Virology **267,** 135–140 (2000)
doi:10.1006/viro.1999.0103, available online at http://www.idealibrary.com on IDEAL®

RAPID COMMUNICATION

Emergence of Porcine Reproductive and Respiratory Syndrome Virus Deletion Mutants: Correlation with the Porcine Antibody Response to a Hypervariable Site in the ORF 3 Structural Glycoprotein

M. B. Oleksiewicz, A. Bøtner, P. Toft, T. Grubbe, J. Nielsen, S. Kamstrup, and T. Storgaard[1]

Danish Veterinary Institute for Virus Research, Lindholm, DK-4771 Kalvehave, Denmark

Received September 7, 1999; returned to author for revision November 11, 1999; accepted November 22, 1999

By using porcine immune sera to select a library of phage-displayed random peptides, we identified an antigenic sequence (RKASLSTS) in the C-terminus of the ORF 3 structural glycoprotein of European-type porcine reproductive and respiratory syndrome virus (PRRSV). Through the use of overlapping reading frames, the same PRRSV genetic locus codes for the ORF 3 "RKASLSTS" sequence, and a previously described ORF 4 epitope (Meulenberg, J. J. M., Van Nieuwstadt, A. P., Van Essen-Zandbergen, A., and Langeveld, J. P. M., 1997, *J. Virol.* 71, 6061–6067). Sequence analysis identified naturally occurring deletion mutants at this ORF ¾ site. Phylogenetic analysis showed the presence of a highly accurate ORF 3 molecular clock, according to which deletion mutants and nondeleted viruses evolved at differing speeds. Furthermore, deletion mutants and nondeleted viruses evolved as separate lineages. These distinctions suggested that deletion mutants were a hitherto unrecognized subtype of European-type PRRSV. Currently, deletion mutants appear to be outcompeting nondeleted viruses in the field, highlighting the importance of the porcine antibody response against the minor structural glycoproteins of European-type PRRSV for viral evolution.

Since its emergence in the late eighties, porcine reproductive and respiratory syndrome virus (PRRSV) has become a wide-spread pathogen, causing significant economical impact. Monitoring the evolution of recently emerged viruses is of basic interest, as such agents may still be adapting to their new host, and thus may be genetically unstable. Evolution of antigenic areas is of particular interest, due to the practical relevance for immune diagnostics and prophylaxis. To examine the specificity of porcine antibodies following PRRSV infection, we used a phage-display library of random 7-mer peptides (Ph.D-7 library, New England Biolabs, Hitchin, Hertfordshire, UK). Sera from pigs 40–60 days post-experimental infection with 111/92, a European-type PRRSV (*1*), were preadsorbed with UV-inactivated wild-type M13 phage, and used to select the Ph.D-7 library, essentially according to the manufacturer's instructions. After five cycles of selection and amplification, peptide sequences from 32 plaque-purified phages were matched to 111/92's structural genes (EMBL Accession No. AJ223078). A better than 4 out of 7 (4/7) residue match between selected peptide and cognate antigen may not appear even when using a monoclonal antibody for library selection (*2–5*). This posed a problem, because with a target sequence as long as the structural genes of PRRSV (1200 amino acids), any single heptamer peptide has an unacceptably large chance of exhibiting a nonsignificant 4/7 match. To overcome this, we required a 4/7 match to the heptamer peptide of two different phage clones to define a putative antigenic site in PRRSV. With this strategy, we identified a putative antigenic site (RKASLSTS) in the C-terminus of the ORF 3 protein; the phage-identified site exactly coincided with a predicted major antigenic peak (Fig. 1). As a negative control, the same 32 phage clones were matched to the structural genes of VR2332 (GenBank Accession No. U87392), an American-type PRRSV antigenically very distinct from European-type PRRSV (*6*); no putative antigenic sites were found. The phage-display data were further corroborated by sequence analysis and Western blotting, as described below.

For sequence analysis, we RT-PCR amplified two overlapping fragments, covering the complete ORF 3 sequence (C-terminus, PCR primers 5′GTGTCGCGCGTCTTCGTGGA3′ and 5′GCACGCAGAAAGCATCAGCA3′, N-terminus, PCR primers 5′ACCGCGTTGAGCTCATCTTC3′ and 5′CCATTCCAAATGGAACCAITTG3′), from 26 Danish, European-type PRRSV isolates covering the period 1992–1998. Sequencing of both strands was done on uncloned, gel-purified PCR product, using the BigDye cycle sequencing kit (Per-

[1] To whom correspondence and reprint requests should be addressed. Fax: +45 5586 0300. E-mail: TS@VETVIRUS.DK.

0042-6822/00 $35.00

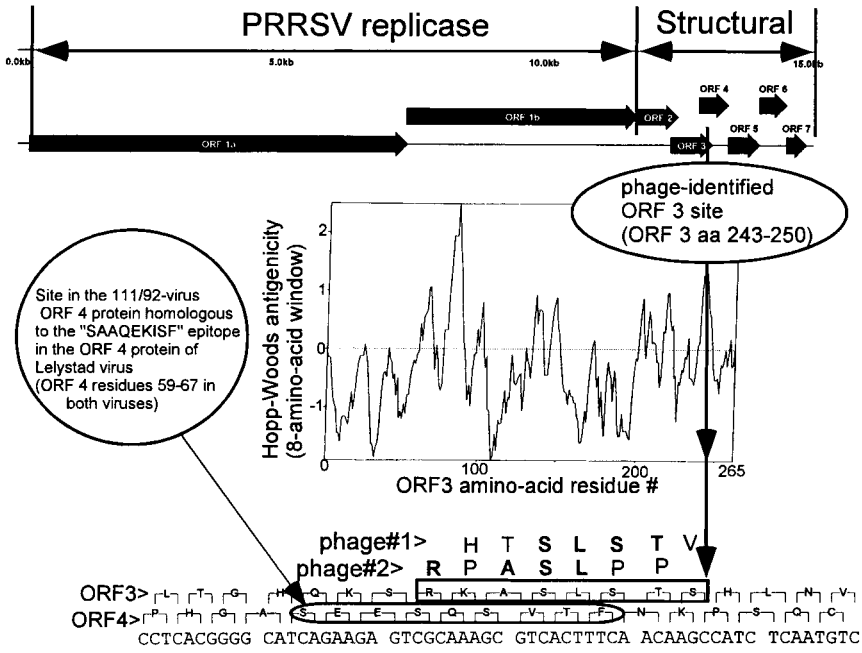


FIG. 1. Identification of an antigenic site in the C-terminus of the ORF 3 protein by phage display. Selecting a naive random heptamer–peptide phage-display library (each random peptide sequence present at $<10^{-7}$ frequency) with immune sera from PRRSV 111/92-infected pigs led to the enrichment of two phage clones (post-enrichment frequencies of 5/32 and 6/32), whose peptide insert partially matched a sequence at the C-terminus of the ORF 3 protein. The amino acid sequences of the two phage-displayed peptides are shown aligned to the PRRSV 111/92 ORF 3 amino acid sequence. Residues matching PRRSV ORF 3 appear bold in the phage–peptide sequences. The PRRSV genetic locus which codes for the ORF 3 site identified by polyclonal porcine sera in this study also codes for an ORF 4 epitope (marked in the ORF 4 amino acid sequence) identified through mAb mapping by Meulenberg *et al.* (*14*). Also shown is information about the PRRSV genomic organization (*16*) and predicted Hopp–Woods antigenicity profile of the ORF 3 protein (*17*).

kin–Elmer, Allerød, Denmark). The sequences, with isolate dates, have been submitted to GenBank (Accession No. AF171671-96). Alignment of the 26 sequences showed that the phage-identified RKASLSTS sequence localized to the most variable segment of the ORF 3 protein, as expected for a site targeted by the immune response (*7, 8*) (Fig. 2). Intriguingly, many isolates exhibited frank deletions in the RKASLSTS sequence. The deletions were 3–24 nt large, localized to the RKASLSTS site, and, with a single exception (see below), did not disturb the ORF 3 and 4 reading frames, which overlap at this location (Figs. 1 and 2). This is the first report of PRRSV deletion mutants; remarkably, the deletions affected antigenic areas in two overlapping ORFs (Fig. 1). Interestingly, deletions have as yet not been described in American-type ORF 3 sequences, possibly reflecting differences in the degree of virion association between EU-type (virion associated (*9*)) and U.S.-type (not virion associated (*10*)) ORF 3 proteins.

Next, to confirm the phage-display data by Western blotting, we RT-PCR amplified a fragment corresponding to the C-terminal 199 aa of the 265-aa-large ORF 3 protein (PCR primers 5′GCTCGAGCCCGGTCGTAAC3′ and 5′GTTATCGTGARTACTGGG3′, ORF 3 stop codon underlined). This fragment contained the predicted major antigenic sites of the protein, while avoiding the hydrophobic N-terminus (Figs. 1 and 2), and has previously been shown to be antigenic *in vivo* (*7*). Cloned RT-PCR products (pT7Blue-2 vector, Novagen, Madison, WI) were confirmed by sequencing and *in vitro* transcribed/translated in the presence of biotinylated lysine (TNT rabbit reticulocyte lysates, Promega, Madison, WI). The ORF 3 proteins were N-terminally fused to a vector-derived S-peptide, allowing precise quantitation with the S-tag assay (Novagen). Equalized amounts of *in vitro* translated ORF3 antigen were electrophoresed on precast 15% SDS–PAGE gels (BioRad, Copenhagen, Denmark), and electrotransferred to PVDF membranes (Amersham, Hørsholm, Denmark) for Western blot analysis. Serum from a 111/92-infected pig reacted with the ORF 3 protein of 111/92, but had significantly reduced reactivity toward the ORF 3 protein of 21191, a naturally occurring deletion mutant (Fig. 3). The 21191 isolate exhibited a near-total deletion of the RKASLSTS phage-identified sequence, as well as several other amino acid changes (Fig. 2). To determine the effect of the deletion of the RKASLSTS sequence alone, targeted deletions were made in the 111/92 protein using mutagenic reverse PCR primers 5′TTATCGTGATGTACTGGGGAATGCCGACGGCTTCACGACATTGAGATGACTCTTCTGATG3′ and 5′TTATCGTGATGTACTGGGGAATGCCGACGGCTTCACGACATTGAGGCCGGGCACAGGTG3′ in conjunction with the forward PCR primer described above (deletion of the

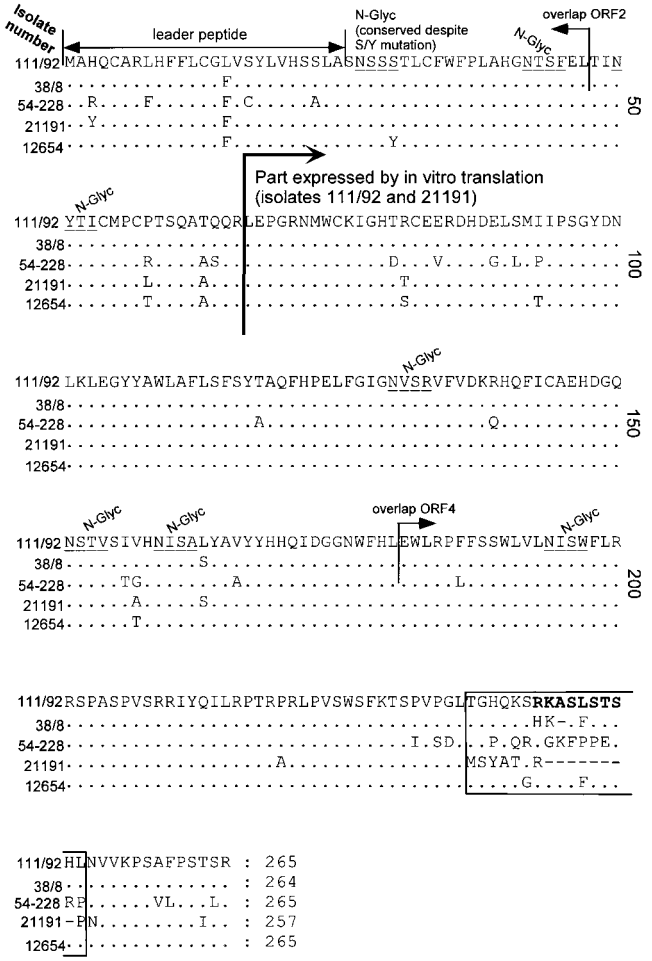


FIG. 2. Sequence variability in the ORF 3 protein of Danish, European-type PRRSV isolates. Of the 26 complete ORF 3 sequences generated in this study (GenBank Accession No. AF171671-96), 5 were selected as representative: 111/92 (GenBank Accession No. AF171671) is the nondeleted isolate used for experimental infections to generate porcine immune sera for phage-display screening (see Fig. 1) and Western blot analysis (see Fig. 3). Isolates 12654 (GenBank Accession No. AF171672) and 54-228 (GenBank Accession No. AF171691) are nondeleted isolates exhibiting low and high variability at the phage-identified RKASLSTS site. Isolates 38/8 (GenBank Accession No. AF171689) and 21191 (GenBank Accession No. AF171679) are deletion-mutant isolates exhibiting small and large deletions at the phage-identified RKASLSTS site. In the shown alignment, sequence identity to 111/92 is indicated by dots. Deletions are indicated by dashes. The leader peptide was predicted using the Signal IP software (*18*). Predicted N-linked glycosylation sites are shown underlined. Glycosylation sites and cysteines were generally well conserved in all ORF 3 sequences. The most variable ORF 3 segment (boxed residues) was identified from an alignment of the 26 full-length ORF 3 sequences generated in this study using the "mutational hotspot" function of the Omiga software (Oxford Molecular Ltd., Oxford, UK), and agreed with the findings of Drew *et al.* (*8*). The part of the ORF 3 protein expressed by *in vitro* translation is indicated. The amino- and carboxy-terminal parts of the ORF 3 protein are encoded by genome areas containing overlapping ORFs, as shown.

RKASLSTS sequence and the 16-aa-large hypervariable segment boxed in Fig. 2, respectively). Deletion of the RKASLSTS phage-identified sequence by PCR reduced the reactivity of the 111/92 ORF3 protein (Fig. 3), with a somewhat larger reduction in reactivity evident if the 16-aa-large hypervariable site containing the RKASLSTS sequence (Fig. 2, boxed residues) was deleted (Fig. 3). Thus, the Western blot results confirmed that the phage-identified RKASLSTS sequence (Fig. 1) was a naturally antigenic site. Further, natural evolution of this and presumably other antigenic sites, exemplified by the 21191 isolate, had significantly changed the antigenicity of the ORF3 protein.

Finally, to examine the phylogenetic relationship between deletion mutants and "nondeleted" viruses, we removed the genome segment subject to deletions (ORF 3 nucleotides 695–742), as current phylogenetic models deal poorly with insertion/deletion information, and constructed a maximum-likelihood phylogenetic tree (Fig. 4A). PRRSV was first isolated in Denmark in March 1992 (*1*). The tree showed that the PRRSV type(s) introduced into Denmark were generally distinct from the first European PRRSV isolate, the Lelystad virus from the Netherlands (*11*) (Fig. 4A). Deletion mutants mostly occupied branches which were free of nondeleted viruses, and vice versa (Fig. 4A). This indicated that generation of deletion mutants from non-

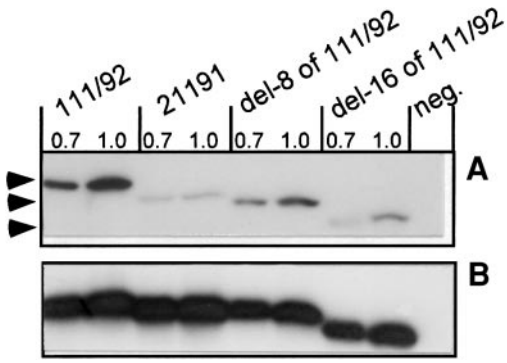
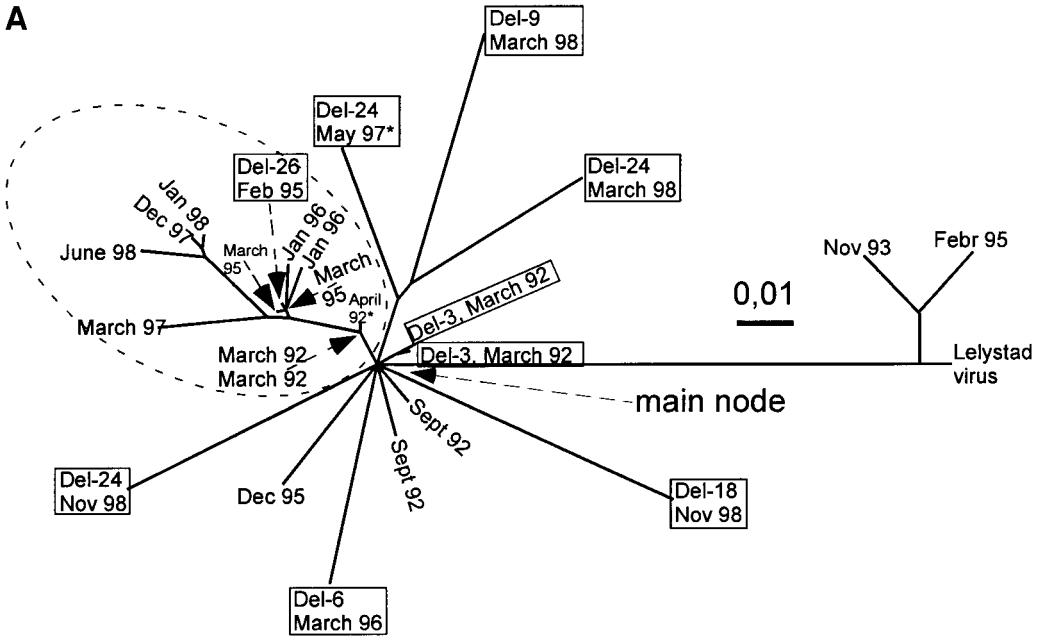
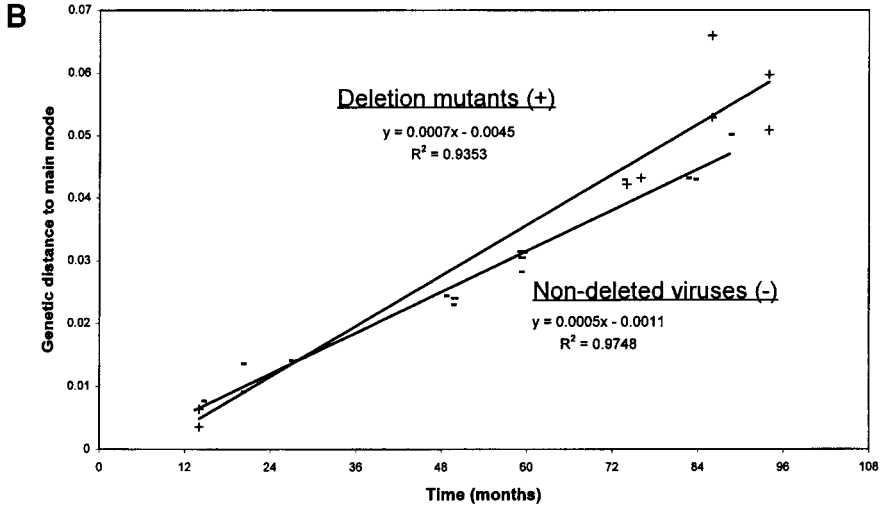


FIG. 3. Examination of the phage-identified RKASLSTS sequence by Western blot. *In vitro* translated, lysine-biotinylated ORF 3 protein was made from the PRRSV field isolates 111/92 and 21191 (see Fig. 2). Deletions of the phage-identified RKASLSTS site and the 16-aa-large hypervariable area (boxed in Fig. 2) containing the RKASLSTS site were made in the 111/92 protein by PCR; these constructs are indicated as "del-8 of 111/92" and "del-16 of 111/92," respectively. Specific bands are indicated by arrowheads, the deletions present in the 21191, del-8 of 111/92, and del-16 of 111/92 ORF3 proteins being reflected in increased mobility. Numbers above lanes indicate the number of ORF3 molecules loaded, as determined by the S-tag assay (relative units). The "negative" lane was loaded with unprogrammed reticulocyte lysate. (A) The Western blot was probed with 1:300 diluted serum from a 111/92-infected pig at 42 days postinfection, HRP-conjugated secondary antibody, and developed with ECL^{Plus} (Amersham). (B) To confirm the presence of equal amounts of protein of the expected size in all lanes, the duplicate blot was probed with streptavidine-conjugated alkaline phosphatase and developed with CDP^{Star} (Tropix, Bedford, MA). The total lysine content (potential biotinylation sites) of the *in vitro* translated proteins was, from left to right as they appear on the Western blot, 10, 7, 9, and 8.





C

Year	Total number of sequenced isolates	Frequency of deletion mutants	Deletion sizes (nucleotides)
1992	7	2/7=0.29	3, 3
1993	2	0	
1995	6	1/6=0.17	26
1996	3	1/3=0.33	6
1997	3	1/3=0.33	24
1998	8	5/8=0.63	9, 18, 18, 24, 24

FIG. 4. del^+ and del^- viruses evolve as separate lineages, at differing speeds, and with differing success. The ORF 3 sequences generated in this study of Danish, European-type PRRSV isolates are presented as a phylogenetic tree (A), a plot of genetic distance versus time (B), and a frequency and size distribution of deletions over time (C). (A) The nucleotide segment affected by deletions (ORF 3 nucleotides 695–742), and ORF 3 start and stop codons were removed from sequences prior to phylogenetic modeling. The tree was constructed using the HKY85+Γ substitution model (*19*), and all parameters were estimated by maximum likelihood (*20*). The isolate date is given for each point in the tree. Deletion mutants are boxed, "del-3" meaning deletion of 3 nt, etc. 111/92 (April 92) and 21191(Del-24, May 97), the isolates used for the *in vitro* translation studies (see Fig. 3), are indicated by *. A branch containing most nondeleted viruses, but largely devoid of deletion mutants, is indicated by a dashed oval. (B) Genetic distances from the main node of the phylogenetic tree were plotted against isolate time. Lelystad virus and the two Danish Lelystad-like isolates were excluded from the analysis. Sequences with and without deletions are shown by "+" and "−," respectively. Covariance analysis revealed that the slope of + and −

deleted viruses was rare and that present-day deletion mutants likely had evolved from deletion mutant parents existing at the start of the Danish epidemic. As an exception to this, a single deletion mutant was situated in the branch occupied by most nondeleted viruses (Fig. 4A, nondeleted virus branch indicated by stipled oval, deletion mutant boxed). However, other highly unusual characteristics of this deletion mutant (a 26-nt frame-shifting deletion, which was predicted to cause the formation of an ORF 3 + 4 fusion protein, GenBank Accession No. AF171696) suggested that it might represent a minor pathway in PRRSV evolution.

When the genetic distances to the main node (main node marked in Fig. 4A) were plotted against isolate time, we found an excellent linear correlation, which corroborated the phylogenetic tree, and showed a highly accurate ORF 3 molecular clock (Fig. 4B). Intriguingly, the molecular clock indicated that deletion mutants evolved significantly quicker than did nondeleted viruses (Fig. 4B). The tree shown in Fig. 4A was then used as input in the codon-based model of Yang and Nielsen (*12*) to estimate underlying ratios of nonsynonymous to synonymous nucleotide substitutions (d_N/d_S ratios). The d_N/d_S ratios were 0.29 and 0.46 for deleted and nondeleted viruses, respectively ($0.025 < P < 0.05$ (*12*)). Thus, the faster evolution of the deletion mutant ORF 3 appeared not to be due to increased selective pressure for change at the protein level. Hypothetically, a faster evolution rate despite a lower d_N/d_S ratio, as observed for the deletion mutants, could for example be due to peculiarities in RNA replication, which could perhaps also explain the generation of deletions per se. Importantly, using a codon-based model avoided confounding of the d_N/d_S ratios by the pronounced transition bias which is observed during PRRSV evolution (the rate ratio (*12*) of transitions/transversions was 10 for the sequences reported in this study). Also, the d_N/d_S analysis yielded same results regardless of whether the parts of ORF 3 that overlap with ORFs 2 and 4 were excluded or not.

The ORF 3 protein of European-type PRRSV has been reported to be capable of inducing protective immunity *in vivo* (*13*), and monoclonal antibodies to the "overlapping" ORF 4 site (Fig. 1) are neutralizing *in vitro* (*14*). That, together with our data above, suggested that deletion mutants might have emerged due to immunological pressure, the nature (humoral or cellular) and target (ORF 3 or 4 proteins) of which cannot presently be decided. In any case, such adaptive PRRSV evolution would be expected to produce viral fitness gain. Our data in fact provided evidence for this: Since 1992, the proportion of deletion mutants has significantly increased among field isolates, with a (nonsignificant) tendency toward larger deletions as well (Fig. 4C). Importantly, this immune response-driven (adaptive) evolution affected the short, hypervariable and deletion-prone ORF3 segment (Fig. 2), whereas the rest of the ORF 3 protein appeared to evolve according to neutral principles, as evidenced by the presence of a highly accurate molecular clock (see above, note that the deletion-prone ORF3 segment was removed from all sequences prior to the phylogenetic modeling, which revealed the molecular clock and on which d_N/d_S calculations were based).

Finally, it should be mentioned that deletion mutants from 1997 and 1998 were isolated from cases of recent PRRSV introduction into previously PRRSV-free farms; that is, deletion mutants were capable of interfarm spread, as opposed to, for example, being evolutionary dead ends of long-term, intrafarm persistence. Deletion mutants exhibited no unusual geographical clustering. Also, based on the type of clinical material from which deletion mutants were isolated, and clinical information from affected herds, we have currently no indication of deletion mutants being associated with apparent changes in PRRSV disease pattern. In summary, this study provided evidence for neutral (molecular clock, Fig. 4B) as well as adaptive (deletions/mutations at an antigenic site) changes in a PRRSV minor structural glycoprotein of unknown but likely essential (*15*) function. It remains to be determined whether genetic changes will exacerbate at the deletion-prone ORF 3/4 site as the virus continues adaptation to its new porcine host, and what effects this might have on PRRSV-induced disease.

ACKNOWLEDGMENTS

We thank P. Normann for introduction to the ABI310 genetic analyzer, and J. Nowicki for help with the artwork. N. Kirkby, K. J. Sørensen, and R. Forsberg participated in helpful discussions. This work was partly supported by a grant from the Norma and Frode Jacobsen foundation, and Grant 9702742 from The Danish Research Councils to M.B.O.

lines was significantly different ($P < 0.05$). (C) Three partial C-terminal ORF 3 sequences not yet submitted to GenBank were included in the table, bringing the total sequence number to 29. The frequency of deletion mutants increased significantly over time (Spearman correlation coefficient 0.81, $P < 0.05$). The apparent increase in the size of deletions over time was only significant if the 26-nt frame-shifting deletion from 1995 was excluded from the analysis (Spearman correlation coefficient 0.71, $P < 0.05$). The 1995 frame-shifting deletion mutant (GenBank Accession No. AF171696) is thought to represent an evolutionary pathway distinct from that followed by the other deletion mutant isolates (see text for details). Statistical analysis was performed in the SAS system, using PROC REG, PROC CORR, and PROC GLM.

REFERENCES

1. Bøtner, A., Nielsen, J., and Bille-Hansen, V. (1994). Isolation of porcine reproductive and respiratory syndrome (PRRS) virus in a Danish swine herd and experimental infection of pregnant gilts with the virus. *Vet. Microbiol.* **40,** 351–360.
2. Jellis, C. L., Cradick, T. J., Rennert, P., Salinas, P., Boyd, J., Amirault, T., and Gray, G. S. (1993). Defining critical residues in the epitope for a HIV-neutralizing monoclonal antibody using phage display and peptide array technologies. *Gene* **137,** 63–68.
3. Bonnycastle, L. L. C., Mehroke, J. S., Rashed, M., Gong, X., and Scott, J. K. (1996). Probing the basis of antibody reactivity with a panel of constrained peptide libraries displayed by filamentous phage. *J. Mol. Biol.* **258,** 747–762.
4. Keller, P. M., Arnold, B. A., Shaw, A. R., Tolman, R. L., van Middlesworth, F., Bondy, S., Rusiecki, V. K., Koenig, S., Zolla-Pazner, S., Conard, P., Emini, E. A., and Conley, A. J. (1993). Identification of HIV vaccine candidate peptides by screening random phage epitope libraries. *Virology* **193,** 709–716.
5. Cortese, R., Felici, F., Galfré, G., Luzzago, A., Monaci, P., and Nicosia, A. (1994). Epitope discovery using peptide libraries displayed on phage. *TIBTECH* **12,** 262–267.
6. Sørensen, K. J., Strandbygaard, B., Bøtner, A., Madsen, E. S., Nielsen, J., and Have, P. (1998). Blocking ELISA's for the distinction between antibodies against European and American strains of porcine reproductive and respiratory syndrome (PRRS) virus. *Vet. Microbiol.* **60,** 169–177.
7. Katz, J. B., Schafer, A. L., Eernisse, K., and Landgraf, J. G. N. E. A. (1995). Antigenic differences between European and American isolates of porcine reproductive and respiratory syndrome virus (PRRSV) are encoded by the carboxyterminal portion of viral open reading frame 3. *Vet. Microbiol.* **44,** 65–76.
8. Drew, T., Lowings, J. P., and Yapp, F. (1997). Variation in open reading frames 3, 4 and 7 among porcine reproductive and respiratory syndrome isolates in the UK. *Vet. Microbiol.* **55,** 209–221.
9. Van Nieeuwstadt, A. P., Meulenberg, J. J., Van Essen-Zandbergen, A., Petersen-den, B. A., Bende, R. J., Moormann, R. J., and Wensvoort, G. (1996). Proteins encoded by open reading frames 3 and 4 of the genome of Lelystad virus (*Arteriviridae*) are structural proteins of the virion. *J. Virol.* **70,** 4767–4772.
10. Mardassi, H., Gonin, P., Gagnon, C. A., Massie, B., and Dea, S. (1998). A subset of porcine reproductive and respiratory syndrome virus GP3 glycoprotein is released into the culture medium of cells as a non-virion-associated and membrane-free (Soluble) form. *J. Virol.* **72,** 6298–6306.
11. Meulenberg, J. J. M., Bos-de Ruijter, J. N. A., Van de Graaf, R., Wensvoort, G., and Moormann, R. J. M. (1998). Infectious transcripts from cloned genome-length cDNA of porcine reproductive and respiratory syndrome virus. *J. Virol.* **72,** 380–387.
12. Yang, Z., and Nielsen, R. (1998). Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J. Mol. Evol.* **46,** 409–418.
13. Durand, J. P., Climent, I., Sarraseca, J., Urinza, A., Cortes, E., Vela, C., and Casal, I. (1997). Baculovirus expression of proteins of porcine reproductive and respiratory syndrome virus strain Olot/91. Involvement of ORF3 and ORF5 proteins in protection. *Virus Genes* **14,** 19–29.
14. Meulenberg, J. J. M., Van Nieuwstadt, A. P., Van Essen-Zandbergen, A., and Langeveld, J. P. M. (1997). Posttranslational processing and identification of a neutralization domain of the GP4 protein encoded by ORF4 of Lelystadt virus. *J. Virol.* **71,** 6061–6067.
15. Hedges, J. F., Balasuriya, U. B. R., Timoney, P. J., McCollum, W. H., and MacLachlan, N. J. (1999). Genetic divergence with emergence of novel phenotypic variants of equine arteritis virus during persistent infection. *J. Virol.* **73,** 3672–3681.
16. Snijder, E. J., and Meulenberg, J. J. M. (1998). The molecular biology of arteriviruses. *J. Gen. Virol.* **79,** 961–979.
17. Hopp, T. P., and Woods, K. R. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA* **78,** 3824–3828.
18. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10,** 1–6.
19. Yang, Z., Goldman, N., and Friday, A. (1994). Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. *Mol. Biol. Evol.* **11,** 316–324.
20. Strimmer, K., and von Haeseler, A. (1996). Quartet-puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13,** 964–969.