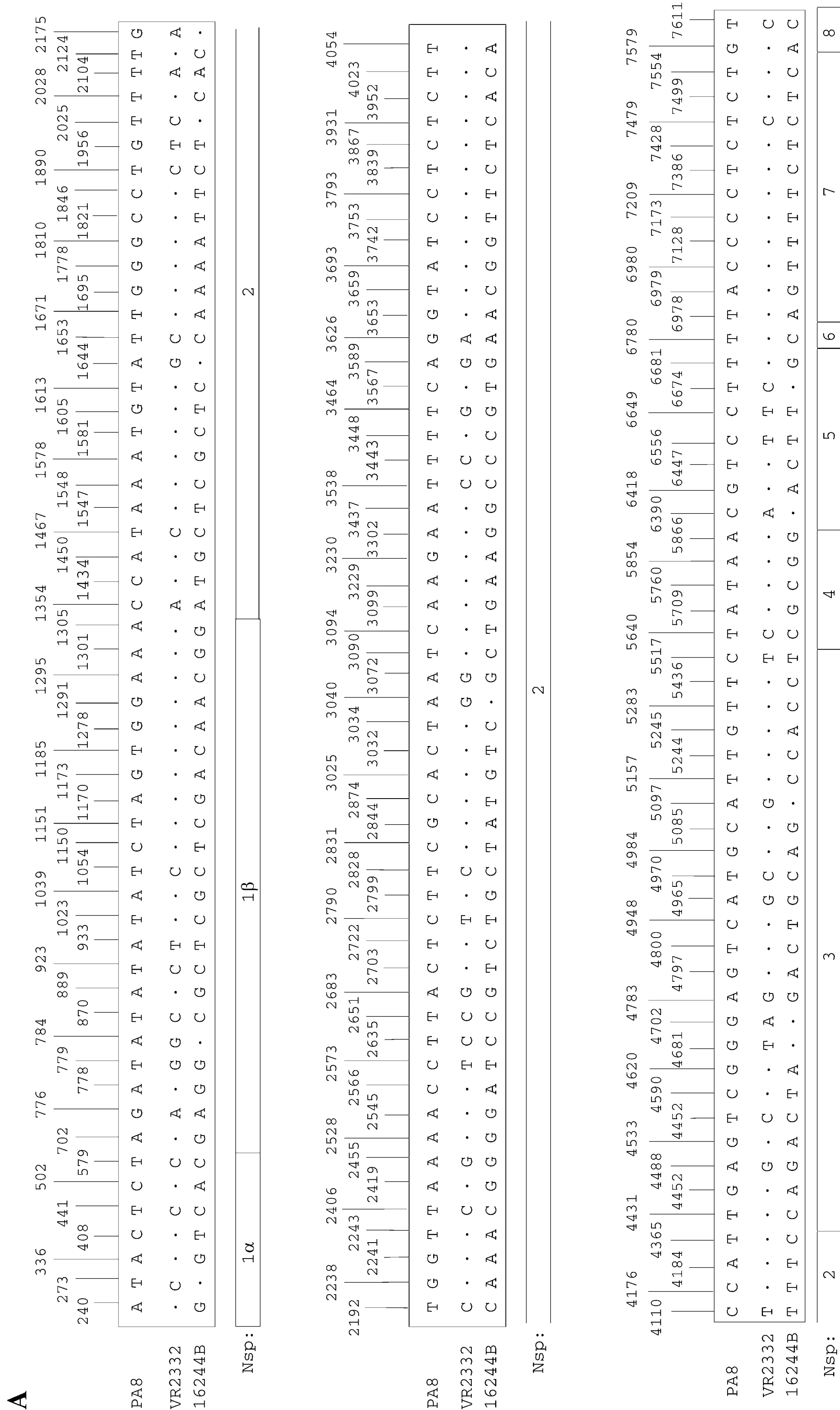

Fig. 6 *(continued)*

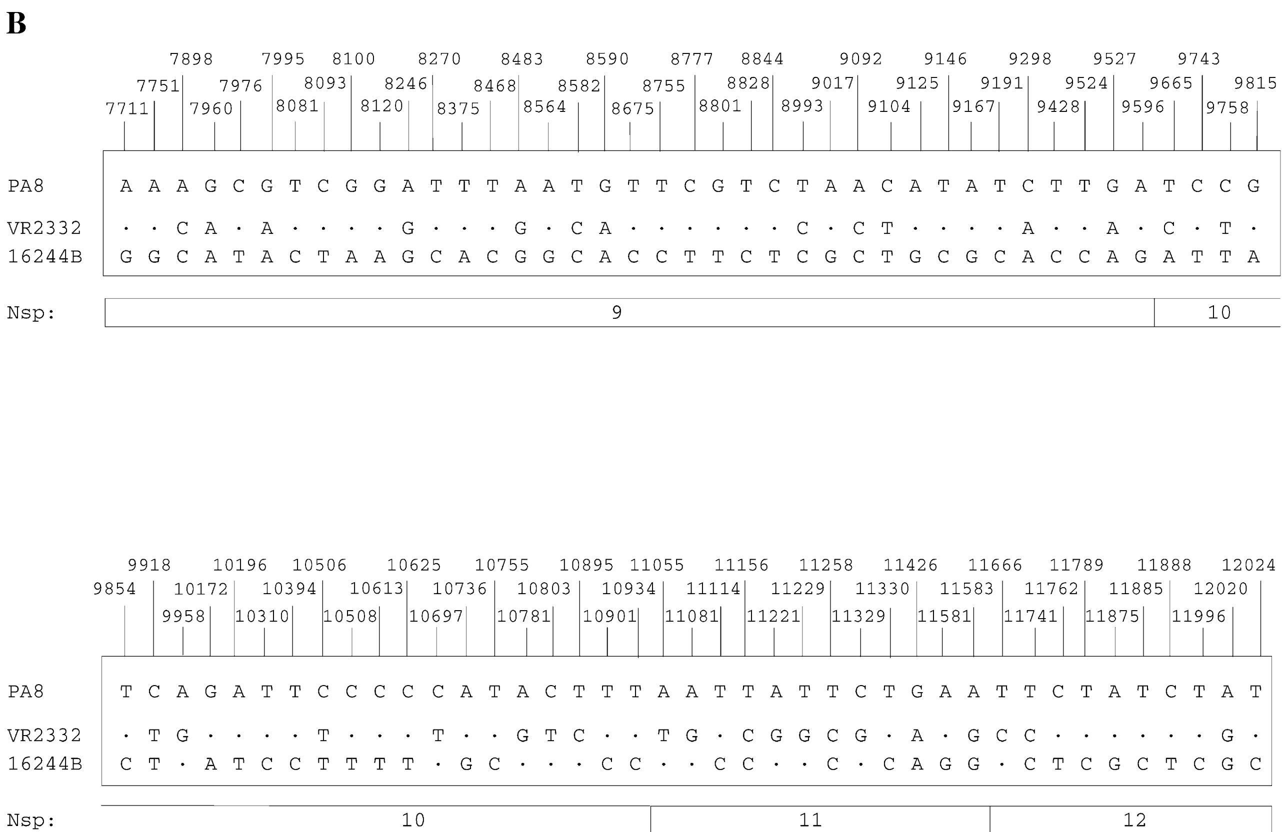

Fig. 6. Nucleotide changes within the ORF1a (**A**) and ORF1b (**B**) coding region of VR2332 and 16244B relative to PA8. Numbers at the top indicate location within the PA8 genome where nucleotide differences occur. Bars representing the putative cleavage products of ORF1a and ORF1b are positioned below such that the number of nucleotide changes that fall within these regions is shown

PRRSV Nsp2 is 981. Despite its variability in length, Nsp2 is, however, well conserved at the amino- and carboxy-termini. In the case of PA8, VR2332 differs by only 14 amino acids (all of which were conservative) while 16244B differs by 43 resulting in 98.6% and 95.6% homology, respectively (Fig. 7). Of the 43 amino acid differences in 16244B, only 7 were conservative.

The Nsp3 proteins of VR2332 and 16244B differed from PA8 by 4 amino acids each, resulting in a 99.1% amino acid identity. Three and two out of the four amino acid changes were nonconservative in VR2332 and 16244B, respectively (Fig. 7). The function of Nsp3 is not known, however it possesses a number of hydrophobic regions and a putative transmembrane domain that is believed to be involved in anchoring the replicase complex within intracellular membranes [47, 35].

The PA8 Nsp4 protein was the one of the most highly conserved ORF1a processing products with only 1 nucleotide difference in VR2332 and 5 in 16244B (Fig. 6A). PA8 shared 100% amino acid identity with VR2332 and 99% identity with 16244B due to 2 nonconservative amino acid changes (Fig. 7). Based on homology with EAV, the PRRSV Nsp4 encodes a putative 3C-like serine protease (SP), the functional domain of which appears to be located between amino acids 1841 to 1947 [44]. The catalytic triad is likely comprised of H1848, D1873 and S1927 [44]. A number of Nsp4 cleavage sites have been predicted, these are: E1809/G1810, E2013/G2014, E2183/G2184, E2199/S2200, and E2458/A2459 which would ultimately serve to release Nsp3, 4, 5, 6 , 7 and 8, respectively [44, 48, 50] (Fig. 5).

In EAV, viral RNA synthesis is localized on membrane-bound complexes [47]. Biochemical studies clearly implicated the hydrophobic domains of Nsp2, Nsp3 and especially Nsp5 in membrane association [47]. The hydropathy profile of the PRRSV Nsp5 indicates that it too is highly hydrophobic (Fig. 5), which suggests that it may serve a similar function during PRRSV infection. The Nsp5 of PA8 differed from VR2332 and 16244B by 1 and 2 amino acids respectively (Fig. 7). None of these changes interfered with the hydrophobicity of the protein. The Nsps 6, 7 and 8 of PA8 were 100% identical to VR2332 at the amino acid level and differed by 3 nonconservative amino acids in 16244B. The function of processing products Nsp 6 to 8 remains to be determined.

Overlapping region between ORF1a and ORF1b

The RNA element containing a slippery sequence and a sequence that is predicted to form an RNA pseudoknot structure in the ORF1a/1b overlap region is essential for the expression of ORF1b via the mechanism of ribosomal frameshifting [3]. The heptanucleotide slippery sequence, UUUAAAC, located 3 nucleotides upstream of the ORF1a stop codon is strictly conserved between both species of PRRSV. Conversely, the sequence forming the stem-loop structure downstream of the slippery sequence, appears to be able to accommodate mutations. The PA8 and VR2332 ORF1a/ORF1b junction sequences are identical, however, the 16244B junction sequence differs by two nucleotides due to purine wobble, one

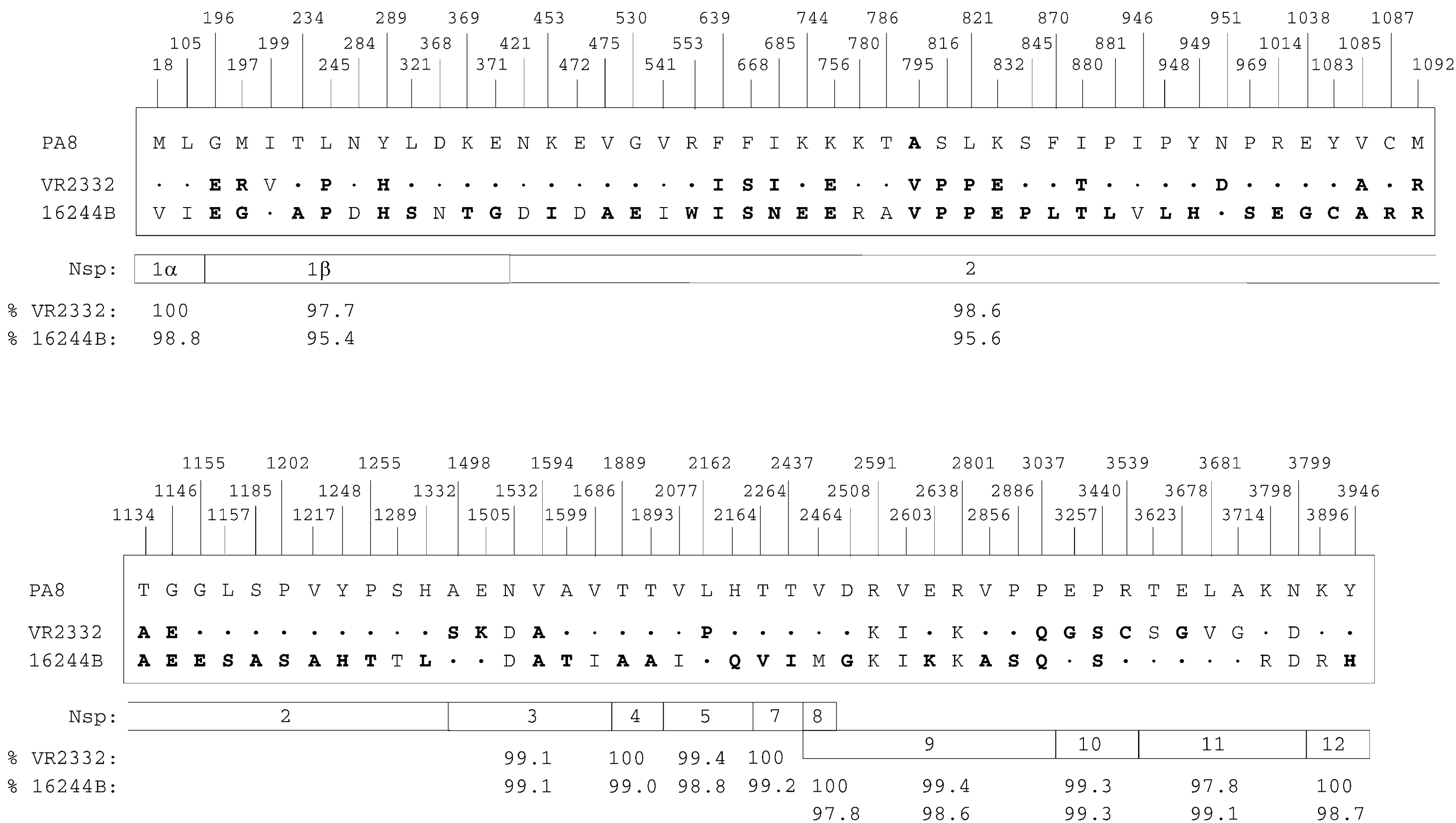

Fig. 7. Amino acid changes within the ORF1ab coding region of VR2332 and 16244B relative to PA8. Numbers at the top indicate location within the PA8 genome where amino acid differences occur. Bars representing the putative cleavage products of ORF1a and ORF1b are positioned below such that the number of amino acid changes that fall within these proteins is shown. Amino acid changes were considered to be conservative if they fell into one of these six groups: [L, I, V, M]; [A, G]; [S, T]; [D, E, N, Q]; [K, R] or [F, Y, W]. Nonconservative amino acid changes are shown in bold. Numbers at the bottom indicate the percent identity between PA8 and VR2332 or 16244B

of which does not appear to affect base-pairing within the stem region (Fig. 8). A change from A to G at position 7711 within the large stem region of the 16244B pseudoknot does however alter base pairing, resulting in a change from A-U as in VR2332 and PA8 to G-U. This change results in a slight reduction of the free energy from $\Delta G^\circ = -29.9$ as in PA8 to -29.0 kJ according to the mfold program [56]. Interestingly, LV possesses the same nucleotide change, however a compensatory nucleotide change was made such that a G-C base pair results. LV differs at 7 other positions, one of which is located within the small stem and results in a base pair change from U-A to U-G (Fig. 8).

Comparison of ORF1b

ORF1b of PA8, like that of VR2332, 16244B and LV, encodes for a 1463 amino acid protein. This value however, does not take into consideration that additional amino acids are added when the amino terminus of Nsp9 is generated by a combination of cleavage at the amino terminus of Nsp8 and ribosomal frameshifting. ORF1b is synthesized as a fusion with ORF1a via a ribosomal frameshifting mechanism which produces a polyprotein of approximately 3960 amino acids [3, 19]. The ORF1b part of the replicase contains a set of highly conserved functions which are essential for viral RNA replication and mRNA transcription [7, 49]. Based on homology with EAV, the PRRSV ORF1b protein contains three putative recognition sequences for the Nsp4 serine protease at positions: E3143/G3144, E3584/G3585 and E3807/G3808 [50]. Cleavage at these sites would release processing products Nsp 9 to 12 (Fig. 5) [48].

The Nsp9 cleavage product contains the putative RdRp domain and, as with all arteriviruses, it possesses the RdRp core SDD motif at positions 3008–3010 [24]. In addition, four other RNA polymerase motifs conserved in positive stranded RNA viruses were present in PA8 at positions 2870–2883, 2890–2908, 2940–2965 and 3003–3014 [36]. The Nsp9 of VR2332 and 16244B differed from PA8 by 4 and 9 amino acids, respectively, 1 and 5 of which were nonconservative (Fig. 7). None of these mutations were found within the RdRp motifs described above.

Nsp10 contains a cysteine-histidine rich putative metal-binding domain at the amino-terminal [15] as well as an NTP-binding helicase motif [20]. The predicted metal-binding domain is located between amino acids 3150–3196 [15] and contains 10 Cys and His residues that are conserved throughout arteriviruses [32]. The NTP-binding motif, defined by the consensus sequence GxxxGKT, where "x" equals any amino acid [20] was found at the beginning of the conserved helicase domain, between amino acids 3292 and 3513 [14]. The Nsp10 of PA8 differs from VR2332 and 16244B by 1 and 3 amino acids, respectively, all of which were nonconservative changes (Fig. 7).

Nsp11 contains the highly conserved "coronavirus-like" domain between amino acids 3712 and 3807. This domain has been found in all other arteriviruses examined to date [7, 12, 28]. PA8 shares 98.7% and 99.0% nucleotide homology, with VR2332 and 16244B, respectively, within this region (Fig. 6B).

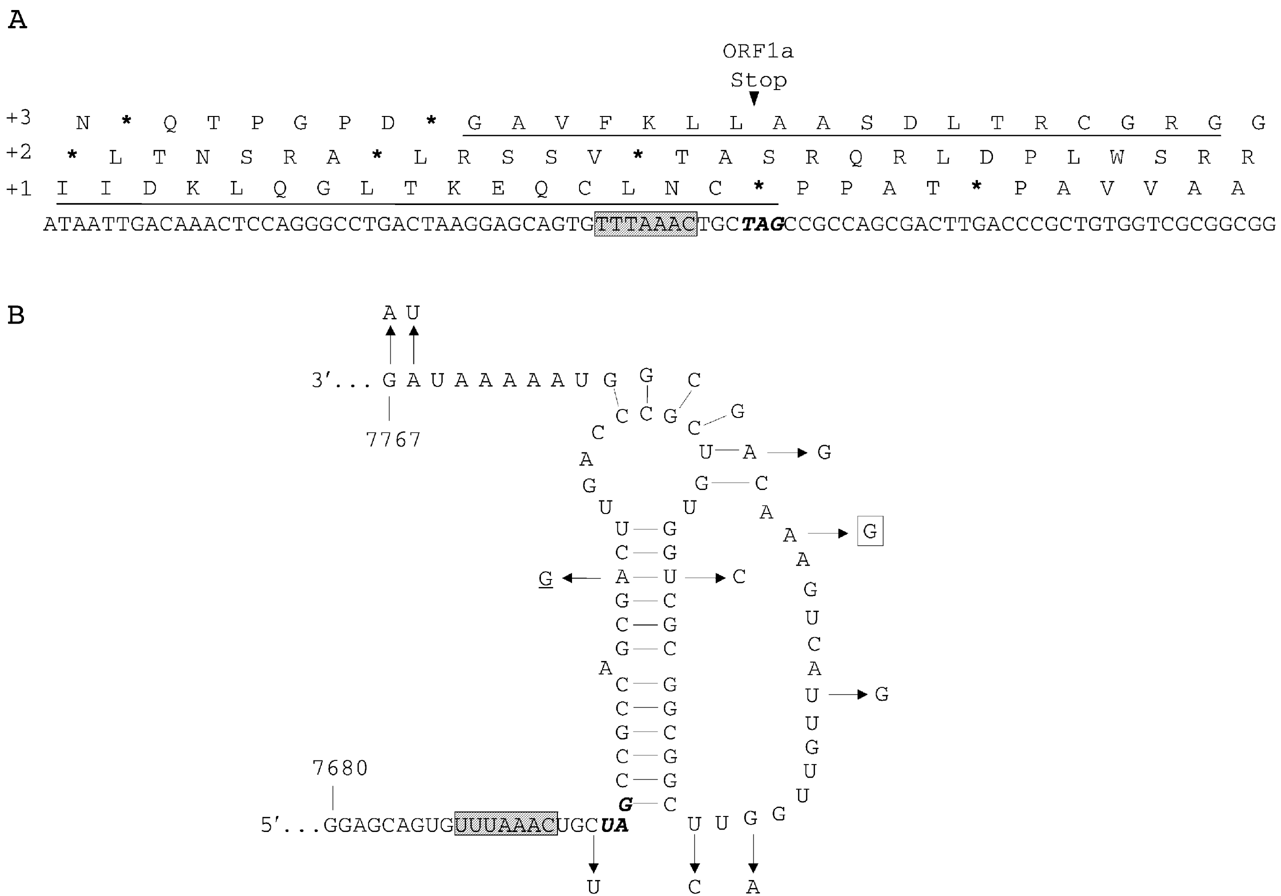

Fig. 8. Production of the ORF1ab fusion protein. **A** Nucleotide and amino acid sequence of the overlapping region between ORF1a and ORF1b. ORF1b (underlined in the +3 frame) is synthesized as a fusion with ORF1a (underlined in the +1 frame) through a (−1) ribosomal frameshift due to the presence of the heptanucleotide slippery sequence (shaded boxed). **B** Predicted pseudoknot structure of the ribosomal frameshift region. The heptanucleotide UUUAAAC slippery sequence is in the shaded-box, and the UAG stop codon of ORF1a is in bold. Numbers indicate nucleotide position. Arrows indicate the sequence differences of PA8 and VR2332 in comparison with that of LV and 16244B (boxed). Nucleotide differences common to both LV and 16244B are underlined

Interestingly, Nsp11 is the only nonstructural protein where 16244B is more similar to PA8, differing by only two conservative amino acids as opposed to five in the case of VR2332, where 4 changes were conservative (Fig. 7).

The carboxy-terminal cleavage product, Nsp12, is highly variable among North American and European species of PRRSV, with only 42% amino acid identity between LV and 16244B [1]. This variability does not extend to the North American PRRSV isolates, however, as PA8 and VR2332 are 100% identical at the amino acid level and 16244B differs by only 2 amino acids, one of which was a conservative change (Fig. 7).

Comparison of structural genes

ORF2a to 7 cover approximately 3 kb of the 3′end of the genome and encode all of the structural proteins (Table 2). Each protein, with the exception of ORF2b, is translated from a separate subgenomic mRNA [6]. The consensus intergenic sequence for these mRNAs in North American PRRSV is 5′-[U/G/A][U/C/A/G][A/C][A/G][C/U]C-3′ [32], whereas in LV it is 5′-[U/C/A]AACC-3′ [6]. The intergenic sequences for all 3 North American strains were perfectly conserved. In the interest of comparing pathogenic and attenuated strains of American type PRRSV, the nucleotide and amino acid sequences for the structural genes of VR2332, 16244B, three Danish isolates of American type PRRSV (DK5163-23, DK5163-17 and DK3016-12) [25] and the American type modified live vaccine (MLV) RespPRRS [16] were compared with PA8. RespPRRS, registered in Europe as Ingelvac PRRSV MLV, is a live vaccine derived from the virulent American field isolate VR2332 [5]. The nucleotide and amino acid differences observed among the American isolates are summarized in Fig. 9A and 9B, respectively. Amino acid differences occurred in all of the structural proteins with the exception of N which shared 100% amino acid identity among all 7 isolates examined. This suggests that N may not be the focus of antibody or other host defenses during the course of PRRSV infection. Relative to all the isolates examined, PA8 differed the most in the GP5 coding region. Only one amino acid difference resulted in the deletion of a potential N-glycosylation site according to the consensus NxT/S, where x cannot be P. The Danish isolate DK5163-17 possessed an N to T mutation at amino acid residue 33 of GP5 (position 601, Fig. 9B). In two of the Danish isolates; DK5163-17 and DK5163-23, a mutation from D to N at position 34 in GP5 (position 602, Fig. 9B) resulted in the addition of a putative N-glycosylation site. GP5 contains a hypervariable region within the ectodomain which is thought to be responsible for generating diversity in field isolates [2, 38] and it is possible that this ectodomain may define a region under selection by antibody or other host defenses [13]. Interestingly, in a study addressing the possible origin of PRRSV diversity in ORF5, the emergence of distinct PRRSV populations was characterized by a single amino acid change from D to N at position 34 in the ectodomain of GP5 [38].

In this report the first full-length sequence of a Canadian PRRSV isolate, PA8, was determined and compared with the genomic sequences of the American

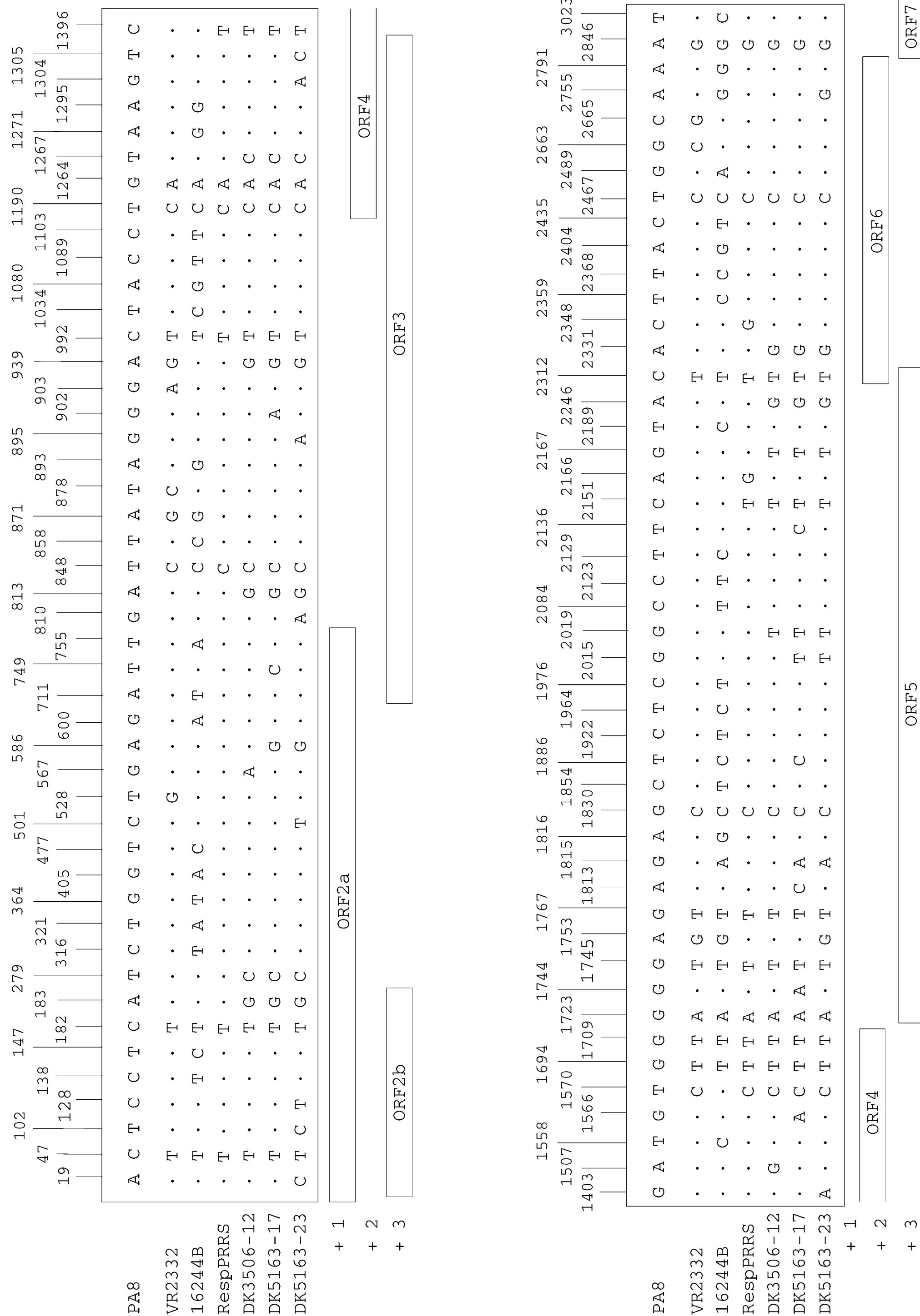

Fig. 9 *(continued)*

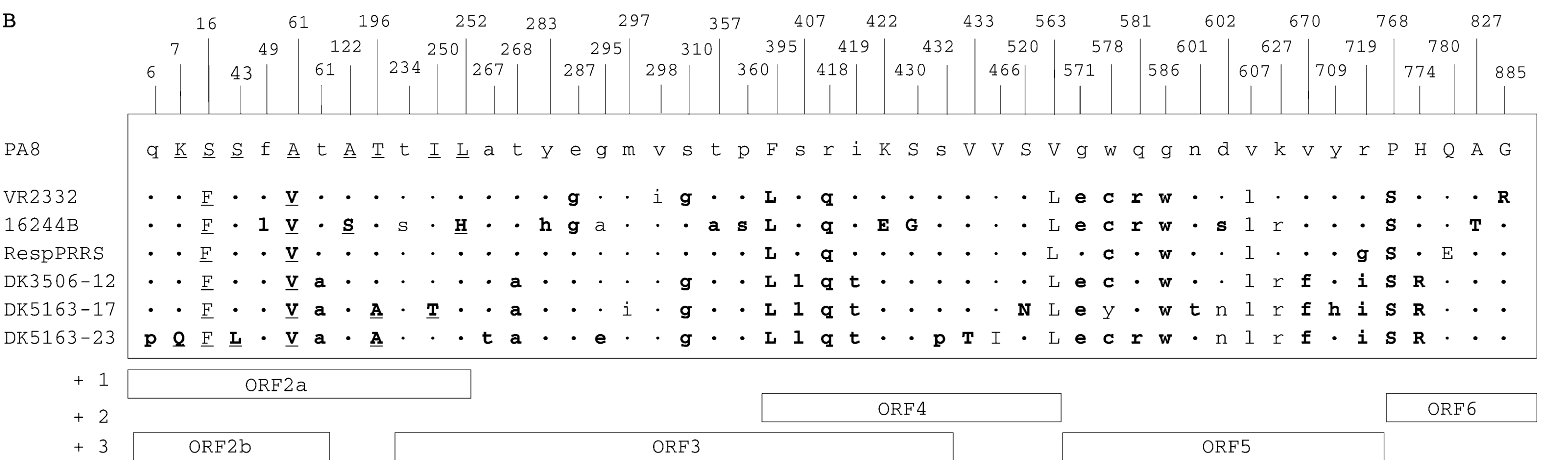

Fig. 9. Nucleotide (**A**) and amino acid (**B**) changes relative to PA8 within the structural coding region of VR2332, 16244B, Danish American type isolates DK3506-12, DK-5163-17, DK5163-23 [25] and the RespPRRS [16] vaccine. Numbers at the top indicate location within the PA8 genome where nucleotide (**A**) and amino acid (**B**) differences occur. Numbering starts at 1 and represents the first nucleotide (**A**) or the first amino acid (**B**) of ORF2. ORF2 to 7 are shown below the table in their respective reading frames. Amino acid changes in the +1 frame are underlined, those in the +2 frame are unchanged and those in the +3 frame are lower case. Nonconservative amino acid changes are shown in bold. Amino acid changes were considered to be conservative if they fell into one of these six groups: [L, I, V, M]; [A, G]; [S, T]; [D, E, N, Q]; [K, R] or [F, Y, W]

isolates VR2332 and 16244B. Completion of the PA8 genomic sequence permits, for the first time, a detailed comparison of PRRSV genomes from independent clones isolated within North America. All three strains displayed high nucleotide and amino acid homology. Of the 119 nucleotide differences within the coding regions between VR2332 and PA8, only 57.1% resulted in amino acid changes and 74% of these changes were nonconservative. Similarly, 16244B differed from PA8 by 274 nucleotides, 38% of which resulted in amino acid changes. Approximately 74% of the amino acid changes in 16244B were nonconservative in nature. Although an excess of nonsynonymous substitutions occurred, the majority of nucleotide changes did not result in amino acid changes, suggesting an evolutionary pressure to conserve amino acid sequences.

Both PA8 and 16244B are more closely related to VR2332 than to each other. Since VR2332 was one of the first isolated and characterized American strains of PRRSV and, in the light of the similarity that PA8 and 16244B share with VR2332, it is possible that these viruses are field strain derivatives of VR2332.

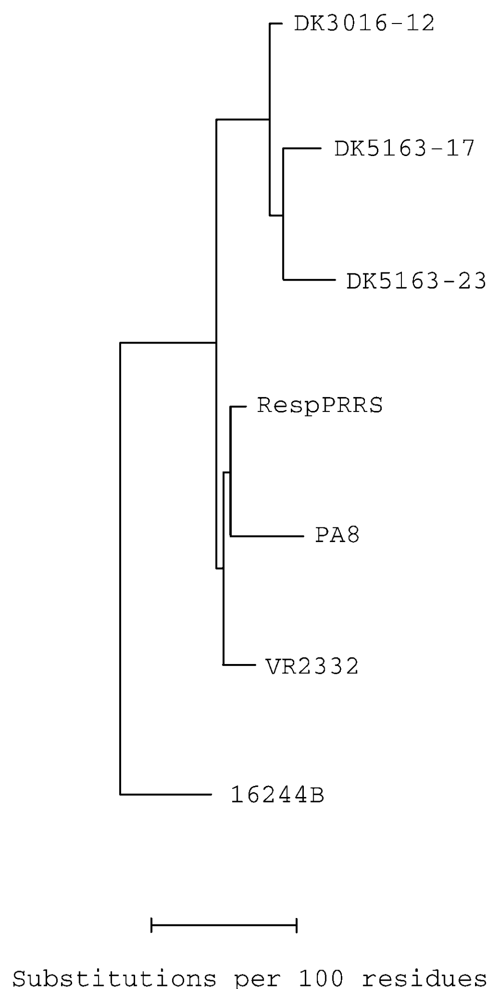

Fig. 10. Phylogenetic tree of the American type PRRSV viruses PA8, VR2332, 16244B, DK3506-12, DK-5163-17, DK5163-23 and RespPRRS. The tree was generated based on alignment of the entire structural coding region of ORF2 to 7 using the evolutionary alignment program of the GCG software package [9]

However, since some of the animals on the farm in Alberta had been vaccinated (MLV RespPRRS is derived from VR2332) this lends support to the idea that spread of the vaccine virus may account for the high degree of similarity between PA8 and VR2332. A phylogenetic tree was generated from the comparison of VR2332, 16244B, three American type Danish isolates and the RespPRRS vaccine with PA8 (Fig. 10). The Danish isolates used in the comparison of the structural proteins of PRRSV were chosen because they are suspected to have originated from the VR2332-based vaccine virus due to spread to uninfected animals [25]. Results from the evolutionary analysis indicate that PA8 is most closely related to RespPRRS and that both of these viruses are very similar to VR2332. These results suggest that PA8 may have originated due to spread of the vaccine virus to uninfected animals in the herd and subsequent reversion to virulence.

Acknowledgements

This work was supported by grants from the Ontario Ministry of Agriculture Food and Rural Affairs (OMAFRA) and Ontario Pork.

References

1. Allende R, Lewis TL, Lu Z, Rock DL, Kutish GF, Ali A, Doster AR, Osorio FA (1999) North American and European porcine reproductive and respiratory syndrome viruses differ in non-structural protein coding regions. J Gen Virol 80: 307–315
2. Andreyev VG, Wesley RD, Mengeling WL, Vowald AC, Lager KM (1997) Genetic variation and phylogenetic relationship of 22 porcine reproductive and respiratory syndrome virus (PRRSV) field strains based on sequence analysis of open reading frame 5. Arch Virol 142: 993–1001
3. Brierley I, Diggard P, Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal, requirement for an RNA pseudoknot. Cell 57: 537–547
4. Cavanagh D (1997) Nidovirales: a new order comprising *Coronaviridae* and *Arteriviridae*. Arch Virol 142: 629–633
5. Collins JE, Benfield DA, Christianson WT, Harris JE, Hennings L, Shaw JC, Goyal DP, Cullough SM, Morrison S, Joo RB, Gorcyca HS, Chladek DW (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
6. Conzelmann K, Visser N, van Woensel P, Thiel H (1993) Molecular characterization of porcine reproductive and respiratory syndrome virus, a member of the arterivirus group. Virology 193: 329–339
7. 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 coronaviruslike superfamily. J Virol 65: 2910–2920
8. Den Boon JA, Faaberg KS, Meulenberg JMM, Wassenaar ALM, Plagemann PGW, Gorbalenya AE, Snijder EJ (1995) Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a protein: identification of two papainlike cysteine proteases. J Virol 69: 4500–4505
9. Devereux J, Haeberli P, Smithies OI (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 18: 3241–3247

10. De Vries AAF, Horzinek MC, Rottier PJM, de Groot RJ (1997) The genome organization of the nidovirales: similarities and differences between arteri-, toro- and coronaviruses. Semin Virol 8: 33–47
11. Dougherty WG, Semler BL (1993) Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. Microbiol Rev 55: 781–822
12. Godney EK, Chen L, Kumar SN, Methven SL, Koonin EV, Brinton MA (1993) Complete genomic sequence and phylogenic analysis of the lactate dehydrogenase-elevating virus (LDV). Virology 194: 585–596
13. Gonin P, Pirzadeh B, Gagnon CA, Dea S (1999) Seroneutralization of porcine reproductive and respiratory syndrome virus correlates with antibody response to the GP5 major envelope glycoprotein. J Vet Diagn Invest 11: 20–26
14. Gorbalenya AE, Blinov VM, Donchenko AP, Koonin EV (1989) An NTP-binding motif is the most conserved sequence in a highly diverged monophyletic group of proteins involved in positive strand RNA viral replication. J Mol Evol 28: 256–268
15. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1989) Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res 17: 4847–4861
16. Gorcyca D, Schlesinger K, Chladek D, Behan D (1995) RespPRRS: a new tool for the prevention and control of PRRS in pigs. Am Assoc Swine Pract 1995: 1–22
17. Halbur PG, Paul PS, Frey ML, Landgraf J, Eernisse K, Meng X-J, Lum MA, Andrews JJ, Rathjc JA (1995) Comparison of the pathogenicity of two US porcine reproductive and respiratory syndrome virus isolates with that of the Lelystad virus. Vet Pathol 32: 648–660
18. Halbur PG, Paul PS, Meng X-J, Lum MA, Andrews JJ, Rathje JA (1996) Comparative pathogenicity of nine US porcine reproductive and respiratory syndrome virus (PRRSV) isolates in a five-week-old cesarean-derived, colostrum-deprived pig model. J Vet Diagn Invest 8: 11–20
19. Jacks T, Madhani HD, Masiarz FR, Varmus HE (1988) Signals for the ribosomal frameshifting in the rous sarcoma virus *gag-pol* region. Cell 55: 447–458
20. Kadare G, Haenni AL (1997) Virus-encoded RNA helicases. J Virol 71: 2583–2590
21. Kapur V, Elam MR, Pawlovich TM, Murtaugh MP (1996) Genetic variation in porcine reproductive and respiratory syndrome virus isolates in the Midwestern United States. J Gen Virol 77: 1271–1276
22. Keffaber KK (1989) Reproductive failure of unknown etiology. Am Assoc Swine Pract Newslett 1: 1–9
23. Kim HS, Kwang J, Yoon IJ, Joo HS, Frey ML (1993) Enhanced replication of PRRS virus in a homogeneous subpopulation of MA-104 cell line. Arch Virol 133: 477–483
24. Koonin EV (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J Gen Virol 72: 2197–2206
25. Madsen KG, Hansen CM, Madsen ES, Strandbygaard B, Botner A, Sorensen KJ (1998) Sequence analysis of porcine reproductive and respiratory syndrome virus of the American type collected from Danish swine herds. Arch Virol 143: 1683–1700
26. Meng X-J, Paul PS, Halbur PG, Morozov I (1995) Sequence comparison of open reading frame 2 to 5 of low and high virulence United States isolates of porcine reproductive and respiratory syndrome virus. J Gen Virol 76: 3181–3188
27. Meng X-J, Paul PS, Halbur PG, Lum MA (1995) Phylogenetic analysis of the putative M (ORF 6) and N (ORF 7) genes of porcine reproductive and respiratory syndrome virus (PRRSV): implication for the existence of two genotypes of PRRSV in the U.S.A. and Europe. Arch Virol 140: 745–755
28. Meulenberg JJM, Hulst MM, de Meijer EJ, Moonen PJM, den Besten A, De Kluyver

EP, Wensvoort G, Moormann RJM (1993) Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. Virology 192: 62–72
29. Meulenberg JJM, Bos-De Ruijter JNA, Van De Graaf R, Wensvoort G, Moormann RJM (1998) Infectious transcripts from cloned genome-length cDNA of porcine reproductive and respiratory syndrome virus. J Virol 72: 380–387
30. Murtaugh MP, Faaberg KS, Laber J, Elam M, Kapur V (1998) Genetic variation in the PRRS virus. Adv Exp Med Biol 440: 787–794
31. Nelson EA, Christopher-Hennings J, Drew T, Wensvoort G, Collins JE, Benfield DA (1993) Differentiation of US and European isolates of porcine reproductive and respiratory syndrome virus by monoclonal antibodies. J Clin Microbiol 31: 3184–3189
32. Nelson CJ, Murtaugh MP, Faaberg KS (1999) Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. J Virol 73: 270–280
33. Oleksiewicz MB, Botner A, Nielson J, Storgaard T (1999) Determination of the 5′-leader sequences from radically disparate strains of porcine reproductive and respiratory syndrome virus reveals the presence of highly conserved sequence motifs. Arch Virol 144: 981–987
34. Paton DJ, Brown IH, Edwards S, Wensvoort G (1991) Blue ear disease of pigs. Vet Rec 128: 617
35. Pedersen KW, van der Meer Y, Roos N, Snijder EJ (1999) Open reading frame 1a-encoded subunits of the arterivirus replicase induce endoplasmic reticulum-derived double membrane vesicle which carry the viral replication complex. J Virol 73: 2016–2067
36. Poch O, Sauvaget I, Delarue M, Tordo N (1989) Identification of four conserved motifs among the RNA dependent polymerase encoding elements. EMBO J 8: 3867–3874
37. Rossow KD, Shivers JL, Yeske PE, Polson DD, Rowland RR, Lawson SR, Murtaugh MP, Nelson EA, Collins JE (1999) Porcine reproductive and respiratory syndrome virus infection in neonatal pigs characterised by marked neurovirulence. Vet Rec 17: 444–448
38. Rowland RR, Steffen M, Ackerman T, Benfield DA (1999) The evolution of porcine reproductive and respiratory syndrome virus: quasispecies and emergence of a virus population during infection of pigs with VR2332. Virology 259: 262–266
39. Rowland RR, Kervin R, Kuckleberg C, Sperlich A, Benfield DA (1999) The localization of porcine reproductive and respiratory syndrome virus nucleocapsid protein to the nucleolus of infected cells and identification of a potential nucleolar localization signal sequence. Virus Res 64: 1–12
40. Salles FJ, Strickland S (1995) Rapid and sensitive analysis of mRNA polyadenylation states by PCR. PCR Methods Appl 4: 317–321
41. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
42. Snijder EJ, Wassenaar ALM, Spaan WJM (1994) Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. J Virol 68: 5755–5764
43. Snijder EJ, Wassenaar ALM, Spaan WJM, Gorbalenya AE (1995) The arterivirus nsp2 protease. J Biol Chem 270: 16671–16676
44. Snijder EJ, Wassenaar ALM, Van Dinten LC, Spaan WJM, Gorbalenya AE (1996) The arterivirus nsp4 protease is the prototype of a novel group of chymotrypsin-like enzymes, the 3C-like serine proteases. J Biol Chem 271: 4864–4871
45. Snijder EJ, Meulenberg JM (1998) The molecular biology of arteriviruses. J Gen Virol 79: 961–979
46. Snijder EJ, van Tol H, Pedersen KW, Raamsman MJ, de Vries AA (1999) Identification of a novel structural protein of arteriviruses. J Virol 73: 6335–6345

47. van der Meer Y, van Tol H, Locker JK, Snijder (1998) ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. J Virol 72: 6689–6698
48. van Dinten LC, Wassenaar ALM, Gorbalenya AE, Spaan WJM, Snijder EJ (1996) Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. J Virol 70: 6625–6633
49. van Dinten LC, den Boon JA, Wassenaar ALM, Spaan WJM, Snijder EJ (1997) An infectious arterivirus cDNA clone: identification of a replicase point mutant which abolishes discontinuous mRNA transcription. Proc Natl Acad Sci USA 94: 991–996.
50. van Dinten LC, Rensen S, Gorbalenya AE, Snijder EJ (1999) Proteolyic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. J Virol 73: 2027–2037
51. Ward CD, Stokes MA, Flanegan JB (1988) Direct measurement of the poliovirus RNA polymerase error frequency in vitro. J Virol 62: 558–62
52. Wassenaar ALM, Spaan WJM, Gorbalenya AE, Snijder E (1997) Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that Nsp2 acts as a cofactor for the Nsp4 serine protease. J Virol 71: 9313–9322
53. Wensvoort G, Tepstra C, Pol JMA, ter Laak EA, Bloemraad M, de Kluyver EP, Kragten C, van Buiten L, den Besten A, Wagenaar F, Broekhuijsen JM, Moonen PLJM, Zestra T, de Boer EA, Tibben HJ, de Jong MF, van Veld P, Groenland GJR, van Gennep JA, Voets MT, Verheijden JHM, Braamskamp J (1991) Mystery swine disease in the Netherlands: the isolation of Lelystad virus. Vet Q 13: 121–130
54. Wensvoort G, de Kluyver EP, Luijtze EA, den Besten A, Harris L, Collins JE, Christianson WT, Chladek D (1992) Antigenic comparison of Lelystad virus and swine infertility and respiratory syndrome (SIRS) virus. J Vet Diag Invest 4: 134–163
55. Wootton SK, Nelson EA, Yoo D (1998) Antigenic structure of the nucleocapsid protein of porcine reproductive and respiratory syndrome virus. Clin Diagn Lab Immunol 5: 773–779
56. Zuker M, Jaeger JA, Turner DH (1991) A comparison of optimal and suboptimal RNA secondary structures predicted by free energy minimization with structures determined by phylogenetic comparison. Nucleic Acids Res 19: 2707–2714

Authors' address: Dr. D. Yoo, Department of Pathobiology, University of Guelph, Guelph, Ontario N1G 2W1, Canada.

Received January 10, 2000