Patent
File 110.03020201

PRRS VIRUSES, INFECTIOUS CLONES, MUTANTS THEREOF, AND METHODS OF USE

CONTINUING APPLICATION DATA

This application claims the benefit of U.S. Provisional Application Serial No. 60/694,021, filed June 24, 2005, which is incorporated by reference herein.

BACKGROUND

Porcine reproductive and respiratory syndrome virus (PRRSV) is the causative agent of a disease characterized by respiratory disorders in young pigs and reproductive failure in sows (Benfield et al., *J. Vet. Diagn. Invest.*, 4:127-133 (1992); Collins et al., *J. Vet. Diagn. Invest.*, 4:117-126 (1992); Wensvoort et al., *Vet. Q.*, 13:121-130 (1991)) and is now endemic in most countries. The syndrome was first recognized as a "mystery swine disease" in the United States in 1987 and was discovered in Europe in 1990. The two prototype viral strains (Lelystad and VR-2332) differ in nucleotide sequence by approximately 40% and represent two distinct genotypes, referred to as European (EU or Type 1, Lelystad; Meulenberg et al., *Virology*, 192:62-72 (1993)) and North American (NA or Type 2, VR-2332; Nelsen et al., *J. Virol.*, 73:270-80 (1999)) strains (Fang et al., *Virus Res.*, 100:229-235 (2004); Mardassi et al., *J. Gen. Virol.*, 75:681-5 (1994); Meng et al., *Arch. Virol.*, 140:745-55 (1995); Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). The disease has also been referred to as Wabash syndrome, mystery pig disease, porcine reproductive and respiratory syndrome, swine plague, porcine epidemic abortion and respiratory syndrome, blue abortion disease, blue ear disease, abortus blau, and seuchenhafter spatabort der schweine. The disease is characterized by reproductive failure in pregnant sows and respiratory problems in pigs of all ages. The disease has a significant negative impact on the swine industry.

PRRSV is an enveloped, positive-sense RNA virus belonging to the family *Arteriviridae* in the order *Nidovirales* (Cavanagh, *Arch. Virol.*, 142:629:633 (1997)). The PRRSV genome varies from 15.1-15.5 kb long (Meulenberg et al., *Virology*, 192:62-72 (1993); Nelsen et al., *J. Virol.*, 73:270-80 (1999)). The first 75% of the

genome encodes the replicase polyprotein essential for virus replication and is comprised of two large open reading frames (ORFs) (1a and 1b) that are processed cotranslationally into smaller proteins by virally encoded proteases (Snijder et al., *J. Gen. Virol.*, 79:961-79 (1998)). The structural proteins are encoded by seven downstream ORFs and are translated from a 3'-coterminal nested set of subgenomic mRNAs (sgmRNA) (Meulenberg et al., *Virology*, 192:62-72 (1993); Pattnaik et al., *Cell*, 69:1011-1020 (1992)). In strain VR-2332, the coding region of the genome (15,411 bases) is flanked by 5' and 3' nontranslated regions of 189 and 151 nucleotides, respectively.

PRRSV strain VR-2332 has been well characterized in terms of its complete genome sequence (Pattnaik et al., *Cell*, 69:1011-1020 (1992)), the ability of PRRSV to constitutively produce defective subgenomic RNA species termed *heteroclites* (*latin*: uncommon forms) (Yuan et al., *Virology*, 275:158-169 (2000)); Yuan et al., *Virus Research*, 105:75-87(2004)), and its growth properties *in vitro* as well as *in vivo* (Murtaugh et al., *Vet. Immunol. Immunopathol.*, 102:105-349 (2004)). In addition, an infectious clone of this 15.4 kb NA PRRSV genome has been produced and examined for its ability to cause disease in swine (pVR-HN; Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)).

PRRSV continues to cause significant economic losses throughout the world. Vaccines are available, but they are based on one PRRSV strain, and there is evidence that PRRSV strains vary at the antigenic and genetic levels. In addition, since the virus was identified in Europe and in the United States, new disease phenotypes have continued to emerge.

SUMMARY OF THE INVENTION

Prior reports had suggested that deletions and/or mutations of any strain of PRRS virus was often extremely detrimental to viral growth. Specifically, individual laboratories had made mutations in the 3' end of the virus, and the resultant virus was either unstable and quickly reverted back to wild-type sequence, or grew very poorly or not at all (Lee et al., *Virol.*, 331:47-62 (2005); Choi et al., *J. Virol.*, 80:723-736 (2006); Lee et al., *Virolog.*, 346:238-250 (2005)). Thus, in comparison of nucleotide sequences of European (Type 1 genotype) and VR-2332 (Type 2 genotype), where to make mutations in VR-2332 NSP2 that were not extremely detrimental was not known.

However, alignment of the full genome sequences of new Type 2 PRRS viruses with VR-2332 began to provide insight as to where viable mutants could be made. Further deletion mutagenesis showed that the region between nsp2 amino acids 324-813 was not necessary for growth *in vitro*.

The present invention provides an isolated infectious polynucleotide having a nucleotide sequence with at least 88% identity to SEQ ID NO:1 and a deletion of at least 39 consecutive nucleotides selected from nucleotide 2062 to nucleotide 3864 of SEQ ID NO:1. Also provided is an isolated infectious polynucleotide having a nucleotide sequence with at least 88% identity to SEQ ID NO:14 and a deletion of at least 39 consecutive nucleotides selected from nucleotide 2061 to nucleotide 3545 of SEQ ID NO:14. The isolated polynucleotide may be present in a vector, in an isolated virus particle, present in a cell, or a combination thereof. When present in a vector an RNA polymerase promoter may be operably linked to the polynucleotide. The isolated polynucleotide may by an RNA. The isolated polynucleotide may include 2 or more deletions, and each deletion may be independently at least 37 consecutive nucleotides. The isolated polynucleotide may further include an exogenous polynucleotide present in the deletion, and the exogenous polynucleotide may encode a polypeptide, such as a detectable marker.

The present invention also provides an isolated polynucleotide having a nucleotide sequence with at least 88 % identity to SEQ ID NO:1 and at least one deletion of at least 39 consecutive nucleotides selected from nucleotide 2062 to nucleotide 3864 of SEQ ID NO:1, and wherein the polynucleotide replicates and produces infectious virus particles when introduced into a cell. Also provided is an isolated polynucleotide having a nucleotide sequence with at least 88 % identity to SEQ ID NO:14 and at least one deletion of at least 39 consecutive nucleotides selected from nucleotide 2061 to nucleotide 3545 of SEQ ID NO:14, wherein the polynucleotide replicates and produces infectious virus particles when introduced into a cell. The isolated polynucleotide may be present in a vector, in an isolated virus particle, present in a cell, or a combination thereof. When present in a vector an RNA polymerase promoter may be operably linked to the polynucleotide. The isolated polynucleotide may by an RNA. The isolated polynucleotide may include 2 or more deletions, and each deletion may be independently at least 37 consecutive nucleotides. The isolated polynucleotide may further include an exogenous polynucleotide present in the

deletion, and the exogenous polynucleotide may encode a polypeptide, such as a detectable marker.

The present invention further provides an infectious clone having a polynucleotide with a nucleotide sequence having at least 88 % identity to SEQ ID NO:1 and at least one deletion of at least 39 consecutive nucleotides selected from nucleotide 2062 to nucleotide 3864 of SEQ ID NO:1. Also provided is an infectious clone having a polynucleotide with a nucleotide sequence having at least 88 % identity to SEQ ID NO:14 and at least one deletion of at least 39 consecutive nucleotides selected from nucleotide 2061 to nucleotide 3545 of SEQ ID NO:14. The infectious clone may be present in a cell. An RNA polymerase promoter may be operably linked to the polynucleotide. The infectious clone may include 2 or more deletions, and wherein each deletion is independently at least 37 consecutive nucleotides. The isolated polynucleotide may further include an exogenous polynucleotide present in the deletion, and the exogenous polynucleotide may encode a polypeptide, such as a detectable marker.

Also provided by the present invention is an isolated infectious polynucleotide comprising a nucleotide sequence SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13, and an nsp2 polypeptide encoded by an infectious polynucleotide comprising a nucleotide sequence SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13.

The terms "comprises" and variations thereof do not have a limiting meaning where these terms appear in the description and claims. Unless otherwise specified, "a," "an," "the," and "at least one" are used interchangeably and mean one or more than one.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. A. Nucleotide sequence (SEQ ID NO:1) of infectious polynucleotide VR-V7 (also referred to herein as V6G7475A). B. Nucleotide sequence (SEQ ID NO:2) of infectious polynucleotide VR-V5. C. Nucleotide sequence (SEQ ID NO:3) of infectious polynucleotide VR-V5G7475A. D. Nucleotide sequence (SEQ ID NO:4) of infectious polynucleotide VR-V6. E. Nucleotide sequence (SEQ ID NO:5) of

infectious polynucleotide MN184A. F. Nucleotide sequence (SEQ ID NO:6) of infectious polynucleotide MN184B. G. Nucleotide sequence (SEQ ID NO:7) of infectious polynucleotide Nsp2 Δ324-434. H. Nucleotide sequence (SEQ ID NO:8) of infectious polynucleotide Nsp2 Δ324-523. I. Nucleotide sequence (SEQ ID NO:9) of infectious polynucleotide Nsp2 Δ543-632. J. Nucleotide sequence (SEQ ID NO:10) of infectious polynucleotide Nsp2 Δ633-726. L. Nucleotide sequence (SEQ ID NO:11) of infectious polynucleotide Nsp2 Δ543-726. L. Nucleotide sequence (SEQ ID NO:12) of infectious polynucleotide Nsp2 Δ727-813. M. Nucleotide sequence (SEQ ID NO:13) of infectious polynucleotide Nsp2 Δ324-726.

Figure 2. Assembly of full-length clones of PRRSV strain VR-2332. The 15.4 genome was amplified in four sections (I – IV) that incorporated unique restriction enzyme cleavage sites present in viral cDNA (FseI, AvrII, BsrGI) or added to the PRRSV sequence at the 5' and 3' ends by insertion mutagenesis (SphI, Pac I respectively). A T7 polymerase promoter and 2 nontemplated G residues and a T residue preceded the viral sequence. The pOK12 vector (24) was modified to include a PacI site and a hepatitis delta ribozyme downstream of a poly adensine tail of 50 nucleotides.

Figure 3. Schematic of nucleotide changes of infectious clones or swine progeny. Diagram of the PRRSV genome organization is presented under which are full genome comparisons. Putative nonstructural protein cleavages are depicted above ORF1a and 1b, represented by downward arrows. Signature motifs are identified below ORF1a and 1b, with upward arrows indicating their placement in the PRRSV genome [papain-like cysteine protease α and β (PCPα, PCPβ); cysteine protease (CP); serine/3C protease (SP/3CP); polymerase (POL); cysteine/histidine rich (C/H); helicase (Hel); *Xenopus laevis* homolog poly(U)-specific endoribonuclease (XendoU); Ivanov et al., *Proc. Natl. Acad. Sci. USA*, 101:12694-12699 (2004); Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)]. Nucleotide differences are represented by vertical bars. 1. *wt* strain VR-2332 (U87392) compared to VR-2332 derived vaccine (Ingelvac® MLV or RespPRRS, AF066183). 2. *wt* strain VR-2332 compared to pVR-V6G7475A. 3. pVR-V5 compared to *in vivo* passaged V5-1-P3 (Sw612). 4. *wt* strain VR-2332 compared to Sw612. Detailed nucleotide changes are listed in Tables 4 and 5.

Figure 4. Seroconversion of swine after PRRSV infection. Growing swine were infected with native *wt* strain VR-2332 (□), Ingelvac® MLV (×), V5-1 P3 (O) or

remained uninfected (■). At days indicated, serum samples were taken and tested by IDEXX Elisa for indication of seroconversion by anti-PRRSV antibodies to the nucleocapsid protein.

Figure 5. A. Plaque assays on P3 progeny (first lineage) of all infectious clones as well as *wt* strain VR-2332 revealed different plaque sizes. B. Progeny of V5-1 P3 after growth in swine (Sw612) produced plaques similar to *wt* strain VR-2332.

Figure 6. A. Plaque assays on P3 progeny (second lineage) of all infectious clones as well as *wt* strain VR-2332 displayed plaque sizes that were different from first lineage virus preparations. B. Titers of P4 virus indicate infectious clone progeny were not replicating as *wt* strain VR-2332 or Sw612 virus in spite of having similar plaque size.

Figure 7. A. P3 progeny of *wt* strain VR-2332 (◆), Sw612 (▲), pVR-HN (□), pVR-V5 (×), pVR-V5G7475A(∗), pVR-V6 (●), pVR-V6G7475A (○) were simultaneously examined for one step growth kinetics as outlined in Example 1. *wt* strain VR-2332 and Sw612 viruses replicated to approximately 10-fold higher titers at all time points. pVR-V6G7475A, with no amino acid changes from native virus or vaccine, produced virus that replicated to a higher titer at all time points than all other infectious clone progeny. The final titer for each virus preparation is listed in the companion table.

Figure 8. Northern blot analysis of different progeny passages of pVR-V6G7475A as well as Sw612 and the initial *in vitro* transcript reveals heteroclites are produced as early as P1 and, along with genomic RNA, are more abundant with passage. However, transcript RNA (Tx) does not contain readily detectable heteroclite species.

Figure 9. A. Diagrammatic representation of the PRRSV genome. Putative nonstructural protein cleavages are depicted above ORF1a and 1b, represented by downward arrows. Signature motifs are identified below ORF1a and 1b, indicating their placement in the PRRSV genome [papain-like cysteine protease α and β (PL1); cysteine protease (PL2); serine/3C protease (3CL); polymerase (RdRp); helicase (Hel); *Xenopus laevis* homolog poly(U)-specific endoribonuclease (N); Ziebuhr et al., 2000; Ivanov et al., 2004; Gorbalenya et al., 2006]. B. Schematic diagram of the comparison of ORF1 protein (replicase) of MN184A and MN184B and putative processing. The

degeneracy seen in nsp2 is included in the comparison. C. Schematic diagram of the comparison of ORF2-7 proteins of MN184A and MN184B.

Figure 10. ORF5 amino acid sequence alignment of divergent PRRSV. Dark grey boxes indicate high amino acid conservation (>80%; between 16 and 19 residues are identical), medium grey (>60%; between 12 and 15 residues are identical), lighter grey (>40%; between 8 and 11 residues are identical) and unshaded (<40%; less than 8 residues are identical) boxes identify less conserved residues. The dashed region indicates the putative signal sequence, the boxed regions identify the proposed transmembrane regions, the hypervariable regions are indicated (HV-1 and HV-2), and the proposed orientation of the protein in the virion is identified in bold italics. The conserved cysteine residue that is proposed to interact with the M protein is identified by the downward arrow (↓). The two conserved putative N-glycosylation sites are identified by stars and hypervariable region 1 contains strain/isolate specific N-glycosylation sites (NxS/T). The following GenBank full-length sequences were used for comparison: VR-2332 (*U87392*), Ingelvac MLV (*AF066183*), 01NP1.2 (*DQ056373*), PL97-1 (*AY58524*), PA -8 (*AF176348*), SP (*AF184212*), BJ-4 (*AF331831*), HN1 (*AY457635*), 16244B (*AF046869*), HB-1 (*AY150312*), HB-2 (*AY262352*), CH-1a (*AY032626*), P129 (*AF494042*), JA142 (*AY424271*), SDPRRS-01-08 (*AY375474*), EuroPRRSV (*AY366525*), Lelystad (*M96262*), IAF-93-653 (*U64931*), IAF-Klop (*AY184209*), 98-3298 (*DQ306877*), 98-3403 (*DQ306878*), 99-3584 (*DQ306879*).

Figure 11. Nsp1β amino acid sequence alignment of divergent PRRSV. The figure derivation and color scheme was described in the Figure 10 legend. The two completely conserved putative catalytic residues are identified by stars and the boxed amino acids identify MN184 sequence conservation with Type 1 isolates and EAV. The proposed cleavage site is identified by the downward arrow (⬇).

Figure 12. Nsp2 amino acid sequence alignment of divergent PRRSV. The completely conserved putative cysteine protease catalytic residues (Cys and His) are identified by stars and the boxed amino acids signify protease sequence conservation within PRRSV and EAV. The proposed cleavage sites are identified by filled arrows (⬇); additional possible cleavage sites are indicated by a hashed arrow; signal peptide, solid grey box; transmembrane regions, shown in hashed black boxes; potential N-

glycosylation sites, indicated by an asterisk (*). The figure derivation and color scheme were described in the Figure 10 legend.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present invention includes infectious clones of the Porcine reproductive and respiratory syndrome virus (PRRSV) VR-2332. As used herein, the term "infectious clone" is a polynucleotide having two components; a vector sequence that replicates in a prokaryotic host cell, and a second polynucleotide referred to herein as an infectious polynucleotide. When transcribed *in vitro* to yield an RNA polynucleotide and introduced into a permissive cell, the infectious polynucleotide replicates (as an RNA) and produces infectious virus particles. Thus, an infectious polunucleotide can be present in a vector as a DNA, as an RNA in a virus particle, or as an isolated DNA or RNA. The term "polynucleotide" refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxynucleotides, and includes both double- and single-stranded DNA and RNA. Unless otherwise noted, a polynucleotide includes the complement thereof. The nucleotide sequence of the complement of a polynucleotide can be easily determined by a person of skill in the art. A polynucleotide may include nucleotide sequences having different functions, including for instance coding sequences, and non-coding sequences such as regulatory sequences and/or untranslated regions. A polynucleotide can be obtained directly from a natural source, or can be prepared with the aid of recombinant, enzymatic, or chemical techniques. A polynucleotide can be linear or circular in topology. A polynucleotide can be, for example, a portion of a vector, such as an expression or cloning vector, or a fragment.

If naturally occurring, a polynucleotide is preferably isolated, more preferably, purified. An "isolated" compound, such as a polynucleotide, polypeptide, or virus particle, is one that is separate and discrete from its natural environment. A "purified" compound is one that is at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated. Compounds such as polynucleotides and polypeptides that are produced outside the organism in which they naturally occur, e.g., through chemical or recombinant means, are considered to be isolated and purified by definition, since they were never present in a natural environment.

An example of an infectious polynucleotide of the present invention includes the infectious polynucleotide VR-V7 (SEQ ID NO:1). VR-V7 is also referred to herein as V6G7475A. Other examples of infectious polynucleotides of the present invention include VR-V5 (SEQ ID NO:2), VR-V5G7475A (SEQ ID NO:3), and VR-V6 (SEQ ID NO:4). It should be noted that while SEQ ID NOs:1, 2, 3, 4, 5, 6 and other virus nucleotide sequences are disclosed herein as a DNA sequence, the present invention contemplates the corresponding RNA sequence, and RNA and DNA complements thereof, as well.

Other infectious polynucleotides of the present invention have a polynucleotide sequence having structural similarity to a reference polynucleotide. Reference polynucleotides include SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, the European prototype strain of PRRS virus, Lelystad (Genbank accession number M96262; SEQ ID NO:14), and the North American prototype strain of PRRS virus, VR-2332 (Genbank accession number U87392; SEQ ID NO:15). The similarity is referred to as "percent identity" and is determined by aligning the residues of the two polynucleotides (i.e., the nucleotide sequence of a candidate infectious polynucleotide and the nucleotide sequence of the reference polynucleotide) to optimize the number of identical nucleotides along the lengths of their sequences; gaps in either or both sequences are permitted in making the alignment in order to optimize the number of shared nucleotides, although the nucleotides in each sequence must nonetheless remain in their proper order. In some aspects of the present invention the gap (also referred to as a deletion) is present in the candidate infectious polynucleotide sequence. A candidate infectious polynucleotide is the polynucleotide that has the nucleotide sequence being compared to the reference polynucleotide. A candidate infectious polynucleotide can be isolated from an animal, such as a pig infected with PRRSV, isolated from a cultured cell line, or can be produced using recombinant techniques, or chemically or enzymatically synthesized. Two nucleotide sequences can be compared using any of the commercially available computer algorithms routinely used to produce alignments of nucleotide sequences. Preferably, two nucleotide sequences are compared using the GAP program of the GCG Wisconsin Package (Accelrys, Inc.) version 10.3 (2001). The GAP program uses the algorithm of Needleman et al. (*J. Mol. Biol.*, 48:443-453 (1970)) to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number

of gaps. Preferably, the default values for all GAP search parameters are used, including scoring matrix = NewsgapDNA.cmp, gap weight = 50, length weight = 3, average match = 10, average mismatch = 0. In the comparison of two nucleotide sequences using the GAP search algorithm, structural similarity is referred to as "percent identity." Preferably, a polynucleotide has structural similarity with a reference polynucleotide of at least 88 %, at least 89 %, at least 90 %, at least 91 %, at least 92 %, at least 93 %, at least 94 %, at least 95 %, at least 96 %, at least 97 %, at least 98 %, or at least 99 % identity when the structural similarity is determined using the GAP program.

Whether a polynucleotide is an infectious polynucleotide can be determined by inserting into a vector a candidate infectious polynucleotide, transcribing the candidate infectious polynucleotide *in vitro*, transfecting a permissive cell with the resulting RNA molecules, and detecting progeny viral RNA, progeny viral nucleocapsid protein, detecting infectious virus particles, or a combination thereof. The vector preferably has the characteristics of being low copy number and remains stable after insertion of large (e.g., 15 kb) inserts. An example of a suitable vector is pOK and pOK12 (GenBank Accession AF223639, Vieira et al., *Gene*, 100:189-194 (1991)), and other vectors having these characteristics are known and available. In the vector the candidate infectious polynucleotide is immediately downstream of a promoter. Useful promoters are those that can be induced to yield high levels of transcription, such as a T7 RNA polymerase promoter, for example TAATACGACTCACTATA (SEQ ID NO:16), or the RNA polymerase promoters SP6 and T3. Transcription of the candidate infectious polynucleotide typically includes restriction endonuclease digestion of the vector to make it linear, and producing RNA transcripts by use of routine and well known *in vitro* transcription methods. Kits for *in vitro* transcription are commercially available (for instance, mMessage mMachine, available from Ambion, Austin, TX).

After *in vitro* transcription the RNA is purified using routine methods and then used to transfect a permissive cell. Examples of permissive cells include, for instance, BHK-21 (which allows one round of virus particle production), CL-2621, MA-104 (ATCC CRL-2378), MARC-145 (Kim et al., *Arch. Virol.*, 133:477-483 (1993)), cell lines cloned from these cell lines, or primary porcine alveolar macrophages. Methods for efficiently transfecting cells include the use of 1,2-dimyristyloxypropyl-3-dimethyl-hydroxy ethyl ammonium bromide and cholesterol (DMRIE-C), and other

commercially available products, preferably, DMRIE-C. Methods for efficiently transfecting primary porcine alveolar macrophages are known to the art (Groot Bramel-Verheige et al., *Virol.*, 278:380-389 (2000)). Generally, 2 to 3 micrograms of RNA can be used for trasnfection, but lower and higher amounts may be used. After a suitable period of time, the presence of progeny viral RNA can be detected by, for instance, reverse transcriptase-polymerase chain reaction (RT-PCR). Likewise, progeny viral nucleocapsid protein can be detected by, for instance, nucleocapsid specific antibody. Further, whether the virus particles produced by cells transfected with a candidate infectious polynucleotide will infect another cell can be detected by exposing uninfected permissive cells to supernatant from infected cells. Optionally, cytopathic effect (CPE) may be observed. A candidate infectious polynucleotide is considered to be an infectious polynucleotide when it produces progeny viral RNA, progeny viral proteins (nucleocapsid, membrane, GP5, and others), and infects other permissive cells.

In some aspects of the present invention an infectious polynucleotide includes a deletion of nucleotides encoding non-structural protein 2 (nsp2), one of several (12 predicted) polypeptides present in the polyprotein encoded by ORF1. In a PRRS virus, and infectious polynucleotides thereof, the nucleotides encoding the first amino acid of nsp2 can be determined by identifying the cleavage site of papain-like protease 1 beta, predicted to be after the ORF1 amino acid glycine at position 383 in VR-2332.

With respect to identifying the nucleotides encoding the last amino acid of nsp2, the exact nsp2 C-terminal cleavage site of the ORF1a-encoded polyprotein has not been empirically determined, thus the nucleotides corresponding to the 3' end of the coding region are unknown. However, two predictions of the C-terminal cleavage site have been proposed, one Gly|Gly (where the vertical line between the two glycine residues indicates the cleavage location) at amino acid 980 in VR-2332, and the other at amino acid 1197 in VR-2332. In alignment of all available PRRSV sequences, there are several completely conserved Gly|Gly doublets within this protein that may also be the nsp2 C terminal cleavage site of the polyprotein (amino acids 646, 980, 1116, 1196, 1197,in VR-2332. The locations of the Gly|Gly doublets in the other viruses and infectious polynucleotides can be identified by comparison to the sequences of nsp2 and the Gly|Gly doublets disclosed in Figure 12. Present studies suggest that there may be at least 3 cleavage sites in nsp2 , corresponding to amino acid 980, 116, 1196 or 1197.

The nsp2 polypeptide includes a highly conserved chymotrypsin-like cysteine protease domain (identified as CP in Figure 3 and PL2 in Figure 9) present at the N-terminus, and 3-4 predicted transmembrane domains near the C terminus of nsp2 (where the number of transmembrane domains varies depending on the location of the C-terminal cleavage site). Typically, deletion of the nucleotides encoding the amino acids of the PL2 domain or all of the predicted transmembrane domains results in a polynucleotide that can replicate in permissive cells but will not produce infectious virus particles. Thus, an infectious clone of the present invention does not typically include deletion of the entire PL2 domain or all of the predicted transmembrane domains.

The nucleotides encoding the chymotrypsin-like cysteine protease domain are nucleotides 1474 to 1776 of VR-V7 (SEQ ID NO:1), nucleotides 1474 to 1776 of VR-2332 (Genbank accession number U87392), and nucleotides 1482 to 1784 of Lelystad (Genbank accession number M96262). The location of a chymotrypsin-like cysteine protease domain in the nucleotide sequence of other PRRS viruses can be identified by aligning the amino acid sequence of the nsp2 polypeptide encoded by a PRRS virus with the amino acid sequence alignment disclosed in Figure 12, and determining which nucleotides encode those amino acids that line up with the chymotrypsin-like cysteine protease domain. Alternatively, the amino acid sequences of nsp2 polypeptides of other PRRS viruses can be identified by aligning the amino acid sequence of the nsp2 polypeptide encoded by a PRRS virus with the amino acid sequence of nsp2 polypeptides produced by other arteriviruses, such as equine arteritis virus (EAV) and lactate dehydrogenase-elevating virus (LDV).

The nucleotides encoding the predicted transmembrane domains of VR-V7 (SEQ ID NO:1), VR-2332 (Genbank accession number U87392), and Lelystad (Genbank accession number M96262) are shown in Table 1.

Table 1. Nsp2 nucleotides encoding predicted transmembrane domains.

	VR-V7	VR-2332	Lelystad
Transmembrane domain I	881 to 901	881 to 901	761 to 781
Transmembrane domain II	913 to 934	913 to 934	793 to 814
Transmembrane domain III	963 to 980	963 to 980	843 to 860
Transmembrane domain IV	985 to 1003	985 to 1003	865 to 883

The location of the transmembrane domains in the nucleotide sequence of other PRRS viruses can be identified by aligning the amino acid sequence of the nsp2 poplypeptide encoded by a PRRS virus with the amino acid sequence alignment disclosed in Figure 12, and determining which nucleotides encode those amino acids that line up with the transmembrane domains. Alternatively, the location of the transmembrane domains can be identified with a computer algorithm, such as the PredictProtein algorithm as described by Rost et al. (*Nucleic Acids Res.*, 32(Web Server issue):W321-326 (2004), or the TMHMM algorithm as described by Krogh et al. (*J. Mol. Biol.*, 305:567-580 (2001)) and available through the World Wide Web.

The deletion present in infectious polynucleotides of the present invention is typically between the nucleotides encoding the chymotrypsin-like cysteine protease domain and the nucleotides encoding the transmembrane domains, and does not result in a frameshift in the reading frame of ORF1. As discussed above, the deletion typically does not include all the nucleotides encoding the chymotrypsin-like cysteine protease domain, all the nucleotides encoding the transmembrane domains, or the combination thereof. In some aspects, for instance when the infectious polynucleotide has structural similarity with SEQ ID NO:1, the 5' boundary of a deletion is at nucleotide 2305, nucleotide 2205, nucleotide 2105, or nucleotide 2062, and the 3' boundary of a deletion is at nucleotide 3774, nucleotide 3804, nucleotide 3834, or nucleotide 3864. In other aspects, for instance when the infectious polynucleotide has structural similarity with SEQ ID NO:14, the 5' boundary of a deletion is at nucleotide 2304, nucleotide 2204, nucleotide 2104, or nucleotide 2061, and the 3' boundary of a

deletion is at nucleotide 3455, nucleotide 3495, nucleotide 3525, or nucleotide 3545. The deletion can be at least 39 nucleotides, 48 nucleotides, or 57 nucleotides. In some aspects, the deletion can be at least 267 nucleotides, at least 276 nucleotides, or at least 285 nucleotides. In some aspects the deletion is no greater than 489 nucleotides, no greater than 459, no greater than 429, or no greater than 402 nucleotides. An infectious polynucleotide may have more than one deletion in the nsp2 region.

Examples of infectious polynucleotides derived from VR-V7 and containing a deletion are disclosed in Table 2.

Table 2. Infectious polynucleotides derived from VR-V7 (SEQ ID NO:1).

Polynucleotide*	deleted nucleotides of SEQ ID NO:1	amino acids of ORF1 deleted	viral titlers (PFU/ml)	Summary of phenotype**
Nsp2 Δ180-323	1876-2304	563-705	-	nonviable
Nsp2 Δ242-323	2056-2304	623-705	-	nonviable
Nsp2 Δ324-434	2305-2637	706-816	+ (~10^5)	small plaque size
Nsp2 Δ324-523	2305-2904	706-905	+ (~10^5-10^6)	intermediate
Nsp2 Δ543-632	2962-3231	925-1014	+ (~10^5)	small plaque size
Nsp2 Δ633-726	3232-3513	1015-1108	+ (~10^5)	small plaque size
Nsp2 Δ543-726	2962-3513	925-1108	+ (~10^5)	small plaque size
Nsp2 Δ727-813	3514-3774	1109-1195	+ (~10^5)	small plaque size
Nsp2 Δ324-726	2305-3513	706 -1108	+ (~10^{1-2})	ND
Nsp2 Δ324-813	2305-3774	706-1195	-	nonviable
Nsp2 Δ727-845	3514-3870	1109-1227	-	nonviable
Nsp2 Δ324-845	2305-3870	706-1227	-	nonviable

* the deletion refers to the amino acids of nsp2 that are deleted, e.g., in the virus Nsp2 Δ180-323, amino acids 180-323 of nsp2 are deleted.

**plaque size is relative to plaques produced by wildtype VR-2332.

An infectious polynucleotide containing a deletion can include an exogenous polynucleotide inserted in place of the deletion. An "exogenous" polynucleotide refers to a foreign nucleotide sequence, i.e., a nucleotide sequence that is not normally present in a PRRS virus or an infectious clone thereof. The exogenous polynucleotide can, and preferably does encode a polypeptide. Suitable exogenous polynucleotides include those encoding a detectable marker, e.g., a molecule that is easily detected by various methods. Examples include fluorescent polypeptides (e.g., green, yellow, blue, or red fluorescent proteins), luciferase, chloramphenicol acetyl transferase, and other molecules (such as c-myc, flag, 6xhis, HisGln (HQ) metal-binding peptide, and V5 epitope) detectable by their fluorescence, enzymatic activity or immunological properties, and are typically useful when detected in a cell, for instance, a cultured cell, or a tissue sample that has been removed from an animal. Other exogenous polynucleotides that can be used are those encoding polypeptides expressed by other entities, such as cells and pathogens. Expression of an exogenous polynucleotide results in an infectious polynucleotide that expresses foreign antigens. Examples of exogenous nucleotide sequences include those encoding proteins expressed by pathogens, preferably porcine pathogens, such as porcine circovirus type 2, *Mycoplasma hyopneumoniae* (e.g., the P46 and P65 proteins of *M. hyopneumoniae*), *Lawsonia intracellularis* (e.g., the outer membrane proteins of *L. intracellularis*), the ORF5 of different strains of PRRSV, and *Streptococcus suis* (e.g., the 38-kDa protein of *S. suis*). The nsp2 polypeptide has B-cell epitopes and is expected to be immunogenic. Inclusion of foreign epitopes in an nsp2 polypeptide is expected to result in an immune response to the foreign epitopes. Additional examples of exogenous polynucleotides include those encoding biological response modifiers, such as, for example, IFN- α, IFN- γ, IL-12, IL-2, TNF-α, and IL-6.

The exogenous polynucleotide is inserted into the deletion region such that it is in frame with the open reading frame encoding nsp1α and nsp1β, and more than one exogenous polynucleotide can be inserted in tandem, for instance, nucleotide sequences encoding three c-myc epitopes can be present. The total size of the infectious polynucleotide containing an exogenous polynucleotide inserted in the place of the deletion is typically no greater than 16,000 bases, no greater than 15,800 based, no greater than 15,600 bases, no greater than 15,400 bases, or no greater than 15,200 based

(including the poly A tail). An insertion can be present in an infectious polynucleotide having the Nsp2 Δ324-434, Nsp2 Δ324-523, Nsp2 Δ543-632, Nsp2 Δ633-726, Nsp2 Δ543-726, Nsp2 Δ727-813, or Nsp2 Δ324-726 deletion, preferably, the Nsp2 Δ324-434, Nsp2 Δ543-632, Nsp2 Δ633-726, Nsp2 Δ543-726, Nsp2 Δ727-813, or Nsp2 Δ324-726 deletion. Preferred examples of infectious clones containing an exogenous polynucleotide in the location of a deletion include an infectious polynucleotide having the Nsp2 Δ324-434 deletion containing a coding region encoding a 238 amino acid green fluorescent protein, an infectious polynucleotide having the Nsp2 Δ543-632 deletion containing a coding region encoding a 238 amino acid green fluorescent protein, an infectious polynucleotide having the Nsp2 Δ324-434 deletion containing a coding region encoding a 10 amino acid *c-myc* epitope (EQKLISEEDL, SEQ ID NO:17), an infectious polynucleotide having the Nsp2 Δ324-434 deletion containing a coding region encoding a 10 amino acid *c-myc* epitope, and an infectious polynucleotide having the Nsp2 Δ324-726 or Nsp2 Δ543-726 deletions each containing a coding region encoding tandem repeat of the 10 amino acid *c-myc* epitope.

An infectious polynucleotide is typically present in a vector, and the combination of infectious polynucleotide and vector is referred to as an infectious clone, which is made through reverse genetics. A vector is a replicating polynucleotide, such as a plasmid, phage, or cosmid, to which another polynucleotide may be attached so as to bring about the replication of the attached polynucleotide. Construction of vectors containing a polynucleotide of the invention employs standard recombinant DNA techniques known in the art (see, e.g., Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989)). A vector can provide for further cloning (amplification of the polynucleotide), i.e., a cloning vector, or for expression of the polypeptide encoded by the coding region, i.e., an expression vector, or the combination thereof. The term vector includes, but is not limited to, plasmid vectors, viral vectors, cosmid vectors, or artificial chromosome vectors. Typically, a vector is capable of replication in a bacterial host, for instance *E. coli*. Preferably the vector is a plasmid.

Selection of a vector depends upon a variety of desired characteristics in the resulting construct, such as a selection marker, vector replication rate, and the like. Preferably, a vector suitable for use as part of an infectious clone is both a cloning vector and an expression vector. Useful vectors have a low copy number in a host cell.

Suitable host cells for cloning or expressing the vectors herein are prokaryote or eukaryotic cells. Preferably the host cell secretes minimal amounts of proteolytic enzymes. Suitable prokaryotes include eubacteria, such as gram-negative organisms, for example, *E. coli* or *S. typhimurium*. Examplary host cells useful for making, manipulating, and maintaining an infectious clone are DH-5α, DH-1 (ATCC 33849), and AG-1, preferably, DH-1 or AG-1.

A vector includes regulatory sequences operably linked to the infectious polynucleotide. The term "operably linked" refers to a juxtaposition of components such that they are in a relationship permitting them to function in their intended manner. A regulatory sequence is "operably linked" to an infectious polynucleotide of the present invention when it is joined in such a way that expression of the coding region is achieved under conditions compatible with the regulatory sequence. Typically, a promoter is one that provides for high specificity binding of an RNA polymerase, and such promoters include T7, SP6, and T3. Typically the promoter is situated immediately upstream of the first nucleotide of the infectious polynucleotide. Preferably, a GGT is inserted between the promoter and the first nucleotide of the infectious polynucleotide. Optionally and preferably the vector also contains a hepatitis delta virus ribozyme downstream of the poly A region.

The vector optionally, and preferably, includes one or more selection marker sequences, which typically encode a molecule that inactivates or otherwise detects or is detected by a compound in the growth medium. For example, the inclusion of a selection marker sequence can render the transformed cell resistant to an antibiotic, or it can confer compound-specific metabolism on the transformed cell. Examples of a selection marker sequence are sequences that confer resistance to kanamycin, ampicillin, chloramphenicol, tetracycline, and neomycin.

When producing a deletion of nucleotides encoding an nsp2 polypeptide in an infectious clone, standard recombinant DNA techniques known in the art can be used (see, e.g., Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989)). As the skilled person will recognize, it is standard practice during construction of an infectious clone (and when construction deletions in an infectious clone) to verify by nucleotide sequence analysis the presence of expected nucleotide sequences, such as deletions or other alterations and the absence of other mutations. Likewise, when a candidate infectious polynucleotide is tested to determine

if it is infectious, it is standard practice to verify by nucleotide sequence analysis the absence of contaminating wild-type virus.

The present invention also includes isolated infectious polynucleotides disclosed at SEQ ID NO:5 and SEQ ID NO:6, and infectious polynucleotides having structural similarity to SEQ ID NO:5 or SEQ ID NO:6. Methods for determining structural similarity are described herein. Preferably, an infectious polynucletoides of this aspect of the present invention has structural similarity to SEQ ID NO:5 or SEQ ID NO:6 of at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%. A polynucleotide having structural similarity to SEQ ID NO:5 or SEQ ID NO:6 is considered to be an infectious polynucleotide if, when present in a virus particle and exposed to permissive cells, the polynucleotide replicates in the permissive cells and produces infectious virus particles.

The present invention also includes isolated virus particles. As used herein, the terms "virus particle" and "viral particle" are used interchangeably and refer to a polynucleotide of the present invention surrounded by an envelope. A virus particle of the present invention can, when added to a permissive cultured cell, can replicate to result in the production of more viral particles.

A virus particle can be grown by passage *in vivo* or in cell culture. Passage *in vivo* includes inoculating a pig (Faaberg et al., U.S. Patent 7,041,443). Passage in cell culture includes exposing cultured cells to the virus particle and incubating the cells under conditions suitable for the virus to reproduce and produce more virus particles. Preferably, the cultured cells are not an immortalized or transformed cell line (i.e., the cells are not able to divide indefinitely). Preferably, primary porcine alveolar macrophages are used for passage in cell culture (Faaberg et al., U.S. Patent 7,041,443).

A virus of the present invention can be inactivated, i.e., rendered incapable of reproducing *in vivo* and/or in cell culture. Methods of inactivation are known to the art and include, for instance, treatment of a virus particle of the invention with a standard chemical inactivating agent such as an aldehyde reagent including formalin, acetaldehyde and the like; reactive acidic alcohols including cresol, phenol and the like; acids such as benzoic acid, benzene sulfonic acid and the like; lactones such as beta propiolactone and caprolactone; and activated lactams, carbodiimides and carbonyl

diheteroaromatic compounds such as carbonyl diimidazole. Irradiation such as with ultraviolet and gamma irradiation can also be used to inactivate the virus.

Also included in the present invention are attenuated virus particles (i.e., viruses having reduced ability to cause the symptoms of mystery swine disease in pigs), and methods of making an attenuated virus particle. Methods of producing an attenuated virus are known to the art. Typically, a virus of the present invention is passaged, i.e., used to infect a cell in culture, allowed to reproduce, and then harvested. This process is repeated until the virulence of the virus in pigs is decreased. For instance, the virus can be passaged 10 times in cell culture, and then the virulence of the virus measured. If virulence has not decreased, the virus that was not injected into the animal is passaged an additional 10 times in cell culture. This process is repeated until virulence is decreased. In general, virulence is measured by inoculation of pigs with virus, and evaluating the presence of clinical symptoms and/or LD_{50} (see, for instance, Halbur et al., *J. Vet. Diagn. Invest.*, 8:11-20 (1996), Halbur et al., *Vet. Pathol.*, 32:200-204 (1995), and Park et al., *Am. J. Vet. Res.*, 57:320-323 (1996)). Preferably, virulence is decreased so the attenuated virus does not cause the death of animals, and preferably does not cause clinical symptoms of the disease.

Typically, a cell culture useful for producing an attenuated virus of the present invention includes cells of non-porcine mammal origin. Examples of non-porcine mammal cell cultures include, for instance, the cell line MA-104 (ATCC CRL-2378), the cell line MARC-145 (Kim et al., *Arch. Virol.*, 133:477-483 (1993)), and the cell line CL-2621 (Baustita et al., *J. Vet. Diagn. Invest.*, 5:163-165 (1993)). Preferably, a mixed cell culture is used for producing an attenuated virus particle of the present invention. In a mixed cell culture there are at least two types of cells present. Preferably, a mixed cell culture includes an immortalized or transformed cell line and a primary cell culture. A mixed cell culture is particularly useful when a virus reproduces slowly, or not at all, in an immortalized or transformed cell line. Preferred examples of an immortalized or transformed cell line for use in a mixed cell culture include, for example, the cell line MARC-145 (Kim et al., *Arch. Virol.*, 133:477-483 (1993)), and the cell line MA-104 (ATCC CRL-2378). Preferably, primary cell cultures for use in a mixed cell culture are porcine in origin. A preferred example of a primary cell culture for use in a mixed cell culture is primary porcine alveolar macrophages.

The present invention further includes the polypeptides encoded by the nsp2 coding regions present in the polynucleotides disclosed in Table 2, including those that are viable. Also included in the present invention are antibodies, including monoclonal and polyclonal antibodies, that specifically bind a polypeptide encoded by the nsp2 coding regions present in the polynucleotides disclosed in Table 2. The term "antibody," unless specified to the contrary, includes fragments of whole antibodies which retain their binding activity for a target antigen. Such fragments include Fv, F(ab') and $F(ab')_2$ fragments, as well as single chain antibodies (scFv). As used herein, an antibody that can "specifically bind" a polypeptide is an antibody that interacts only with the epitope of the antigen that induced the synthesis of the antibody, or interacts with a structurally related epitope. An antibody that "specifically binds" to an epitope will, under the appropriate conditions, interact with the epitope even in the presence of a diversity of potential binding targets. As used herein, the term "polypeptide:antibody complex" refers to the complex that results when an antibody specifically binds to a polypeptide, or a subunit or analog thereof. In some aspects, an antibody of the present invention include those that do not specifically bind to a full length nsp2 polypeptide encoded by VR-2332 (e.g., Genbank accession number U87392, ORF1 amino acids 384-1363 (also see Allende et al. *J. Gen. Virol.*, 80:307-315 (1999) or ORF1 amino acids 384-1580 (also see Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)). Such antibodies can be identified using routine methods known in the art.

Antibodies of the present invention can be prepared using the intact polypeptide. Optionally, an nsp2 polypeptide described herein can be covalently bound or conjugated to a carrier polypeptide to improve the immunological properties of the polypeptide. Useful carrier polypeptides are known in the art.

The preparation of polyclonal antibodies is well known. Polyclonal antibodies may be obtained by immunizing a variety of warm-blooded animals such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice, hamsters, guinea pigs and rats as well as transgenic animals such as transgenic sheep, cows, goats or pigs, with an immunogen. The resulting antibodies may be isolated from other proteins by using an affinity column having an Fc binding moiety, such as protein A, or the like.

Monoclonal antibodies can be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see, for example,

Antibodies: A Laboratory Manual, Harlow et al., eds., Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, (1988)). Monoclonal antibodies can be isolated and purified from hybridoma cultures by techniques well known in the art.

In some embodiments, the antibody can be recombinantly produced, for example, by phage display or by combinatorial methods. Phage display and combinatorial methods can be used to isolate recombinant antibodies that bind to a polypeptide described herein, or a biologically active subunit or analog thereof (see, for example, Ladner et al., U.S. Pat. No. 5,223,409). Such methods can be used to generate human monoclonal antibodies.

The present invention also provides compositions including an infectious polynucleotide, PRRS polynucleotide, virus particle, or antibody of the present invention. Such compositions typically include a pharmaceutically acceptable carrier. As used herein "pharmaceutically acceptable carrier" includes saline, solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Additional active compounds can also be incorporated into the compositions.

A composition may be prepared by methods well known in the art of pharmacy. In general, a composition can be formulated to be compatible with its intended route of administration. Examples of routes of administration include perfusion and parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), and transmucosal. Solutions or suspensions can include the following components: a sterile diluent such as water for administration, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates; electrolytes, such as sodium ion, chloride ion, potassium ion, calcium ion, and magnesium ion, and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. A composition can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Compositions can include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile solutions

or dispersions. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). A composition is typically sterile and, when suitable for injectable use, should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile solutions can be prepared by incorporating the active compound (i.e., an infectious polynucleotide or PRRS virus of the present invention) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle, which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, e.g., gelatin capsules. These compositions may also be formed into a powder or suspended in an aqueous solution such that these powders and/or solutions can be added to animal feed or to the animals' drinking water. These compositions can be suitably sweetened or flavored by various known agents to promote the uptake of the vaccine orally by the pig.

The active compounds can also be administered by any method suitable for administration of polynucleotide agents, e.g., using gene guns, bio injectors, and skin patches as well as needle-free methods such as the micro-particle DNA vaccine technology disclosed by Johnston et al. (U.S. Pat. No. 6,194,389). Additionally, intranasal delivery is possible, as described in, for instance, Hamajima et al., *Clin. Immunol. Immunopathol.*, 88:205-210 (1998). Liposomes and microencapsulation can also be used.

The active compounds may be prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Such formulations can be prepared using standard techniques. The materials can also be obtained commercially from, for instance, Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art.

Toxicity and therapeutic efficacy of such active compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds which exhibit high therapeutic indices are preferred.

The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in the field. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration used.

The compositions can be administered one or more times per day to one or more times per week, including once every other day. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present.

Moreover, treatment of a subject with an effective amount of a polypeptide can include a single treatment or, preferably, can include a series of treatments.

The present invention includes methods for using the compositions described herein. In one aspect the invention includes methods for treating one or more symptoms of mystery swine disease in an animal that may be caused by infection by a PRRS virus. The method includes administering an effective amount of a composition of the present invention to an animal having or at risk of having mystery swine disease, or symptoms of mystery swine disease.

Treatment of mystery swine disease, or symptoms of mystery swine disease, can be prophylactic or, alternatively, can be initiated after the development of disease or symptoms thereof. As used herein, the term "symptom" refers to objective evidence in a subject of mystery swine disease. Symptoms associated with mystery swine disease and the evaluations of such symptoms are routine and known in the art. Examples of symptoms include abortion, anorexia, fever, lethargy, pneumonia, red/blue discoloration of ears, labored breathing (dyspnea), and increased respiratory rate (tachypnea). Treatment that is prophylactic, for instance, initiated before a subject manifests symptoms of a condition caused by a PRRS virus, is referred to herein as treatment of a subject that is "at risk" of developing the disease or symptoms thereof. Typically, an animal "at risk" is an animal present in an area where animals having the disease or symptoms thereof have been diagnosed and/or is likely to be exposed to a PRRS virus. Accordingly, administration of a composition can be performed before, during, or after the occurrence of the conditions described herein. Treatment initiated after the development of a condition may result in decreasing the severity of the symptoms of one of the conditions, or completely removing the symptoms.

In some aspects, the methods typically include administering to an animal a composition including an effective amount of a virus particle of the present invention. An "effective amount" is an amount effective to prevent the manifestation of symptoms of mystery swine disease, decrease the severity of the symptoms of the disease, and/or completely remove the symptoms. Typically, the effective amount is an amount that results in a humoral and/or cellular immune response that protects the animal during future exposure to a PRRS virus. The virus particle used in the composition may contain an infectious polynucleotide that has a deletion as described herein. Optionally, the infectious polynucleotide also includes an exogenous polynucleotide present at the

location of the deletion. An advantage of using a virus particle having a deletion (or an exogenous polynucleotide present in the location of the deletion) is it can be easily distinguished from other PRRS viruses, including wild type PRRS viruses present in the field. The virus particle can be identified by isolation of the virus from an animal followed, for instance, by sequencing, restriction enzyme digestion, or PCR-based amplification of specific nucleotides. Such a "marked" virus particle is often referred to in the art as a marker vaccine.

In other aspects of the present invention the infectious clones and/or infectious polynucleotides described herein can be used to investigate viable gene insertions, to investigate alternative expressed RNA or proteins other than full length virus, to investigate viral recombination, and to investigate immunogenic proterties of full-length nsp2 as relative to truncated nsp2.

EXAMPLES

Example 1

Full-length cDNA clones of North American porcine reproductive and respiratory syndrome virus (PRRSV) prototype VR-2332 strain were developed, with each progressive version possessing less nucleotide changes than prior versions when compared to *wt* strain VR-2332. Progeny virus of each infectious clone was recovered and analyzed for nucleotide sequence verification, *in vitro* growth rate and plaque size. Progeny from one infectious clone confirmed robust *in vivo* replication, seen by the appearance of α-PRRSV antibodies at the same rate as *wt* virus. Northern blot analysis of the *in vivo* progeny also revealed that defective subgenomic RNA species, termed heteroclites (uncommon forms), were present along with full-length genomes. Concurrent northern blot analysis of a passage series of infected MA-104 cell cultures revealed that recombinant virus only gradually gained a profile of both full-length and heteroclite RNA similar to the RNA species seen in *in vivo* infection.

Materials and Methods

Cells and viral strains. MA-104 cells or its descendent MARC-145 cells (ATCC CRL-11171), an African green monkey kidney epithelial cell line which supports PRRSV replication (Meng et al., *J. Vet. Diagn. Invest.*, 8:374-81 (1996)), were

maintained in Eagle's minimal essential medium (EMEM) (JRH Biosciences 56416), supplemented with 1 mg/ml $NaHCO_3$ and 10% fetal bovine serum (FBS), at 37°C with 5% CO_2. The cultured cells were transfected with RNA or infected with virus when monolayer growth had reached 70-80% confluency. PRRSV North American prototype strains VR-2332 and Ingelvac® MLV have been described previously (Yuan et al., *Virus Res.*, 79:189-200 (2001)). Strain VR-2332 grows to equivalent titers on both cell lines.

Viral RNA purification. Viral RNA (vRNA) was purified as described. (Chen et al., *J. Gen. Virol.*, 75:925-930 (1994); Yuan et al., *Virus Res.*, 79:189-200 (2001)). Briefly, supernatant from MARC-145 cells infected with VR-2332 was harvested on day 4 post-infection (p.i.). After removal of cellular debris by centrifugation at 12,000 rpm, the supernatants were layered onto a 2 ml 0.5 M sucrose cushion and centrifuged at 76,000 x g for 4 hours. The pelleted virions were resuspended in 0.5 ml LES (0.1 M LiCl/5 mM EDTA/1.0% SDS) and further digested by addition of 100 µg proteinase K at 56°C to remove all protein. After 10 minutes of incubation, vRNA was extracted several times with acid phenol and phenol/chloroform and then precipitated in 70% v/v ethanol. Pelleted vRNA was immediately resuspended into 50 µl H_20 or RNase-free TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and stored at –80°C.

Construction of full-length viral cDNA. cDNA synthesis was performed with Enhanced Avian HS RT-PCR Kit (Sigma, HSRT-100). Eight PCR primers (Table 3) were used to amplify four overlapping cDNA fragments covering the complete VR-2332 genome (Figure 2). The cycling conditions were 94°C for 2 minutes, then 35 cycles of 94°C for 15 seconds, 68°C for 4-5 seconds, followed by 68°C for 5 minutes. Each PCR fragment was purified with the QIAEX II Gel Extraction Kit (Qiagen) and cloned into pCR®2.1-TOPO® vector with TOPO TA Cloning® Kit (Invitrogen K450001). Plasmids representing each fragment were submitted for nucleotide sequence analysis. The fragments with the minimum nucleotide mutations compared to parental VR-2332 sequence (GenBank submission number U87392) were used to assemble the full-length cDNA, as shown in Figure 2. In each overlap region, a unique restriction enzyme site was utilized to join flanking fragments. Four digested fragments, representing full-length genomic sequence, were precisely assembled stepwise into a modified low copy plasmid vector (pOK12HDV-PacI). The vector was modified to include the HDV ribozyme by inserting a 244 bp *Sma*I to *Sac*II fragment

containing the HDV antigenome ribozyme and a T7 RNA polymerase terminator sequence from Transcription vector 2.0 (Johnson et al., *J. Virol.*, 71:3323-3327 (1997); Pattnaik et al., *Cell*, 69:1011-1020 (1992)) into the corresponding sites in pOK12 (Vieira et al., *Gene*, 100:189-194 (1991)). The *Nco*I restriction enzyme site in this 244 bp fragment was replaced with a unique *Pac*I site by oligonucleotide mutation with primer sets 5'pOK12HDV-2157/3'pOK12HDV-257 and 5'pOK12HDV-257/polyA-modified (Table 3), followed by fusion PCR. In the full-length cDNA clones, viral genomic sequence was preceded by the T7 RNA polymerase promoter, 1 or 2 G residues and a T residue, and followed by a polyadenylic acid tail of 50 nucleotides. Assembled clones were propagated in the DH5α strain of *Eschericia coli* and then submitted for full-genome nucleotide sequence confirmation.

Table 3. Oligonucleotide primers used in this study. Forward primers are indicated with a slash (/) after the designator, reverse primers are preceded by a slash. Inserted restriction enzyme sites are shown in underlined italics.

Primer	Genome Position*	Sequence
Cloning:		
T7Leader-VR long/	1-31	5'-ACAT*GCATGC*TTAATACGACTCACTATAGTATGACGTATAGGTGTTGGCTCTATGCCTTGG (SEQ ID NO:18)
/3'-4300	4617-4635	5'-CTGGGCGACCACAGTCCTA (SEQ ID NO:19)
5'-4056-*AscI*/	4055-4080	5'-CTTCTC*GGCGCGCC*CGAATGGGAGT (SEQ ID NO:20)
/3'-7579	7578-7603	5'-TCATCATA*CCTAGG*GCCTGCTCCACG (SEQ ID NO:21)
5'-7579/	7578-7603	5'-CGTGGAGCAGGC*CCTAGG*TATGATGA (SEQ ID NO:22)
/P32	13293-13310	5'-TGCAGGCGAACGCCTGAG (SEQ ID NO:23)
VR1509/	11938-11958	5'-GTGAGGACTGGGAGGATTACA (SEQ ID NO:24)
/3'end-FL	15405-15411	5'-GTCT*TTAATTAA*CTAG$(T)_{30}$AATTTCG (SEQ ID NO:25)
Mutagenesis:		
5'-pOK12HDV-257/(SphI, PacI)	pOK12HDV-PacI 257-282	5'-GAT*GCATGC*CA*TTAATTAA*GGGTCGGC (SEQ ID NO:26)
/3'-pOK12HDV-257(SphI, PacI)	pOK12 HDV-PacI 257-282	5'-GCCGACCC*TTAATTAA*TG*GCATGC*ATC (SEQ ID NO:27)
T7leader-VR-2G/	1-5	5'-ACATGCATGCTTAATACGACTCACTATAGGTATGAC (SEQ ID NO:28)
7475G2A/	7453-7477	5'-5Phos/CTGTGTGGACATGTCACCATTGAAA (SEQ ID NO:29)
13860C2T/	13843-13867	5'-5Phos/GTGTATCGTGCCGTTCTGTTTTGCT (SEQ ID NO:30)
14979A2G/	14958-14982	5'-5Phos/CAGATGCTGGGTAAGATCATCGCTC (SEQ ID NO:31)
Northern Blot Analyses:		
/3'-UTR	15298-15336	5'-GCACAATGTCAATCAGTGCCATTCACCACACATTCTTCC (SEQ ID NO:32)
/1a-p222	221-261	5'-TAGACTTGGCCCTCCGCCATAAACACCCTGGCATTGGGGGT (SEQ ID NO:33)

* Genome position is based on GenBank Submission U87392

Modification and sequence analysis of full-length cDNA clones. QuikChange® Multi Site-Directed Mutagenesis Kit (Stratagene) was used to modify all cDNA clones from pVR-V4 to pVR-V6G7475A. The complete genomic cDNA plasmid inserts were then submitted to the University of Minnesota Advanced Genetic Analysis Center (AGAC) for nucleotide sequence analysis with appropriate sequencing primers (Table 3). Sequence differences between pVR-V4 through pVR-V6G7475A, as well as to those of parental VR-2332, its corresponding attenuated vaccine strain, Inglevac MLV, and pVR-HN, the first infectious clone of VR-2332, are listed in Table 4 (Nelsen et al., *J. Virol.*, 73:270-80 (1999); Yuan et al., *Virus Res.*, 79:189-200 (2001); Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)).

Table 4. Nucleotide differences between PRRSV strains and VR-2332 infectious clones. Only positions where nucleotide differences were noted are shown. Nucleotides that are represented in strain VR-2332 are shown in unshaded boxes. Light shaded boxes represent nucleotide differences that are unique to the infectious clone, medium shaded boxes highlight those nucleotides that are also seen in Ingelvac® MLV, and boxes that are shaded black indicate swine unique nucleotides. Regions that were not sequenced are indicated by a slash.

Base*	Region	VR-2332	V4	V5	V5-1-P3	V5-2-P3	V5-Swine 612	V5G7475A	V5G7475A -P3	V6	V6G7475A	V6G7475A -P3	VR-HN	MLV
-4						T								
-3			G	G		T		G		G	G		G	
-2	5'UTR		G	G		T		G		G	G	T/Ø	G	
-1			T	T		T		T	T/Ø	T	T	T	T	T
48		A	A	A	A	A	A	A	A	A	A	R (G/A)	A	A
102		A	A	A	A	A	G	A	A	A	A	A	A	A
258		C	C	C	C	C	C	C	C	C	C	C	A	C
309	NSP1a	A	G	G	G	G	G	G	G	G	G	G	A	A
415		T	T	T	-	Y (C/T)	T	T	-	T	T	-	T	T
642		T	C	C	-	C	C	C	-	C	C	-	T	T
784		G	G	G	-	G	G	G	-	G	G	-	G	A
827		C	C	C	-	C	T	C	-	C	C	-	C	C
1074		C	C	C	-	C	C	C	-	C	C	-	T	C
1107	NSP1b	A	G	G	-	G	G	G	-	G	G	-	A	A
1122		A	A	A	-	R (G/A)	A	A	-	A	A	-	A	A
1181		C	C	C	-	C	C	C	-	C	C	-	C	T
1294		A	A	A	-	R (G/A)	A	A	-	A	A	-	A	A
1379		C	C	C	-	C	T	C	-	C	C	-	C	C
1595		C	A	C	-	C	C	C	-	C	C	-	C	C
2192	NSP2	C	C	C	-	C	C	C	-	C	C	-	C	T
3040		G	G	G	-	G	G	G	-	G	G	-	G	A
3457		G	G	G	-	G	G	G	-	G	G	-	G	A
3657		C	C	C	-	Y (C/T)	C	C	-	C	C	-	C	C
4407		T	C	C	-	C	C	C	-	C	C	-	T	T
4593		A	G	G	-	G	G	G	-	G	G	-	A	A
4681		T	G	G	-	G	G	G	-	G	G	-	G	G
4865		T	T	T	-	Y (C/T)	T	T	-	T	T	-	T	T
4866	NSP3	A	G	G	-	G	G	G	-	G	G	-	A	A
5097		G	A	A	-	A	A	A	-	A	A	-	A	A
5247		T	C	C	-	C	C	C	-	C	C	-	T	T
5519		C	C	C	-	C	C	C	-	C	C	-	T	C
5610		T	T	T	-	T	T	T	-	T	T	-	A	T
6345	NSP5	A	A	A	-	A	A	A	-	A	A	-	A	T

6674		**C**	T	T	-	T	T	T	-	T	T	-	C	T
6853	NSP7	**G**	G	G	-	G	G	G	-	G	G	-	A	G
6966		**T**	T	T	T	T	T	T	T	T	T	T	C	T
7183		**A**	A	A	A	A	A	A	A	A	A	R (G/A)	A	A
7188		**C**	C	C	C	C	C	C	Y (C/T)	C	C	Y (C/T)	C	C
7189		**C**	C	C	C	C	C	C	C	C	C	Y (C/T)	C	C
7213		**C**	C	C	C	C	C	C	M (A/C)	C	C	M (A/C)	C	C
7329		**G**	A	A	A	A	A	A	A	A	A	A	G	G
7475		**A**	G	G	G	G	A	A	A	G	A	A	A	A
7554		**T**	C	C	C	C	C	C	C	C	C	C	C	T
9220	NSP9	**T**	C	T	-	T	T	T	-	T	T	-	T	T
9649	NSP10	**G**	A	G	-	G	G	G	-	G	G	-	G	G
9918		**T**	T	T	-	T	T	T	-	T	T	-	T	C
9958		**G**	A	G	-	G	R (G/A)	G	-	G	G	-	A	A
10040		**A**	G	A	-	A	A	A	-	A	A	-	A	A
10533		**T**	T	T	-	T	T	T	-	T	T	-	T	C
10643		**T**	T	T	-	T	T	T	-	T	T	-	C	T
10697		**T**	C	T	-	T	T	T	-	T	T	-	C	C
10739		**C**	T	C	-	C	C	C	-	C	C	-	C	C
10781		**G**	A	G	-	G	G	G	-	G	G	-	A	A
10803		**T**	C	T	-	T	T	T	-	T	T	-	C	C
10895		**C**	C	C	-	C	C	C	-	C	C	-	C	T
11055	NSP11	**T**	A	T	-	T	T	T	-	T	T	-	A	A
11081		**G**	A	G	-	G	G	G	-	G	G	-	A	A
11221		**G**	A	G	-	G	A	G	-	G	G	-	A	A
11229		**G**	G	G	-	G	G	G	-	G	G	-	G	T
11259		**C**	G	C	-	C	C	C	-	C	C	-	C	C
11327		**C**	T	C	-	C	C	C	-	C	C	-	C	C
11329		**G**	C	C	-	C	C	C	-	C	C	-	C	C
11501		**A**	G	A	-	A	A	A	-	A	A	-	A	A
11666	NSP12	**C**	T	C	-	C	C	C	-	C	C	-	T	T
11744		**G**	C	G	-	G	G	G	-	G	G	-	G	G
11760		**G**	G	G	-	R (A/G)	G	G	-	G	G	-	G	G
11882		**A**	G	A	-	A	A	A	-	A	A	-	A	A
12076	ORF2a/b	**A**	G	A	-	A	A	A	-	A	A	-	A	A
12102		**G**	G	G	-	G	G	G	-	G	G	-	G	T
12153		**A**	G	A	-	A	A	A	-	A	A	-	A	A
12432		**A**	G	A	-	A	A	A	-	A	A	-	A	A
12501		**T**	A	T	-	T	T	T	-	T	T	-	T	T
12600		**G**	G	G	-	G	G	G	-	G	G	-	G	T
12943	ORF3	**G**	G	G	-	G	G	G	-	G	G	-	G	A
12950		**C**	C	C	-	C	C	C	-	C	C	-	C	T
12973		**T**	G	T	-	T	T	T	-	T	T	-	T	T
13011		**G**	G	G	-	G	G	G	-	G	G	-	A	A
13787	No ORF	**T**	T	T	-	T	T	T	-	T	T	-	C	T

13825	ORF5	G	G	G	-	G	G	G	-	G	G	-	G	A
13860		T	T	C	-	C	C	C	-	T	T	-	T	T
14238		A	A	A	-	A	A	A	-	A	A	-	A	G
14336		T	T	C	-	C	Y (T/C)	C	-	C	C	-	T	T
14404	ORF6	T	T	C	-	C	C	C	-	C	C	-	T	T
14420		C	C	C	-	C	C	C	-	C	C	-	C	C
14686		A	A	A	-	A	G	A	-	A	A	-	A	A
14735		C	G	G	-	G	G	G	-	G	G	-	G	G
14737		G	C	C	-	C	C	C	-	C	C	-	C	C
14979	ORF7	G	A	A	-	A	A	A	-	G	G	-	G	G
15281	3'UTR	A	G	A	-	A	A	A	-	A	A	-	A	A
15334		T	C	T	-	T	T	T	-	T	T	-	T	T
15339		C	C	C	-	Y (T/C)	C	C	-	C	C	-	C	C
15411		T	T	T	-	K (T/G)	Y (T/C)	T	-	T	T	-	T	T

* The negative bases refer to those nucleotides present in the RNA after transcription and derived from the RNA polymerase promoter immediately upstream of the infectious polynucleotide. These promoter-derived nucleotides are typically no longer present in an infectious polynucleotide after it has been passaged 9 times.

In vitro transcription. The full-length cDNA clone was linearized by cleavage with *Pac*I, which cuts downstream of the poly(A) tail. Capped [m^7G(5')ppp(5')G cap analog] RNA transcripts were produced using the mMESSAGE MACHINE™ Kit (Ambion) and an optimized 2:1 ratio of methylated cap analogue to GTP. Approximately 50 to 60 μg of RNA was generated from 2 μg of DNA template in a 20-μl of reaction mixture. Increasing the ratio of cap analogue to GTP substantially reduced the RNA yield. The RNA was subsequently purified by acid phenol-chloroform followed by isopropanol precipitation and resuspended in nuclease-free TE buffer (pH 8.0). RNA was evaluated for quality by size comparison with wild-type VR-2332 viral RNA on a 1% glyoxal denaturing agarose gel, and quantified by spectrophotometry at OD_{260}.

MARC-145 cell transfection. A modified transfection procedure was generated based on the approached described by Nielsen (Nielsen et al., (*J. Virol.*, 77:3702-3711 (2003)). For transfection, MARC-145 cells were seeded onto six-well plates (2-3 x 10^5 cells/well) in 3 ml of complete medium [EMEM supplemented with 10% fetal bovine serum (FBS)] and then incubated at 37°C, 5% CO_2 for 20-24 hours until approximately 80% confluent (Collins et al., *J. Vet. Diagn. Invest.*, **4**:117-126 (1992)). 4 μg of *in vitro* transcribed RNA diluted in 500 μl Opti-MEM® I Reduced Serum Medium (Invitrogen) and 2 μl of 1,2-dimyristyloxypropyl-3-dimethyl-hydroxy ethyl ammonium bromide and cholesterol (DMRIE-C; Invitrogen) diluted in 1 ml Opti-MEM® medium were combined and vortexed briefly. The MARC-145 cells were washed once with 2 ml Opti-MEM® medium and then immediately overlayed with the lipid:RNA complex solution. DMRIE-C without RNA (2 μl) was used as a negative control and DMRIE-C with 10 - 100 ng strain (wild type) *wt* VR-2332 purified viral RNA was used as a positive control. After 4 hours of exposure to the lipid:RNA complexes, the monolayers were washed and fresh complete medium (EMEM with 10% FBS) was added. Supernatants from transfected cells were monitored daily for appearance of cytopathic effect (CPE) and passaged onto fresh MARC-145 at 72-96 hours posttransfection.

Detection of progeny viral RNA. To detect progeny viral RNA, cell culture supernatant from transfected and infected MARC-145 cells were harvested. RNA was isolated with QiaAmp viral RNA Kit (Qiagen). RT-PCR was performed with select primer pairs, specific to the VR-2332 strain nucleotides that were indicative of

infectious clone mutated residues (Table 3). Confirmation of infectious clone progeny was obtained by nucleotide sequence verification of clone specific nucleotides present in the RT-PCR products.

Detection of progeny viral nucleocapsid protein. Indirect immunofluorescence assays (IFA) were used to detect viral protein expression in *in vitro* transcript RNA transfected, or progeny virus infected, MARC-145 cells prepared on coverslips. Infected cells were fixed in 3.7% paraformaldehyde with phosphate buffered saline (PBS), pH 7.5, at room temperature for 10 minutes. The fixed cells were washed with PBS, incubated at 37°C for 45 minutes in PRRSV nucleocapsid protein specific monoclonal antibody SDOW17 (Magar et al., *Can. J. Vet Res.*, 59:232-234 (1995)) and further incubated with goat anti-mouse immunoglobulin G (IgG) conjugated with fluorescein isothiocyanate at 37°C for another 45 minutes (1:100 dilution) (Sigma). The coverslips were washed with PBS, mounted to a slide using gel mount oil, and observed under a fluorescence microscope.

Viral plaque assay. MARC-145 cell monolayers on six-well plates were infected with cell supernatant (in 10-fold dilutions) from transfected or infected MARC-145 cells by incubation at room temperature for 1 hour. Infected monolayers were subsequently washed once with fresh EMEM/10% FBS, overlaid immediately with sterile 1% SeaPlaque Agarose (BioWhittaker Molecular Applications, Rockland, Maine) in 1X MEM (Sigma M4144)/10% FBS/2% (w/v) $NaHCO_3$/1X glutamine/1X nonessential amino acids/10 mM HEPES/2% (v/v) gentamycin, and incubated at 37°C/5% CO_2, inverted, for 5 days. After careful removal of the agarose, cells were stained with 5% crystal violet in 20% ethanol for 10-30 minutes for visualization of plaque size.

Viral growth curve. MARC-145 monolayers in T-75 flasks were inoculated with either parental or recombinant PRRSV diluted in serum-free EMEM at a multiplicity of infection (MOI) of 0.001. After 1 hour attachment at room temperature with gentle mixing, the inocula were removed and the monolayers washed three times with serum-free EMEM. After washing, 4 ml complete medium was added and the flasks were subsequently incubated for up to 5 days at 37°C, 5% CO_2. Aliquots (0.5 ml) were harvested immediately after the addition of medium (0 hour time point) and at 24, 48, 72, 96 and 120 hours and stored at -80°C. Serial dilutions of the samples were used to infect fresh MARC-145 cells and the cells then processed as described above.

After removal of the agarose, plaques were visualized and counted. Growth curve results were expressed as PFU/ml.

In vivo inoculation of progeny virus. Ten 4-week-old pigs of mixed breed and sex from a PRRSV-seronegative herd were divided into three groups, each consisting of two animals. The first group received $10^{3.5}$ 50% tissue culture infectious dose ($TCID_{50}$) of cloned virus (pVR-V5, third passage on MARC-145 cells) per ml, the second group received $10^{5.4}$ $TCID_{50}$ per ml of the parental virus strain VR-2332 (fourth passage on MARC-145 cells), and the third group was mock inoculated with EMEM. All of the animals received 2 ml of inoculum by intramuscular injection. The animals were kept in separate rooms throughout the experiment and observed daily for clinical signs. All pigs were euthanized on day 28 postinfection. To recover virus, individual serum samples were diluted 5-fold with incomplete EMEM and placed on fresh MARC-145 monolayers for 1 to 2 hours at room temperature with gentle agitation. The inocula were then removed and complete EMEM was added. Infected cells were incubated at 37°C, 5% CO_2 and observed daily. Once CPE was evident, infected cell supernatants were frozen at -80°C until further characterized.

Northern Blot Analysis. pVR-V6G7475A transcripts were transfected into MA104 cells and then passaged onto fresh cells for several passages. For subsequent northern blot analysis, supernatants from passage 1 (P1), P3, P6, P8 and P10 were diluted 1:50 and then used to infect cells (1 ml/T75 flask) on the same day. At the same time, infected swine serum was diluted 10-fold and then used (1 ml) to infect a separate T75 flask. Cytopathic effect was seen on day 3 p.i. for all flasks. Intracellular RNA was extracted using a RNeasy Midi kit (Qiagen) and electrophoresced (15 μg/sample) on a glyoxal denaturing gel as described previously (Nelsen et al., *J. Virol.*, 73:270-80 (1999)). pVR-V6G7475A transcript RNA (100 ng) was run as a control. After RNA transfer to 0.45 micron MagnaGraph Nylon Transfer Membrane (Osmonics), the membrane was probed with labeled oligonucleotide /1a-p222, end labeled with γ-^{32}P-ATP (Amersham) using polynucleotide kinase (Promega) as described previously (Nelsen et al., *J. Virol.*, 73:270-80 (1999)).

Nucleic acid sequence analysis of progeny virus. 5'- and 3'- rapid amplification of cDNA ends (RACE) was performed with SMART™ RACE cDNA Amplification Kit (BD Bioscience) or 5' or 3'-Full Race Core Set (TaKaRa Bio Inc) on viral RNA isolated with the QIAmp®Viral RNA Mini Kit (Qiagen). The remaining nucleotide

sequence was determined from RT-PCR products of primer pairs developed to cover the entire genome of strain VR-2332 (Table 3), as described previously (Yuan et al., *Virus Res.*, 79:189-200 (2001)). The products were submitted for nucleic acid sequence determination at the Advanced Genetic Analysis Center at the University of Minnesota. Complete viral sequence with at least three fold coverage was initially assembled with the SeqMan suite of the Lasergene® sequence analysis software (DNASTAR, Inc.), and further analyzed using GCG Wisconsin Package Version 10.3 software (Accelrys Inc.). Strain VR-2332 (GenBank Accession U87392) strain Ingelvac® MLV (GenBank Accession AF066183) and cDNA clone pVR-HN (GenBank Accession AY150564; Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)) were used in all nucleotide comparisons to recombinant virus strains.

Results

Modification of pOK12 Vector. pOK12 (GenBank Accession AF223639; Vieira et al., *Gene*, 100:189-194 (1991)), a low copy cloning vector, was modified by digestion with SmaI (enzyme site at 273 bp in pOK12) and SalI (site at 307 bp) and inserting the 244 bp SmaI-SalI fragment of Vector 2.0 (7) containing the hepatitis delta virus (HDV) ribozyme. The vector (pOK12HDV) was then further modified by mutagenesis of an existing KpnI site (pOK12HDV site at 273 bp) to insert a PacI restriction enzyme site through the use of the primer pair 5'-pOK12HDV-257SphIPacI/3'-pOK12HDV-257SphIPacI. The HDV ribozyme was added to provide for effective cleavage precisely at the 3'end of the polyA tract. Studies revealed that the modification was not necessary to obtaining infectious progeny virus.

Construction of full-length cDNA clones. The cloning strategy is depicted in Figure 2. Four overlapping genome fragments were amplified from purified VR-2332 viral RNA by RT-PCR using the primer pairs indicated (Figure 2, Table 3). Each fragment was individually cloned into the pCR®2.1-TOPO® vector to generate intermediate clone pCR-SphI-FseI (segment I), pCR-FseI-AvrII (segment II), pCR-AvrII-BsrGI (segment III), and pCR-BsrGI-PacI (segment IV). The cDNA clones were then digested with two unique restriction enzymes, as indicated by the clone name. Four fragments were gel-purified and stepwise ligated to vector pOK12HDV-PacI to generate a full-length cDNA clone of PRRSV (pVR-V4). In the full-length cDNA clone, viral genomic sequence was driven by T7 RNA polymerase promoter and followed by polyadenylic acid tail of 50 nucleotides. RNA transcripts of clone pVR-V4

did not display typical PRRSV infectivity when transfected into permissive cells, although viral RNA could be detected over several passages. When compared to strain VR-2332, a total of 45 nucleotide mutations (Table 4) leading to 21 amino acid changes were detected (Table 5), although several mutations were the same as previously identified in Ingelvac® MLV (Yuan et al., *Virus Res.*, 61:87-98 (1999)).

Table 5. Amino acid differences between PRRSV strains and VR-2332 infectious clones. Only positions where nucleotide differences were noted are shown with corresponding amino acid position within the identified genomic region. Amino acids that are represented in strain VR-2332 are shown in unshaded boxes and infectious clone amino acid identities with VR-2332 are represented by blank boxes. Text in each individual box represent silent or amino acid changes due to nucleotide differences shown in Table 2. Light shaded boxes represent nucleotide differences that are unique to the infectious clone, medium shaded boxes highlight those nucleotides that are also seen in Ingelvac® MLV, and boxes that are shaded black indicate swine unique nucleotides. Amino acids separated by slashes indicate ORF2a/ORF2b amino acid numbers. Regions that were not sequenced are indicated by a slash.

NT Position	AA Position	Region	VR-2332	V4	V5	V5-1-P3	V5-2-P3	V5-Sw612	V5G7475A	V6	V6G7475A	VR-HN	MLV
258	23		V									Silent	
309	40	NSP1α	Q	Silent	Silent	Silent	Silent	Silent	Silent	Silent	Silent		
642	151		P	Silent	Silent	-	Silent	Silent	Silent	Silent	Silent		
784	199		V			-							L
827	213		A			-		V					
1074	295	NSP1β	Y			-						Silent	
1107	306		L	Silent	Silent	-	Silent	Silent	Silent	Silent	Silent		
1181	331		S			-							P
1379	397		A			-		V					
1595	469		A	D		-							
2192	668	NSP2	S			-							P
3040	951		D			-							N
3457	1090		D			-							N
4407	1406		P	Silent	Silent	-	Silent	Silent	Silent	Silent	Silent		
4593	1468		Q	Silent	Silent	-	Silent	Silent	Silent	Silent	Silent		
4681	1498		S	A	A	-	A	A	A	A	A	A	A
4866	1559		V	Silent	Silent	-	Silent	Silent	Silent	Silent	Silent		
5097	1636	NSP3	R	Silent	Silent	-	Silent	Silent	Silent	Silent	Silent	Silent	Silent
5247	1686		V	Silent	Silent	-	Silent	Silent	Silent	Silent	Silent		
5519	1777		T			-						I	
5610	1807		L			-						Silent	
6345	2052	NSP5	P			-							Silent

6674	2162		P	L	L	-	L	L	L	L	L		L
6853	2222		D			-						N	
6966	2259		D			-						Silent	
7329	2380	NSP7	K	Silent	Silent	Silent	Silent	Silent	Silent	Silent	Silent		
7475	2429		E	G	G	G	G			G			
7554	2455		V	Silent	Silent	Silent	Silent	Silent	Silent	Silent	Silent	Silent	
9220	3011	NSP9	L	P		-							
9649	3154		G	E		-							
9918	3244		L			-							Silent
9958	3257		G	E		-		G/E				E	E
10040	3284		V	Silent		-							
10533	3449		Y			-							H
10643	3485	NSP10	V			-						Silent	
10697	3503		A	Silent		-						Silent	Silent
10739	3517		H	Silent		-							
10781	3531		T	Silent		-						Silent	Silent
10803	3539		C	R		-						Silent	R
10895	3569		D			-							Silent
11055	3623		S	T		-						T	T
11081	3631		P	Silent		-						Silent	Silent
11221	3678		G	E		-		E				E	E
11229	3681	NSP11	V			-							L
11259	3691		R	G		-							
11327	3738		H	Silent		-							
11329	3739		G	A	A	-	A	A	A	A	A	A	A
11501	3771		E	Silent		-							
11666	3826		P	Silent		-						Silent	Silent
11744	3852	NSP12	W	C		-							
11882	3898		K	Silent		-							
12076	2		K	E		-							
12102	10/9		L/D			-							L/Y
12153	27	ORF2a/b	P/I	P/V		-							
12432	120		E	Silent		-							
12501	143		D	E		-							
12600	176		G			-							Silent
12943	83	ORF3	G			-							E
12950	85		D			-							Silent

12973	93		M	R		-					S	S	
13011	106		G			-						Q	
13825	13	ORF5	R			-							
13860	25		F		L	-	L	L	L			G	
14238	151		R			-							
14336	183		G		Silent	-	Silent	Silent	Silent	Silent	Silent		
14404	10		H		Silent	-	Silent	Silent	Silent	Silent	Silent		
14420	16		Q			-						E	
14686	104	ORF6	L			-		Silent					
14735	121		R	G	G	-	G	G	G	G	G	G	
14737	121		R	G	G	-	G	G	G	G	G	G	
14979	31	ORF7	A	T	T	-	T	T	T				

Because many mutations in pVR-V4 occurred in the critical region encoding putative helicase, polymerase and other Nidovirus motifs (Figure 3, Table 4), additional clones of genomic segment III (pCR-AvrII-BsrGI) were generated and sequenced in their entirety. After replacing segment III of pVR-V4 with the most sequence accurate fragment obtained, we again determined the nucleotide sequence of the entire genomic full-length clone (pVR-V5). Except for the replaced region and for four spontaneous mutations (nucleotides 1595, 13860, 14336, and 14404), these two genomic clones were identical (Table 4). Sequence analysis of pVR-V5 showed that this clone harbored a total of 23 mutations compared to strain VR-2332. Of these 23 changes, only 8 nucleotide mutations coded for a change in amino acid and five of the amino acid residue mutations were identical to Ingelvac® MLV and thus not predicted to adversely effect *in vitro* replication (Table 4).

Clone pVR-V6 was derived from site-directed mutagenesis of genome segment IV to repair nucleotides 13860 and 14979 using primers 13860C2T/ and 14979A2G/, respectively. Mutation of these two nucleotides would correct amino acid residue 25 of GP5 (L→F) and residue 31 of the nucleocapsid protein (T→A). Sequence analysis of clone pVR-V6 confirmed that the nucleotides had been corrected back to wild-type (*wt*) VR-2332 nucleotides and had not resulted in any other nucleotide changes elsewhere in the genome when compared to pVR-V5 (Tables 4 and 5). Finally, site-directed mutagenesis on genome segment III using oligomer 7475G2A was completed on both pVR-V5 and pVR-V6 in order to correct an alteration from *wt* VR-2332 at nt 7475. The change of G→A at nt 7475 resulted in a glycine (G) at ORF1 amino acid 2429 in the two recombinant clones to the glutamic acid (E) seen in the parental VR-2332 viral strain. The final two clones, pVR-V5G7475A and pVR-V6G7475A were again sequenced in their entirety and found to have only (nt 7475) altered from the original recombinant plasmids pVR-V5 and pVR-V6, respectively (Table 5). pVR-V6G7475A thus contains 11 nucleotide and no amino acid changes from strain VR-2332, besides those also seen in Ingelvac® MLV.

As can be seen schematically in Figure 3 for the final construct (pVR-V6G7475A), and detailed in Tables 4 and 5, all full-length clones still possess nucleotide changes scattered throughout the genome, primarily in the poorly defined regions of ORF1. However, the large cluster of ORF1b nucleotide changes that presumably prevented pVR-V4 from completing viral replication were repaired in later

versions of the full-length genome clones. Only one nucleotide mutation (nt 11329 coding for G3739A mutation) remained in ORF1b of pVR-V5 and later clones, and this mutation does not prevent Ingelvac® MLV from infecting and replicating efficiently in cultured cells. Tables 4 and 5 also display the residue information for the previously published infectious clone, pVR-HN (Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)), shown to replicate in animals. There is a substantial increase in the number of residues in pVR-HN (15 nucleotides) that directly display the sequence of Ingelvac® MLV over the final construct, pVR-V6G7475A (7 nucleotides).

Characterization of recombinant virus. Full-length RNA transcripts of each cDNA clone were produced. MARC-145 cell transfection with the cDNA transcripts or *wt* VR-2332 viral RNA (vRNA) resulted in CPE, characterized by cell clumping followed by lysis, at 48 to 72 hours post transfection. CPE caused by the recombinant transcripts were delayed and somewhat distinct compared to that induced by *wt* VR-2332 vRNA in which CPE presents as vigorous aggregation, detachment, and disruption. At 96 hours posttransfection, most of the cells transfected with VR-2332 vRNA had undergone lysis and detached from the plate, whereas less severe CPE was apparent in cells transfected with the cloned *in vitro* derived RNA transcripts.

Virus (P0) was harvested from the transfected cells and an aliquot (10 µl diluted to 1 ml in culture medium) was used to infect MARC-145 cells for progeny virus amplification. After CPE was detected, virus (P1) was again harvested and an aliquot used for reinfection of MARC-145 cells. Recombinant virus in the cell supernatant (P2) was utilized for purification of viral RNA, which was then used to obtain RT-PCR fragments with primer pairs 5'-6800/3'-ORF1b (nt 6796-7614) and P51/05P4 (nt 13757-14341). The PCR fragments obtained were submitted for nucleotide sequence analysis to confirm that the infectivity seen was due to transfected full-length RNA transcripts of the infectious construct and not a result of contamination due to *wt* virus. Nucleotide mutations at residues 7329, 7475, 7554, and 13860 nucleotide differences were seen in progeny virus from pVR-V5, and 7329, 7554, and 13860 were detected in virus from pVR-V5G7475A. Similarly, mutations at residues 7329, 7475, and 7554 were detected in pVR-V6 progeny and mutations at 7329 and 7554 were detected in virus resulting from pVR-V6G7475A (Tables 4 and 5). Corresponding mutations were not seen in P2 virus from *wt* vRNA transfections.

Immunofluorescence analysis of recombinant viruses. Direct immunofluorescence assays were used to detect the expression of PRRSV nucleocapsid protein in infected MARC-145 cells. All cells infected by recombinant virus transcripts (P2 and on) as well as vRNA were positive by this method. Massive nucleolar accumulation of the nucleocapsid protein was readily apparent, as previously reported by Rowland et al. (*Virus Res.*, 64:1-12 (1999)).

In vivo infection with pVR-V5 derived recombinant virus. Recombinant viruses recovered from P3 of MARC-145 cells transfected with RNA transcripts of cDNA clone pVR-V5 were inoculated into young swine in parallel with *wt* VR-2332, vaccine virus Ingelvac® MLV and saline (negative control). Blood samples were collected on 0, 3, 5, 7, 14, 21 and 28 days p.i. and analyzed for seroconversion by HerdChek PRRS 2XR ELISA (IDEXX) and for virus recovery. At day 28, all infected animals had seroconverted with approximately the same kinetics, revealing that pVR-V5 recombinant viruses replicated well *in vivo* (Figure 4). Clinical signs were absent from all animals during the course of the experiment, but this was not unexpected as *wt* strain VR-2332 often does not produce overt disease in young swine and results in enlarged lymph nodes only transiently, typically at day 14 p.i.

A serum sample from one animal infected with progeny of pVR-V5 (Sw612), taken at 14 days p.i., was used to infect fresh MARC-145 monolayers for recovery of *in vivo* passaged recombinant virus. As described previously, the virus derived from *in vitro* transfection of clone pVR-V5 RNA transcripts caused only minimal CPE (evidenced by aggregation of infected cells) while virus recovered from day 14 serum of the test animal caused typical CPE (cell aggregation, detachment, and disruption) at 96 hours postinfection. This suggested that a shift in viral genotype or phenotype had occurred while pVR-V5 replicated *in vivo*.

In order to elucidate the reason for the apparent change in phenotype, full-genome sequence analysis was completed on virus recovered from one pig (Sw612) and then passaged once in MARC-145 cells to amplify the Sw612 progeny (Figure 3, Tables 4 and 5). When compared to the virus used to infect swine, pVR-V5, 17 infectious cDNA clone-specific nucleotide changes were retained in Sw612, some of which are also seen in Ingelvac® MLV (7/17 nucleotides). The two non-viral G residues followed by a T residue present at the 5' end of the original pVR-V5 clone transcript were not seen in the virus derived from *in vivo* infection. Degeneracy was

seen at nucleotide positions 9958 (R), 14336 (Y) and 15411 (Y). The *wt* VR2332-like nucleotide (G) at position 9958 showed degeneracy with an Ingelvac® MLV-like nucleotide (A). This change results in a mutation of a glycine residue to a glutamic acid residue, respectively (Table 2). At position 14336, degeneracy was detected as an infectious clone-specific base (C) and a *wt* VR-2332-specific base (T), which reflected a silent mutation. Another mutation (nt 7475) occurred in which a G residue had reverted to the *wt* residue A. However, there were another 5 nucleotide differences (nt 102, 827, 1379, 14686 and 15411) not seen in any of the other viruses in this study. Nucleotide 102 is located in the leader sequence, thought not be translated. However, if the leader sequence were translated, the encoded ORF (VR-2332 nucleotides 1-100) would be extended by one amino acid residue (W). The mutations at residues 827 and 1379 led to mutations in ORF1a, in both cases resulting in an amino acid change of *wt* VR-2332 encoded alanine for a Sw612 valine. The guanine residue at nt 7475 of pVR-V5 had mutated to *wt* adenine. This resulted in a G3294A non-conservative amino acid mutation, which lies in ORF1a predicted protease cleavage product NSP7 and this genomic region has no defined function to date. Nucleotide 14686, located in ORF6, showed a change from a *wt* VR-2332 guanine to an alanine in Sw612, which still encodes the amino acid glycine. The other unique nucleotide change occurred at the very 3' end of the viral sequence (nt 15411), before the start of the polyA tail. In this case, a previously conserved thymine residue revealed degeneracy with a cytosine residue. These genetic changes, although informative, did not immediately reveal the cause(s) of the change in growth phenotype observed. However, it did reveal the errant nature of PRRSV replication *in vivo* and suggests that a moderately different viral genomic sequence from *wt* VR-2332 was able to replicate efficiently (Figure 3).

Comparison of viral plaque size. Plaque size determinations of the recombinant viruses as well as *wt* VR-2332 were completed in parallel on MARC-145 cells at 120 hours p.i. (Fig. 5A). Strain VR-2332 formed plaques that averaged 3 mm in size, while passage 3 progeny of pVR-HN cDNA clone formed slightly smaller plaques (2.5 mm average). In contrast, only pinpoint plaques were obtained from recombinant viruses derived from pVR-V5 and pVR-V6, and these were only readily apparent through microscopic examination (Fig. 5A). Recombinant virus recovered from clones pVR-V5G7475A and pVR-V6G7475A formed, on average, 1.5 mm and 2mm plaques respectively. However, in another assay, the plaques produced by the viral progeny

(Sw612) recovered from *in vivo* infection of VR-FLV5 derived recombinant virus were much larger, approximately equal in both size and number as those derived from *wt* VR2332 (Fig 5B).

Only minimal volumes of the cell supernatants containing each recombinant virus remained. Therefore, in order to fully examine the role of nucleotide change in determining plaque size, we transfected fresh RNA transcripts produced from pVR-V5, pVR-V6, pVR-V5G7475A and pVR-V6G7475A into MARC-145 cells (termed second lineage). Passage 3 progeny viruses of each infectious clone at 5 days post-infection were again analyzed for plaque size in comparison to *wt* VR-2332, VR-HN and Sw612 viruses. In contrast to the previous plaque assay, all plaque sizes appeared similar, with the recombinant viruses obtained from pVR-V5, pVR-V6, pVR-V5G7475A only slightly smaller than the *in vivo* derived *wt* VR-2332, Sw612 and pVR-V6G7475A viruses (Fig. 6A). The recombinant viruses, however, were not yet directly mimicking authentic viral infection as shown by the approximately 10-fold lower titers when compared to *wt* VR-2332 or to pVR-V5 recombinant virus that had been passaged through swine (Sw612)(Fig. 6B).

Nucleotide Sequence Analysis of First and Second Lineage Virus Preparations. Limited nucleotide sequence analysis (due to virus stock limitation) of passage 3 pVR-V5-derived virus inoculated into swine (V5-1-P3) and complete nucleotide sequence analysis of passage 3 pVR-V5-derived virus obtained above (V5-2-P3) were completed in order to reveal the genetic reason for the plaque size discrepancies. Such analyses revealed that the two independently prepared V5 viruses differed in sequence at the 5' end (Table 4). The virus that had produced pinpoint plaques (V5-1-P3) had no extraneous 5'-end nucleotides, as shown in the nucleotide sequence of *wt* strain VR-2332, while that producing larger plaques (V5-2-P3) possessed 4 non-templated thymidine residues at the 5' terminus (Table 4). The remaining V5-1-P3 viral nucleotide sequence we could obtain exactly matched that of V5-2-P3 virus, as well as that of the parental clone. However, complete sequence analysis of V5-2-P3 virus revealed that the virus displayed nucleotide degeneracy at several genomic sites. Similar findings were obtained when analyzing limited regions of second lineage viruses VR-FLV5G7475A-P3 and VR-FLV6G7475A-P3. These last two infectious clone progeny displayed different 5'-termini as well as exhibiting degeneracy in sequence.

Viral Growth Curves. Simultaneous one-step viral growth curve determinations were completed using MARC-145 cells and passage 3 viruses (second lineage) (Fig. 7). The recombinant viruses recovered from pVR-V5, pVR-V5G7475A, pVR-V6, and pVR-V6G7475A and pVR-HN displayed similar one-step viral growth rates, but their peaks of replication were all significantly lower than *wt* strain VR-2332 and Sw612, the *in vivo* progeny of pVR-V5. Also, the replication rates of the recombinant virus preparations derived from pVR-V5, pVR-V6 and pVR-HN were somewhat decreased as compared to the virus derived from pVR-V5G7475A and pVR-V6G7475A. The last two infectious clones code for as little as 13 and 11 nucleotide differences, respectively, resulting in 2 and zero amino acid changes, from *wt* VR-2332 sequence besides the changes seen in Ingelvac® MLV. These data then reveal that viruses with as little as 11 nucleotide changes from *wt* VR-2332 and its attenuated offspring Ingelvac® MLV are somehow impaired in replication. Correspondingly, the resultant titers of *wt* VR-2332 and Sw612 viruses were approximately 6-15 fold higher than that of the recombinant viruses that had not been passaged in swine (Fig. 7).

Northern analyses of vRNA. PRRSV defective sgRNA species, identified previously as heteroclite subgenomic RNAs (latin: uncommon forms), have been shown to be a constituent of PRRSV infection and cannot be separated from full-length viral genomes by standard methods such as cultured cell passage at low multiplicities of infection or sucrose gradient centrifugation (Yuan et al., *Virology*, 275:158-169; 30 (2000); Yuan et al., *Virus Res.*, 105:75-87 (2004)). To explore whether or not PRRSV heteroclites are produced during *in vitro* transcription of full-length cDNA genome clones or appear after subsequent transfection/infection, northern blot analysis was completed. The full-length RNA transcript and passages 1, 3, 6, 8 and 10 of the virus produced from transfected MA-104 cells were used to inoculate fresh T-75 flasks of MA-104 cells with 10 µl supernatant diluted 1:100, as well as Sw612 serum diluted 1:10 (2 ml total/flask). After 4 days, intracellular PRRSV RNA was harvested and 15µg of each preparation was separated by electrophoresis through a denaturing agarose gel and transferred to a nylon membrane. After RNA crosslinking, the membrane was hybridized with a ^{32}P-radiolabeled probe complementary to the 5'end of ORF1a that selects for full-length VR-2332 genomes as well as heteroclites (/1a-222; 29). As shown in Figure 8, the RNA transcript is mostly a single band, migrating as full-length vRNA, while PRRSV RNA species from passage 1 and later migrate as both

full-length and subgenomic-sized species previously identified as heteroclites. In addition, the strength of hybridization increases over passage. Since the virus was harvested from an equal volume of infected cell supernatant at the same time point, this observation suggests that the vRNA becomes more efficient at replication over time. Lastly, when comparing virus generated from Sw612 with the cell culture generated virus, the RNA banding pattern is indistinguishable, strongly suggesting that the defective RNA species are readily formed and replicated *in vitro* as well as *in vivo* and thus are a natural part of PRRSV infection.

Discussion

In theory, an infectious cDNA clone of a virus should be identical to the parental sequence in order to generate a reverse genetic system that mimics wild-type infection. Considerable effort was exerted to reproduce a fully faithful PRRSV strain VR-2332 genome, yet due to unpredictable spontaneous mutations at several sites, we have not yet been successful at deriving an infectious clone that has no differences from the *wt* strain VR-2332 sequenced in our laboratory. High fidelity DNA polymerases, used in this study, are available to decrease artificial mutations, but such mutation cannot be avoided during reverse transcription (Malet et al., *J. Virol. Methods*, 109:161-70 (2003)). In addition, the fact that PRRSV exhibits astonishing viral evolution and strain variation (Chang et al., *J. Virol.*, 76:4750-6 (2002); Murtaugh et al., *Adv. Exp. Med. Biol.*, 440:787-94 (1998); Yoon et al., *Adv. Exp. Med. Biol.*, 494:25-30 (2001)) recombines readily at high frequency to result in intergenic recombinants between strains (Yuan et al., *Virus Res.*, 61:87-98 (1999)), undergoes intragenic recombination to form PRRSV subgenomic RNAs and heteroclites (Nelsen et al., *J. Virol.*, 73:270-80 (1999); Yuan et al., *Virology*, 275:158-169 (2000); Yuan et al., *Virus Research*, 105:75-87 (2004)) and often displays nucleotide degeneracy at unpredictable nucleotide sites in field isolates serve to make this initial goal time-consuming and of negligible gain. An infectious DNA construct possessing as little as 11 nucleotide mutations, as compared to strain VR-2332, outside of domains known to be involved in viral replication (5' and 3' ends, ORF1b) was thought sufficient for *wt* virus production and the downstream goals of infectious clone use for pathogenesis queries and structure:function studies. pVR-HN is more similar to Ingelvac® MLV in the region of the virus encoding the helicase motif (NSP 10). Further pathogenic comparison of these two infectious clones

may shed light on the differences between the parental strain, VR-2332, and its vaccine strain offspring, Ingelvac® MLV.

Valuable information can be derived from the construction and evaluation of the infectious clones for PRRSV strain VR-2332. First of all, PRRSV strain VR-2332 cannot tolerate all mutations for survival. Particular nucleotide or amino acid mutations may help or hinder viral replication, and the challenge is to ascertain which are lethal to survival. In clone pVR-V4, which did not produce infectious virions, there were total of forty-two nucleotide differences from *wt* parental strain VR-2332. In these forty-two nucleotide changes, several nucleotides result in silent mutations (20 residues) or exist in other known PRRSV strains (9 amino acid residue mutations directly mimic Ingelvac® MLV) allowed prediction that these changes may be non-lethal for virus replication. Eleven nucleotide changes leading to 12 amino acid changes and two 3'UTR nucleotide mutations, each not seen in Ingelvac® MLV, were thus predicted to be lethal to PRRSV strain VR-2332. In pVR-V5 and later constructs, 19 changes were corrected, including several silent mutations and 9 aberrant amino acid changes not seen in the genome of Ingelvac® MLV and 8 other changes seen in the vaccine strain. This lead to the first evidence that the constructs were infectious, although in pVR-V5 two amino acid mutations were still present, one of which was altered through site directed mutagenesis to produce pVR-V6. The remaining amino acid change was repaired in pVR-V5G7475A and pVR-V6G7475A, although these clones still harbor silent mutations that are not found in strain VR-2332 and the derived vaccine strain.

Several unique observations were obtained from this study. First of all, each lineage of produced virus may result in a unique 5' terminal sequence that was not detected in *wt* strain VR-2332. We also cannot yet correlate plaque size with nucleotide sequence. Secondly, we saw unique nucleotide changes after replication in swine, which may reflect the inherent nature of the PRRSV polymerase. All nucleotide changes were transitional in nature and did not exhibit a bias (5 A/G and 4 C/T). Although the G A reversion at nucleotide 7475 was seen after *in vivo* passage, we could not correlate this site with the subsequent increased plaque size because other non-templated changes had occurred. In addition, full-genome sequence analyses of passage 3 of a V5-derived virus that produced larger plaques (V5-2-P3) revealed a different 5'terminal sequence from the pinpoint plaque-producing V5 virus used to infect swine (V5-1-P3). However, we can conclude that the mutations were not lethal

to virus replication because this virus, after passage in swine, produced *wt*-sized plaques on MARC-145 cells ad grew at almost the same rate as the parental virus (Fig 5A, 6 and 7).

Of considerable interest is the fact that sequence analysis of the third *in vitro* passage of V5, V5G7475A and V6G7475A seemed to suggest that the PRRSV replicase complex allows frequent transitions, and infrequent transversions, to occur while undergoing viral replication. This may reflect a viral replicase that has evolved so that it may generate new genetic forms of a PRRSV genome and then assess their competence amid other variants, resulting in an optimally "fit" virus. These observations have also been noted during PRRSV sequential passage *in vivo* (Chang et al., *J. Virol.*, 76:4750-63 (2002)). Present sequencing efforts are to examine the full-length genomes of later passages, when a more robust replication is detected. Finally, it is now clear that PRRSV strain VR-2332 replicase readily synthesizes heteroclites at the same time it is producing full-length vRNA. This prototype strain, isolated and characterized in 1992, may be unique in the gradual acquisition of replication fitness, as other investigators producing infectious clones of more recent strain have not observed the same effect (Truong et al., *Virology*, 325:308-319 (2004)). The role of heteroclite formation and the concomitant appearance of vigorous viral replication suggest that there is an advantageous role for heteroclites in PRRSV evolution.

Example 2

Many virulent isolates of a seemingly novel PRRSV were recently identified in the State of Minnesota, USA. ORF5 nucleotide sequence analysis and comparison to the University of Minnesota Veterinary Diagnostic Laboratory PRRSV database (>5000 isolates) revealed that the isolates were of Type 2 lineage, but were significantly different than previous isolates. Furthermore, they were most closely related to those isolates previously seen in Canada in the early 1990s (Mardassi et al., *J. Gen. Virol.*, 75:681-685 (1994)) and in the State of Minnesota in 1998. Restriction fragment length polymorphism (RFLP) analysis of ORF5 also demonstrated that they belonged to the same group of viruses as these early cases, known as 1-8-4 isolates (Wesley et al., *J. Vet. Diagn. Invest.*, 10:140-144 (1998)) and were thus named MN184 isolates. Because of the striking dissimilarity with all but one previously isolated MN PRRSV isolate, two of these new isolates were amplified just one time on porcine alveolar macrophages (PAM), the host cell, and full-length genome analyses was completed on the viruses, designated as MN184A and MN184B. These two isolates were collected at different times from two separate farms.

Materials and Methods

To sequence the MN184 isolates, viral RNA (vRNA) was extracted from PRRSV infected cell supernatant with QIAmp® Viral RNA Mini Kit (Qiagen, Valencia, CA)) and RT-PCR was performed (Qiagen® OneStep RT-PCR Kit). Primers (available on request) were designed based on the published sequences of different strains of PRRSV deposited in GenBank as well as newly generated MN184 sequence. The 5′ nucleotide sequence of the two PRRSV isolates was derived using the 5′-Full RACE Core Kit (TaKaRa Bio, Madison, WI). 3′-RACE was performed with SMART™ RACE cDNA Amplification Kit (Clontech, Mountain View, CA). RT-PCR products were gel purified (QIAquick®, Qiagen), cloned into the pGEM-T Vector (Promega, Madison, WI) and 3 to 5 clones for each RT-PCR product were chosen for sequencing. The nucleotide sequence determination was completed in both directions with the PCR specific primers or the vector encoded SP6 and T7 promoter primers. The products were submitted to the Advanced Genetic Analysis Center at the University of Minnesota for sequence determination with an ABI 377 automated DNA fragment analyzer. A quality sequence representing at least three-fold genome coverage was

obtained. Sequence data was assembled and analyzed by using the GeneTool sequence analysis program (BioTools Inc., Edmonton, Alberta CA) and Lasergene (DNASTAR, Madison, Wis.).

Multiple sequence alignments were generated with CLUSTALX (Thompson et al., *Nucleic Acids Res.*, 24:4876-4882 (1997)) or Wisconsin Package Version 10.3 (Accelrys Inc., San Diego, CA). Full-length PRRSV sequences were aligned using ClustalX (version 1.83.1; IUB DNA weight matrix, gap penalty 15.00, gap length penalty 6.66). The resulting alignment was further analyzed using the Wisconsin Package Version 10.3 Distances Program (Jukes-Cantor distance method, partial matches due to degenerate symbols considered). For Figure 10, sequences were aligned with the Pileup program of the Wisconsin Package (Blosum62 Scoring Matrix, Gap Weight = 8, Length Weight = 2, Weighted Ends). The alignment was scored for redundancy and colored for percent identity using Jalview (Clamp et al., *Bioinformatics*, 12:426-427 (2004)) and then transferred to Adobe® Photoshop® CS, version 8.0, for grayscale transformation. For Figure 11, sequences were aligned with the Pileup program of the Wisconsin Package (Blosum62 Scoring Matrix, Gap Weight = 8, Length Weight = 2, Weighted Ends). For Figure 12, a signal peptide was predicted using the SignalP server (Bendtsen et al., *J. Mol. Biol.*, 340:783-795 (2004)). Transmembrane regions were derived by PHDhtm (Rost et al., *Protein Sci.*, 5:1704-1718 (1996)) and potential N-glycosylation sites were identified by PROSITE (Bairoch et al., *Nucleic Acids Res.*, 25:217-221 (1997)) using the PredictProtein server (Rost et al., *Nucleic Acids Res.*, 32:W321-W326 (2003)). Sequences were aligned with the Pileup program of the Wisconsin Package (Blosum62 Scoring Matrix, Gap Weight = 8, Length Weight = 2, Weighted Ends).

Results

Genomic alignment demonstrated that these two PRRSV were quite distinct (> 14.5% nucleotide dissimilarity) from other North American Type 2 full-length sequenced genomes, yet comparison with Type 1 (European) full-length sequences confirmed that the isolates were solely of Type 2 genotype origin as they were only approximately 59% similar at the nucleotide level to both EuroPRRSV and Lelystad strains. Strikingly, these Type 2 MN184 isolates represented the shortest PRRSV genomes detected to date (15019 nucleotides, not including the poly A tail). In

addition, no specific area was discerned that suggested that these isolates were derived from viral recombination between Type 1 and Type 2 strains.

Full-length sequence analysis revealed that the two MN184 isolates were actually genetically distinct. They shared 98.0% nucleotide similarity or 2% difference. This percentage of dissimilarity was unexpected due to their sudden simultaneous appearance in Minnesota, with no clear recent related isolate seen in our PRRSV database at that time. Table 6 presents the detailed nucleotide and amino acid comparison between the two isolates and Figure 9 depicts the amino acid differences seen between these two strains. Both of these isolates possessed nucleotide degeneracy in several regions of the genome, predominantly in the predicted nsp2 region of ORF1 (Table 6). The fact that nucleotide degeneracy was seen in these isolates suggested that PRRSV can be made up of several individual species, often referred to as a swarm of related but distinct viral sequences, within infected animals.

Table 6. Detailed analysis of individual PRRSV genomic regions and translated proteins, and number of degenerate bases detected in each region. Degeneracy is defined as more than one nucleotide detected for a particular base on separate trace files of three or more trace files.

Region	Bases	Nucleotide length	% Nucleotide Similarity	% Nucleotide Identity	Number of Degenerate Bases (184A/184B)	Amino Acid Length	%Amino Acid Similarity	%Amino Acid Identity
5' UTR	1-190	190	99.5	98.9	1/0	-	-	-
ORF1A	191-7309	7119	98.5	96.7	16/109	2372	96.8	96.5
NSP1a	191-688	498	98.8	98.5	1/0	166	97.6	97.6
NSP1b	689-1339	651	98.3	97.5	2/3	217	97.2	95.9
NSP2	1340-3886	2547	98.0	94.6	10/76	849	94.2	94.2
NSP3	3887-5224	1338	98.7	98.7	0/0	446	99.3	98.9
NSP4	5225-5836	612	98.5	96.4	0/13	204	97.1	97.1
NSP5	5837-6346	510	99.2	95.3	3/17	170	97.1	97.1
NSP6	6347-6394	48	100.0	100.0	0/0	16	100	100
NSP7	6395-7171	777	99.3	99.3	0/0	259	99.6	99.2
NSP8	7172-7309	138	99.3	99.3	0/0	46	97.6	97.6
ORF1B	7306-11679	4374	99.2	98.9	5/4	1457	99.5	99.2
NSP9	7288-9225	1938	98.9	98.8	1/1	646	99.4	98.9
NSP10	9226-10548	1323	99.3	98.9	3/3	441	99.8	99.3
NSP11	10549-11217	669	99.3	99.3	0/0	223	99.5	99.5
NSP12	11218-11679	462	99.6	99.4	1/0	153	99.3	99.3
ORF2a/GP2	11681-12451	771	99.0	98.3	1/0	222	98.0	97.3
ORF2b/E	11686-11907	222	99.6	99.6	0/0	73	100	100
ORF3/GP3	12304-13068	765	98.6	98.6	0/0	254	97.6	97.6
ORF4/GP4	12849-13385	537	98.5	98.5	0/0	178	98.9	98.9
ORF5/GP5	13396-13998	603	97.8	97.7	1/0	200	96.5	96.5
ORF6/M	13983-14507	525	99.6	97.4	0/0	174	100	100
ORF7/N	14497-14868	372	98.9	98.9	0/0	123	97.6	97.6
3' UTR	14869-15019	151	100	98.0	1/1	-	-	-

In order to more closely pinpoint the individual regions of these MN184 isolates that showed the most dissimilarity from other PRRSV strains and to assign the region(s) accounting for the difference in Type 2 viral genome length, these two isolates were compared to the sequence of the prototype Type 2 strain VR-2332. The differences between the two isolates could again be discerned, with isolate MN184B possessing slightly increased similarity to strain VR-2332 than isolate MN184A. The nucleotide and amino acid comparisons to VR-2332 showed individual MN184 isolate regions varied from 81.5-94.7% and 78.4-100%, respectively, but the regions corresponding to ORF5 (86.4-86.7% and 87.0-87.5%, respectively) predicted nsp1β (83.8-84.0% and 84.8-85.4%, respectively, and nsp2 (81.5-85.5% and 78.4-79.5%, respectively) were the most variable. Most interesting was that only the predicted nsp2 genomic region showed a difference in nucleotide length and that both MN184 isolates possessed the same nsp2 deletion, detailed below. The comparison also revealed that the 5′ and 3′ UTR's were the most conserved regions of the genome (94.7% and 94.0%, respectively), indicating sequence conservation in important regions for viral replication and transcription.

ORF5 encodes a heterogeneous PRRSV structural protein (GP5) and is often used for PRRSV diagnostic identification (Kapur et al., *J. Gen. Virol.*, 77:1271-1276 (1996)). GP5 is a predicted three transmembrane protein with an endodomain and ectodomain. The 30 amino acid ectodomain is composed of a short highly conserved domain usually containing at least two N- glycosylation sites bounded by two hypervariable regions. The highly conserved domain of this 30 amino acid region has been shown to code for the viral attachment epitope in Type 2 strains (Plagemann, *Virology*, 290:11-20 (2001); Ostrowski et al., *J. Virol.*, 76:4241-4250 (2002); Plagemann et al., *Arch. Virol.*, 147:2327-2347 (2002)). GP5 of the same set of full-length genomes, as well as the original RFLP184 isolates identified in Canada (IAF-93-653, IAF-Klop) and in 1998-1999 in Minnesota (98-3298, 98-3403, 99-3584) were aligned (Fig. 10). The alignment of PRRSV GP5 revealed amino acid identities ranging from 82.5% to 87.7% between the new MN184 isolates and other non-RFLP184 Type 2 strains. Interestingly, the amino acid differences between the new MN184 isolates and the older RFLP184 isolates were quite large (5.7% - 12.2%) and thus we detected no clear origin of the new RFLP184 virus. The limited alignment shows that most of the amino acid differences observed were found in the hypervariable

regions (Fig. 10). The two conserved N-glycosylation sites were maintained in the MN184 isolates, except for detected nucleotide degeneracy coding for amino acid 44 in isolate MN184B.

Nsp1β encodes a papain-like cysteine protease (den Boon et al., *J. Virol.*, 69:4500-4505 (1995)). An amino acid alignment of the MN184 isolates with a non-redundant set of available Type 2 nsp1β sequences as well as Type 1 strains EuroPRRSV and Lelystad was completed (Fig. 11). The nsp1β protein possesses a number of completely conserved amino acids, and the proposed catalytic residues were maintained in all sequenced genomes (den Boon et al., *J. Virol.*, 69:4500-4505 (1995)). The alignment, ordered by amino acid similarity, indicates that the MN184 isolates are more similar to Type 1 strains than the other sequenced full-length Type 2 sequences. In particular, five amino acids (boxed in Fig. 11) directly mimic the Type 1 strains. However, the amino acids that were conserved in the other non-redundant Type 2 sequences were also mostly conserved in the MN184 isolates, but scattered amino acids and the amino acid similarity (84.8-85.4%) revealed a more divergent Type 2 protein than had been evidenced to date. Thus, the alignment further defines maintained residues of nsp1β that may be critical to the replication cycle of PRRSV.

An amino acid alignment of non-redundant sequences of nsp2, ordered by pairwise identity, is shown in Figure 12. A highly conserved chymotrypsin-like cysteine protease (PL2) domain is present at the N-terminus, previously predicted by alignment with equine arteritis virus (EAV) nsp2 (Snijder et al., *J. Gen . Virol.*, 79:961-979 (1998); Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)). There are 3-4 predicted transmembrane domains near the C terminus of this protein (McGuffin et al., *Bioinformatics*, 16:404-405 (2000)), but the exact C terminal cleavage site has not been empirically determined. Two predictions of the C-terminal cleavage site have been proposed, one G|G at VR-2332 nsp2 amino acid 980 (Allende et al., *J. Gen. Virol.*, 80:307-315 (1999)) and the other at amino acid 1197 (Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)), but there are several completely conserved G|G doublets within this protein (VR-23332 nsp2 amino acids 646, 980, 1116, 1196, 1197; downward arrows in Fig. 12). Prior work had also shown that the predicted nsp2 protein is proline rich and contains multiple potential B-cell epitopes (Oleksiewicz et al., *J. Virol.*, 75:3277-3290 (2001); Fang et al., *Virus Res.*, 100:229-235 (2004); Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). The large middle region of PRRSV nsp2 (VR-2332

nsp2 amino acids 148-880) has no assigned function but is highly variable in length. Furthermore, the length difference between sequenced Type 1 and Type 2 strains of PRRSV has been mapped to this variable middle region of nsp2 (Fig. 12). Until now, sequenced Type 1 genomes have been shown to be 313-364 bases shorter than most Type 2 PRRSV (Meulenberg et al., *Virology*, 192:62-72 (1993); Fang et al., *Virus Res.*, 100:229-235 (2004), Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). However, the multiple sequence alignment established that the MN184 genome contains the shortest predicted nsp2 to date (2547bp), 393bp shorter than prototype Type 2 strain VR-2332. Furthermore, it contained three discontinuous deletions in the translated protein with deletion sizes consisting of 111, 1 and 19 amino acids, respectively, corresponding to the amino acid positions in PRRSV strain VR-2332 nsp2 of 324-434, 486 and 505-523, respectively (Fig. 12). The three deletions resulted in the loss of several proline residues and predicted B-cell epitopes. Besides these deletions, significant alterations in nsp2 amino acid sequence from other Type 2 strains were also seen, sometimes corresponding to the Type 1 amino acid seen at the same relative position (Fig. 12). Comparison of the nsp2 predicted protein of the two PRRSV genotypes demonstrated that the amino acid identity within Type 2 viruses ranged from 66% to 99% and from 88-90% within Type 1 viruses, but differed greatly between genotypes (<45% similarity). In particular, the MN184 isolates displayed 66-80% amino acid identity to all Type 2 nsp2 predicted proteins and only 43-45 % identity to Type 1 strains. When surveying the multiple sequence alignment in Fig. 12, we also noted that all instances of insertion or deletion in both genotypes occurred in this hypervariable middle region. To this point, Shen et al. (*Arch. Virol.*, 145:871-883 (2000)) first reported that PRRSV North American Type 2 strain SP has a unique insertion of 36 aa relative to the position between aa 813 and 814 of PRRSV VR-2332 nsp2. Another investigator found a unique 12 aa deletion at position 466-477 in PRRSV isolate HB-2(sh)/2002 nsp2 (Gao et al., *Arch. Virol.*, 149:1341-1351 (2004)). A 17 aa deletion occurred in newly identified European-like PRRSV isolates when compared to strain LV (Fang et al., *Virus Res.*, 100;229-235 (2004); Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). The instances of mutation did not consistently occur along the same stretch of amino acids, although the deletions seen between the MN184 isolates and other Type 2 viruses encompass most of the largest deletion detected between Type 1 and other Type 2 PRRSV. All of these data suggested that the nsp2 ORF contains a conserved protease

motif and predicted transmembrane spanning regions that may be necessary for replication of PRRSV, but is highly susceptible to mutation in the large middle section.

The sudden appearance of field isolates of PRRSV in Minnesota reflecting the 184 RFLP pattern is still a mystery, but the consequences of this event are even now being realized. The Minnesota Veterinary Diagnostic Laboratory now performs routine sequencing on similar 184 RFLP isolates from approximately one fourth of the total number of ORF5 sequence requests. In addition, the 184 RFLP pattern has now been detected not only in Minnesota, but in Iowa, Wisconsin, South Dakota, Kansas, Missouri, Illinois, Nebraska, Kentucky, Oklahoma and Wyoming as well. We chose to derive the full-length sequences from two isolates because of the need to understand if this could be more than a single virus type and the fact that the swine herd diagnosed with isolate MN184A presented with a milder case of PRRS than the herd infected with isolate MN184B, as reported by the attending pathologist. The strains have not been inoculated into naïve animals to verify the case presentations, but it is interesting to note that isolate MN184B had many more nucleotide degeneracies detected when analyzing the genome and this might reflect the severity of the disease reported.

This genome analysis increased our understanding of the immense nucleotide and amino acid sequence variation that exists in the field. Factors driving this variation may be related to the way swine are now managed, the interstate and international transport of swine and boar semen, the intermixing of different PRRSV isolates within herds and the nature of the virus itself. Full genome sequence generation also allows us to monitor where on the genome variation is tolerated and which regions are more conserved. As a result of this study, as well as a previous publication (Ropp et al., *J. Virol.*, 78:3684-3703 (2004)), a picture is emerging that indicates nsp2, nsp1β and ORF5 are extraordinarily versatile proteins.

This study has also provided clear evidence that nsp2 size can no longer be used to differentiate between the two PRRSV genotypes. The novel finding that nsp2 evolved to display a Type 2 genome with three discontinuous deletions, leading to the shortest genome to date (15,019 kb), suggests that PRRSV may be evolving to eliminate dispensable genomic regions and make the genome more compact. Finally, although the significance of genetic variations in PRRSV can only be surmised at present, the evolutionary change seen in ORF5, nsp1β and nsp2 should reasonably be related to the biological fitness of PRRSV during selection pressure.

The complete disclosure of all patents, patent applications, and publications, and electronically available material (including, for instance, nucleotide sequence submissions in, e.g., GenBank and RefSeq, and amino acid sequence submissions in, e.g., SwissProt, PIR, PRF, PDB, and translations from annotated coding regions in GenBank and RefSeq) cited herein are incorporated by reference. The foregoing detailed description and examples have been given for clarity of understanding only. No unnecessary limitations are to be understood therefrom. The invention is not limited to the exact details shown and described, for variations obvious to one skilled in the art will be included within the invention defined by the claims.

Unless otherwise indicated, all numbers expressing quantities of components, molecular weights, and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about." Accordingly, unless otherwise indicated to the contrary, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties sought to be obtained by the present invention. At the very least, and not as an attempt to limit the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits and by applying ordinary rounding techniques.

Notwithstanding that the numerical ranges and parameters setting forth the broad scope of the invention are approximations, the numerical values set forth in the specific examples are reported as precisely as possible. All numerical values, however, inherently contain a range necessarily resulting from the standard deviation found in their respective testing measurements.

All headings are for the convenience of the reader and should not be used to limit the meaning of the text that follows the heading, unless so specified.

Claims

1. An infectious polynucleotide comprising a nucleotide sequence having at least 88% identity to SEQ ID NO:1 and a deletion of at least 57 consecutive nucleotides corresponding to nucleotide 2062 to nucleotide 3864 of SEQ ID NO:1.

2. The polynucleotide according to claim 1, wherein the polynucleotide replicates and produces infectious virus particles when introduced into a cell.

3. The polynucleotide according to claim 1, further comprising a vector sequence that replicates in a prokaryotic host cell.

4. An infectious polynucleotide comprising a nucleotide sequence having at least 88% identity to SEQ ID NO:14 and a deletion of at least 57 consecutive nucleotides corresponding to nucleotide 2061 to nucleotide 3545 of SEQ ID NO:14.

5. The polynucleotide according to claim 4, wherein the polynucleotide replicates and produces infectious virus particles when introduced into a cell.

6. The polynucleotide according to claim 4, further comprising a vector sequence that replicates in a prokaryotic host cell.

7. The polynucleotide of claims 1, 2, 4, or 5 wherein the polynucleotide is present in a vector.

8. The polynucleotide of claims 1, 2, 4, or 5 or the infectious clone of claims 3 or 6 wherein the polynucleotide comprises 2 or more deletions, and wherein each deletion is independently at least 57 consecutive nucleotides.

9. The polynucleotide of claims 1, 2, 4, or 5 wherein the polynucleotide is present in an isolated virus particle.

10. The polynucleotide of claims 1, 2, 4, or 5 wherein the polynucleotide is an RNA polynucleotide.

11. The polynucleotide of claims 1, 2, 4, or 5 or the infectious clone of claims 3 or 6 wherein the polynucleotide is present in a cell.

12. The polynucleotide of claims 1,2,4, or 5 or the infectious clone of claims 3 or 6 wherein an RNA polymerase promoter is operably linked to the polynucleotide.

13. The polynucleotide of claims 1, 2, 4, or 5 or the infectious clone of claims 3 or 6 wherein the polynucleotide further comprises an exogenous polynucleotide present in the deletion.

14. The polynucleotide or the infectious clone of claim 13 wherein the exogenous polynucleotide encodes a detectable marker.

15. An infectious polynucleotide comprising the nucleotide sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13.

16. An nsp2 polypeptide encoded by an infectious polynucleotide comprising the nucleotide sequence of SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO: 10, SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13.

17. An isolated infectious polynucleotide having the sequence of SEQ ID NO:1.