Review

PRRSV structure, replication and recombination: Origin of phenotype and genotype diversity

Matthew A. Kappes, Kay S. Faaberg*

Virus and Prion Research Unit, USDA-ARS-National Animal Disease Center, Ames, IA, USA

Article info

Article history:
Received 15 December 2014
Returned to author for revisions
23 January 2015
Accepted 9 February 2015
Available online 7 March 2015

Keywords:
Porcine reproductive and respiratory syndrome virus
Arterivirus
Replication
Recombination
Heteroclite
Subgenomic RNA

Abstract

Porcine reproductive and respiratory disease virus (PRRSV) has the intrinsic ability to adapt and evolve. After 25 years of study, this persistent pathogen has continued to frustrate efforts to eliminate infection of herds through vaccination or other elimination strategies. The purpose of this review is to summarize the research on the virion structure, replication and recombination properties of PRRSV that have led to the extraordinary phenotype and genotype diversity that exists worldwide.

Published by Elsevier Inc.

Introduction

Porcine reproductive and respiratory syndrome virus (PRRSV) is the etiological agent of a worldwide epidemic designated porcine reproductive and respiratory syndrome (PRRS). PRRSV is highly host and tissue restricted to swine cells of the monocyte lineage, preferentially infecting the porcine alveolar macrophage (AMΦ) and is a persistent virus (Duan et al., 1997; Villarreal et al., 2000). PRRSV first emerged in the late 1980s as a “mystery” disease progressing through swine populations in both Europe and North America (Baron et al., 1992; Collins et al., 1992; Hopper et al., 1992; Morin et al., 1991; Wensvoort et al., 1991). Prevailing clinical symptoms were noted to be respiratory distress in young swine and widespread reproductive failure in pregnant sows including mummified, stillborn and aborted fetuses (Goyal, 1993). Initial characterization of circulating European (Type 1) and North American (Type 2) genotype isolates was found to be surprisingly genetically divergent. Although overall disease phenotype, gross
clinical symptoms, genomic organization and temporal emergence were all similar, these strains differed by ~40% at the nucleotide level (Benfield et al., 1992; Collins et al., 1992; Meulenberg et al., 1994; Nelsen et al., 1999; Wensvoort et al., 1991). The degree of genetic heterogeneity suggests a protracted period of independent evolution on the two continents (Nelsen et al., 1999). Molecular clock analysis predicts the divergence of the two genotypes from a common ancestor between a decade to over a century prior to clinical recognition (Forsberg et al., 2001; Yoon et al., 2013), presumably from another host species (Hanada et al., 2005; Plagemann, 2003). The origin of PRRSV remains unknown and no secondary transmission has been reported to date (Otak et al., 2003a, 2003b; Schurrer et al., 2005, 2004; Zimmerman et al., 1997). In the ~25 years since the first emergence of PRRSV, a global near worldwide epidemic has been sustained by a set of emerging and re-emerging strains supported by high frequency mutation and recombination (Goldberg et al., 2003; Meng, 2012; Murtaugh et al., 2010). PRRSV remains the most economically devastating disease of swine and contributes to the deterioration of animal health through disease and the continual emergence of increasingly divergent and often virulent strains (Gauger et al., 2012; Han et al., 2006; Holtkamp, 2011; Tian et al., 2007). Other reviews have recently assessed the Nidovirus family in broad terms, concentrating on what is similar between the family members and comparisons to other nidoviruses such as some coronaviruses, which are considered to be the closest relatives (Moennig, 1992). The similar genomic organization, characteristic genetic elements and common functionality of orthologous proteins however have led to the acceptance of many putative functions for PRRSV proteins, often derived from studies of the arterivirus prototype species EAV and the more distantly related nidoviruses. Each arterivirus infects only one animal species, in contrast to other nidoviruses such as some coronaviruses, which have been shown to transmit between species (Hilgenfeld and Peiris, 2013).

The PRRSV genome varies from 14.9 kb to 15.5 kb in length and expresses a range of accessory and structural proteins through two distinct transcription mechanisms. The genomic organization and associated expression profiles are depicted in Fig. 1. The PRRSV genome encodes a 5’ proximal noncoding element (5’-untranslated region; 5’ UTR) of 217–222 nucleotides (nt; Type 1) and 188–191 nt (Type 2) in length (Faaberg et al., 2012; Yun and Lee, 2013). The 5’ UTR functions will be discussed below.

Directly downstream of the 5’ UTR has the large overlapping replication open reading frames (ORF). The ORF1a/b share a single translational start site but are augmented by two ribosomal frameshift (RFS) sites at genomic positions 3889 nt (RFS1; nonstructural protein (nsp) 2) and 7695 nt (RFS2; nsp8/9) [VR-2332 (U87392) reference sequence] (Fang et al., 2012; Meulenberg et al., 1993b; Nelsen et al., 1999; Li et al., 2014). Two products are generated from the RFS1 site: a -1 RFS occurs approximately 7% of the time and results in an immediate termination of translation (nsp2N) (Li et al., 2014), and a -2 RFS event occurs at ~20% efficiency and yields a translational extension in the -2 coding frame through the putative transmembrane domain of nonstructural protein (nsp2) to generate polyproteins.

**Genome**

PRRSV is a member of the Arteriviridae family within the order Nidovirales, which also includes the Coronaviridae and Roniviridae families. The nidovirus order constitutes a group of single-stranded positive-sense RNA viruses which share a hallmark replication/transcription strategy, similar genomic organization, and a defining set of genetic elements, but differ in host species and range, disease phenotype, virion morphology, cellular tropism, genomic size and encoded content (Baker, 2012). The Arteriviridae family is composed of five viruses which share similar genetic and biological characteristics such as genomic organization and content, morphology and a cellular tropism for the macrophage lineage (Faaberg et al., 2012). Viral species include PRRSV, simian hemorrhagic fever virus (SHFV), lactate-dehydrogenase elevating virus (LDV), newly recognized wobbly possum disease virus (WPDV) in free-ranging Australian brushtail possums (Trichosurus vulpecula) and equine arteritis virus (EAV; arterivirus prototype species) (Baker, 2012; Dunowska et al., 2012; Plagemann and Moennig, 1992). The similar genomic organization, characteristic genetic elements and common functionality of orthologous proteins however have led to the acceptance of many putative functions for PRRSV proteins, often derived from studies of the arterivirus prototype species EAV and the more distantly related nidoviruses. Each arterivirus infects only one animal species, in contrast to other nidoviruses such as some coronaviruses, which have been shown to transmit between species (Hilgenfeld and Peiris, 2013).

**Fig. 1.** PRRSV genome, transcription, and translation. PRRSV replication progresses by a range of genetic and protein regulatory mechanisms. Expression of the first three-fourths of the 14.9–15.5 kb genome yields 4 known polyproteins (pp1a, pp1a-nsp2N, pp1a-nsp2TF, pp1ab) through two documented programmed RFS (●). Polyproteins are co-translationally and post-translationally processed into at least 16 distinct replicase (accessory) nonstructural proteins (nsp5–16) by four viral encoded proteases PLP1α, PLP1β, PLP2, and SP. Recognized polymerase motifs in pp1b are the RNA dependent RNA polymerase (RdRp), the zinc-finger domain (Z), the helical domain (HEL) and the nidovirus uridylate-specific endoribonuclease (NendoU; U) domain. Canonical structural proteins are expressed exclusively through a set of subgenomic RNAs (sgRNA; 2–7) via a coterminal discontinuous transcription strategy via a negative-sense strand intermediate.
the product nsp2TF (Fang et al., 2012). The large replicase polyproteins pp1a, pp1a-nsp2N, pp1a-nsp2TF, and pp1ab are generated from full-length genomic RNA. Using the Type 2 prototype strain VR-2332 (U87392) for a reference, ORF1a/β is responsible for the cleavage of the RFSs and four distinct polyproteins including papain-like cysteine proteinases 1 (PLP1; nsp1α), PLP1β (nsp1β), and PLP2 (nsp2), and the main serine proteinase (SP; nsp4) (Faaberg et al., 2012; Snijder et al., 2013). PLP1α and PLP2 function to cleave the nsp1α-nsp1β and the nsp1β-nsp2 junction, respectively; PLP2 is responsible for the cleavage of the nsp2-nsp3 junction and the main SP processes all remaining nsp products (nsp3–12) (Han et al., 2009; Snijder et al., 2013). ORF1a encodes pp1a encompassing nsp1–8 and ORF1ab encodes pp1ab composed of all known nsps (nsp1αβ, nsp2–6, nsp7αβ, nsp8–12) whereas nsp9 is a translational extension of nsp8 via a programmed -1 RFS at position VR-2332 7695 nt (RFS2) (Fig. 1) (Meulenberg et al., 1993b; Nelsen et al., 1999).

In contrast, the previously recognized structural proteins are encoded and individually expressed by a set of subgenomic RNAs (sgRNA) generated through a negative-strand intermediate (sgRNA2–7; Fig. 1) (van Marle et al., 1999a). sgRNAs are genetically polycistronic (except RNA7) but are assumed to be functionally monocistronic/bicistronic, where only the 5′ terminal ORF(s) is expressed (Fig. 1). sgRNA 2 encodes ORF2a/b which is translated to yield glycoprotein 2 (GP2) and a small uncleaved envelope protein (E); ORF3 is expressed from sgRNA3 to yield GP3; and sgRNA4 encodes ORF4 yielding GP4. Together GP2, GP3, and GP4 form a trimeric complex resulting in the minor glycoprotein complex which functions in viral entry and is heavily N-glycosylated (Das et al., 2010; Wissink et al., 2005). sgRNA5 encodes ORF5 and ORF5a. ORF5a codes for the ORF5a protein, a small uncleaved glycoprotein that is required for virus viability and ORF5 codes for GP5, the major glycoprotein with a variable number of N-glycan residues surrounding the cell attachment domain (Johnson et al., 2011; Mardassi et al., 1996; Meulenberg et al., 1995; Robinson et al., 2013). ORF6 is expressed from sgRNA6, resulting in the generation of the membrane protein (M). GP5 and M form a disulfide-linked heterodimer and together constitute the major glycoprotein complex on the virion, as was first shown for LDV (Faaberg et al., 1995; Mardassi et al., 1996). Finally, the nucleocapsid protein (N) is encoded by ORF7 and is expressed from sgRNA7. N is the major structural element within the PRRSV virion which forms disulfide-linked homodimers, functions to package the viral genomic RNA (gRNA), and is the only known structural protein which does not encode a transmembrane domain or to not have an ectodomain upon the PRRSV virion (Bautista et al., 1996; Dea et al., 2000; Doan and Dokland, 2003; Loemba et al., 1996; Spilman et al., 2009; Wissink et al., 2005; Wootten and Yoo, 2003). Recently, the nsp2 protein, coded for by the most variable region of the genome with insertions and deletions, was also shown to be incorporated into or onto ultra-purified virions of several PRRSV strains as a set of differently sized protein isomers, presumably through its four to five membrane spanning regions near the C-terminal end (Kappes et al., 2015) (Han et al., 2010; Kappes et al., 2013). This surprising result increases the number of viral proteins, 10 or more (full-length nsp2 and its isomers, nsp2TF, GPs 2–5, E, M, N, ORF5a), that are exposed to the porcine immune system on entry of PRRSV into swine alveolar macrophages (Fang et al., 2012; Veit et al., 2014). The original work suggests that only 3 of these (GP5, M, and N) make up the majority of the protein content of PRRSV (Drew et al., 1995; Mardassi et al., 1994a, 1995; Meulenberg et al., 1995; Nelson et al., 1994). However, the immense genetic and protein variation of all of these structural proteins, from the least conserved nsp2 region (Han et al., 2006; Tian et al., 2007) to the most conserved M protein (Murtaugh et al., 1995; Veit et al., 2014), shows the complexity and the plasticity of the PRRSV genome and virion structure.

The 5′ and 3′ UTRs flank the core protein coding regions of the PRRSV genome (Fig. 1). Both the 5′ and 3′ UTRs are implicated as essential components contributing to the viral strategies imparting replicative and translational functionality; however, the exact functions of the 5′ and 3′ UTRs, and the associated mechanisms of interaction, are poorly understood. Both encode conserved putative RNA secondary structures important to replicative function. The 5′ UTR is encoded first within the PRRSV genome and possesses a putative type 1 5′ Cap structure (Sagripari et al., 1986). The 5′ UTR is genetically variable, Type 1 and 2 strains share approximately 50% genetic homology, and within each genotype, the pairwise identity is about 96% (Lin et al., 2002; Meulenberg et al., 1993a; Nelsen et al., 1999; Oleksiewicz et al., 1999; Tan et al., 2001). Detailed studies of the distantly related coronavirus species, as well as arterviruses, have shown the 5′UTRs are regulators of genomic replication, transcription, and mRNA translation, and are considered a necessary docking site for a variety of viral and host factors to complete these functions (Choi et al., 2006; Gao et al., 2012; Liao and Lai, 1994; Lu et al., 2011; Tahara et al., 1994; Zhang et al., 1994). The 3′ UTR is located directly downstream of ORF7 and is encoded by approximately 150 nt excluding the polyadenylation site. The 3′ UTR [114 nt (Type 1), 148 nt (Type 2)] is also generally diverse, sharing approximately 70% nucleotide identity between Type 1 and Type 2 sequences that is about 96% pairwise nucleotide identity within each genotype (Beereks and Snijder, 2007; Choi et al., 2006; Verheije et al., 2002; Yin et al., 2013). Recent reviews extend upon the brief description presented here (Snijder et al., 2013; Wang et al., 2014; Yun and Lee, 2013).

Replication

Due to the unique attributes of nidovirus transcription and replication, including uncharacteristically large polycistronic RNA genomes and the transcription of a nested set of 5′, 3′ co-terminal sgRNAs through a discontinuous transcription strategy, which is in itself a mechanism of recombination, nidoviral RNA synthesis mechanisms have been suggested to be of unparalleled complexity among positive strand RNA viruses (Gorbaleyena et al., 2006; van Hemert et al., 2008). PRRSV replication closely ties three key features: rearrangement of host membranes to establish viral replication complexes (RC), synthesis and expression of gRNA, transcription of sgRNA for the efficient expression of structural proteins, at the same time as the unique ability to produce aberrant PRRSV sgRNAs known as heteroclitics (Yuan et al., 2000, 2004). Genesis of gRNA (replication) and sgRNA is inherently tied to the shared negative strand synthesis mechanism. Modulation of negative-sense transcription through a non-stochastic mechanism yields either gRNA or one of six standard sgRNAs (RNA2–7) through an abortive disjoining/rejoining discontinuous transcription strategy (Fig. 1) (Meng et al., 1996b; Nelsen et al., 1999; Pasternak et al., 2001; van Marle et al., 1999a).

The characterization of the cellular entry mechanism that PRRSV utilizes has been studied in detail, and will not be covered in this review (Van Breedam et al., 2010). Little is known about the establishment of PRRSV infection, however, from the point post-entry to the development of RC, including the formation of
characteristic perinuclear double-membrane vesicles (DMVs) (Knoops et al., 2012). DMVs are believed to be derived from the endoplasmic reticulum (ER) which are apparent sites of viral replication (Pedersen et al., 1999). It has been shown that the EAV replicase proteins (ORF1a/b) are sufficient to support viral replication (Molenkamp et al., 2000b), but that infectivity is dependent on the presence of the structural genes encoded at replication (Molenkamp et al., 2000b; Verheije et al., 2002). Upon entry, the gRNA serves as the mRNA for immediate translation of the large replicase polyproteins. Within ORF1a, three proteins are recognized putative transmembrane proteins (nsp2, nsp3, and nsp5). The EAV nsp2 and nsp3 were shown to be sufficient to modulate host cellular membranes into structures similar to those observed during viral infection (Snijder et al., 2001). It is believed that membrane integration and possibly protein–protein interactions of these transmembrane proteins function to torque the existing membrane structures to form the DMVs; tethering the genesis and processing of the polyprotein (s) at the site of replication. Additional viral or cellular interacting partners are not well defined. The mechanism of DMV formation is unknown but may include the modulation of autophagy and/or apoptosis pathways (Brekenridge et al., 2003; Chen et al., 2012; Costers et al., 2008; Cottam et al., 2014; Huo et al., 2013; Labarque et al., 2003; Razi et al., 2009; Sun et al., 2012; Wang et al., 2014; Yin et al., 2012; Yu et al., 1999).

The core replicative machinery of PRRSV – the RNA dependent RNA polymerase (RdRp; nsp9), the zinc-binding domain (ZBD; or Z; nsp10), the RNA helicase (nsp10), and the conserved nidovirus uridylate-specific endoribonuclease, (NendoU or U; nsp11) – is encoded within ORF1b (Ulferts and Ziebuhr, 2011). The calculated RFS efficiency of the RFS1 (6–7%–1 RFS; 16–20%–2 RFS) (Fang et al., 2012) and the RFS2 (~20%) (den Boon et al., 1991) demonstrates that ORF1b (nsp9–10) is generated approximately once out of every six translational events (15%), suggesting the stoichiometric requirements for the core replicative machinery is low compared to the 5' encoded replicase proteins.

RdRps form a characteristic right hand configuration (thumb, palm, finger(s)) with the thumb and fingers in contact to create a pocket for substrates (Ferrer-Orta et al., 2006). Comparison of single- and double-stranded RNA virus RdRps show structural similarity even though there is low sequence homology between classes (Ferrer-Orta et al., 2006). All polymerases share a core set of conserved motifs, suggesting a common ancestor (Sabanadzovic et al., 2009). The structure of the RdRp possesses an additional conserved motif (O'Reilly and Kao, 1998). The RdRp of nidoviruses is phylogenetically clustered with the picorna-like virus superfamily (Gorbaleya et al., 2006; Koonin, 1991) but possesses a SDD (Ser-Asp-Asp) signature, located within the active site on the palm side of the RdRp. The nidoviral SDD motif is a hallmark of the viral family that discriminates it from all other positive-sense RNA virus groups that contain a GDD (Gly-Asp-Asp) motif (den Boon et al., 1991). The SDD motif at this position was shown to be critical for EAV replication (van Dinten et al., 1999); surprisingly a S→G mutation within the PRRSV RdRp was replication competent (gRNA) but displayed deficiencies in sgRNA synthesis (Zhou et al., 2011). Another salient fact is that the arterivirus RdRp does not possess the 3' proofreading abilities that other nidoviruses display (Lauber et al., 2013). The rate of random mutation introduction is therefore elevated (Forsberg, 2005; Forsberg et al., 2001, 2002), contributing to an abnormally high evolution rate estimated at between 4.71 x 10^2 and 9.8 x 10^2/synonymous site/year (Hanada et al., 2005).

Nsp10 encodes the PRRSV helicase protein (Bautista et al., 2002). The PRRSV multi-domain helicase (HEL) is composed of the core 1A and 2A canonical domains found in super-family 1 type helicases, a flexible accessory domain (1B), and a unique zinc-binding domain (ZBD) (Deng et al., 2014). The PRRSV HEL functions to unwind dsRNA in a 5' to 3' polarity (Bautista et al., 2002). Both the flexible accessory domain and the ZBD are critical to replicative function of EAV including generation of gRNA and sgRNA (van Dinten et al., 2000; van Marle et al., 1999b). The helicase is predicted to function in concert with the RdRp to facilitate replication and transcription; however, it is not understood how the 5' to 3' directionality of the helicase and the 3' to 5' RdRp synthesis coordinate these activities (Fang and Snijder, 2010).

Nsp11 harbors the nidoviral uridylate-specific endoribonuclease (NendoU) domain (Ulferts and Ziebuhr, 2011). Originally described in coronaviruses, the NendoU of EAV was shown to be required for genome replication, and mutation of critical residues had varying deleterious effects on replication, and particularly on sgRNA synthesis (Posthuma et al., 2006). In examining the core nuclease region of EAV and PRRSV, single-stranded RNA was the preferred substrate, as was shown for the coronavirus sudden acute respiratory virus (SARS). However, no dependence on the divergent cation Mn2+ was seen. SARS nsp15, the coronavirus orthologue, was previously shown to Mn2+ dependent. EAV NendoU protein was also shown cleave 3' of pyrimidines, preferring uridine over cytidine, and releasing products with 2',3'-cyclic phosphate and 5'-OH ends. In addition, cleavage after unpaired over paired pyrimidines was preferred (Nedialkova et al., 2009; Ulferts and Ziebuhr, 2011). The function of NendoU in nidovirus replication has not been established.

How the PRRSV RdRp initiates replication, either by a primer dependent mechanism or by de novo synthesis is also unknown. To assess the activity of the EAV RdRp, a recombinant version of the polymerase was generated and assessed. EAV RdRp activity was found in the absence of a primer with poly(U) or poly (C) templates but not with poly(A) templates, indicating a de novo initiation method in a template specific manner (Beereens et al., 2007). Introduction of primers to either the poly(U) or poly (C) templates reduced RdRp activity (Beereens et al., 2007). Incubation with non-complementary bases (i.e. for Poly (U) template=GTP and UTP) did not result in isotopic labeling, which shows the RdRp did not function as a terminal transferase, and radioactively labeled primers were not incorporated into the synthetic non-viral templates (Beereens et al., 2007). De novo polymerase activity could not be detected on viral templates however. This finding suggests the arterivirus RdRp is catalytically active without other viral factors and capable of de novo synthesis, but may require other viral or cellular co-factors to efficiently perform replication or transcription processes. Similar research on PRRSV has not been accomplished to date.

It is generally believed that positive-sense RNA viruses use conformational switches in their terminal noncoding regions in the form of higher order RNA secondary structure to regulate translation, transcription of sgRNAs, and genomic replication (Beereens and Snijder, 2007). To define the minimal cis-acting 3' genomic element required for efficient PRRSV replication, progressive 3' deletions were introduced into self-limiting PRRSV replicons encoding an internal ribosome entry site (IRES)-driven luciferase (LUC) reporter within the deleted region (Choi et al., 2006). Only the smallest deletion, encoding the full M and N proteins, replicated to similar levels as the positive control. The next smallest deletion removing the M protein coding region but maintaining the complete N ORF resulted in significant loss of genome replication (Choi et al., 2006). Taken together, including the RdRp, HEL, 5'UTR, 3'UTR, ORF6, ORF7, and other unknown viral proteins, these data show that key genetic elements or protein interacting partners are required for efficient PRRSV replication and transcription, and are interspersed within multiple coding regions of the genome.
Viruses require the ability to selectively regulate transcription and translation processes both temporally and quantitatively in a highly ordered and balanced process (Pasternak et al., 2000). The expression mechanisms between the nonstructural replicase proteins and the structural proteins are fundamentally separated within the PRRSV genome facilitating rapid expression of nsps from the gRNA and subsequent amplification of sgRNA transcripts through a differential transcription cascade (Pasternak et al., 2004). sgRNAs are synthesized by the viral RdRp through a highly ordered process encoded within the PRRSV genome. The set of nested (Nido Latin: nested) sgRNAs encode noncontiguous gRNA sequence including both the 5'UTR and the polyadenylated 3'UTR as well as one or more ORFs from the 3' region of the genome (ORF2–7) but lack the entire large ~12 kb replicase coding region (ORF1a/b) (Fig. 1) (Gorbalenya et al., 2006; Pasternak et al., 2006; van Berlo et al., 1982). It was originally hypothesized that nidoviral sgRNAs could be generated through a free 5'UTR priming stage (Baker and Lai, 1990; Baric et al., 1983) but was ultimately shown...
that nidoviruses utilize a discontinuous sgRNA transcription strategy (Sawicki and Sawicki, 1995; van Marle et al., 1999a). Discontinuous replication proceeds through a replicative fusion of the viral genome 5'UTR to one of many downstream 3' sites through base pairing interactions between sense and antisense stem-loop (SL) structures via long-range RNA–RNA interactions during negative strand synthesis (van Marle et al., 1999a). Specifically, an antisense transcription–regulating sequence (TRS) at or near the 5' end of each structural protein coding region (ORF2–7) can each individually form a kissing-loop interaction with a conserved TRS sequence (UUAACC), located at the 3' terminus of the 5'UTR. sgRNAs synthesis and gRNA replication utilize a similar initial synthesis mechanism, whereby negative strand transcriptional extension from the 3' termini is completed until a TRS signal sequence within the body of the genome is encountered (Fig. 2A) (den Boon et al., 1996; Pasternak et al., 2004). The body TRS is ordered into a SL structure encoding a conserved heptanucleotide primary sequence (body TRS signal) (den Boon et al., 1996) within the loop structure. This signal halts transcription of the negative strand and a “decision” is made between transcriptional read-through and continuation of synthesis, or a disjoining of the transcriptional machinery and rejoining to the common leader TRS (antisense to body–TRS) by complementary base-pairing with a second SL structure within 5'UTR (Pasternak et al., 2003; van Marle et al., 1999a). The leader TRS is located within the 5'UTR directly upstream of the AUG start codon of nsp1α and is part of a highly ordered and well conserved RNA secondary structural motif between the 5'UTR and the 5' nsp1α coding region. Transcriptional read-through of all body TRS sites will result in gRNA synthesis. Decoupling from the genomic strand at the body TRS and rejoining at the leader TRS results in noncontiguous transcription that lacks a large central region of the genome, yielding one of at least six sgRNA products (dependent on which body TRS is utilized) (Fig. 2A) (Pasternak et al., 2001). The structural integrity of the EAV nsp1 region, composed of two papain-like protease domains and a predicted N-terminal zinc finger, was also indispensable for transcription, and has been shown to interact with the cellular cofactor p100 (Tijms et al., 2007, 2001; Tijms and Snijder, 2003). Again, all research completed on the arterivirus mechanism of sgRNA synthesis was through the use of the model EAV strain adapted for growth in tissue culture. Similar research on PRRSV, with strain differences in replication rates and its immense viral strain in depth. Furthermore, the body TRS motifs and the distance upstream from the starting AUG for the expression of sgRNAs differ between Type 1 and Type 2 PRRSV, except for the predominant sgRNA7 transcript (Meulenberg et al., 1993a; Nelsen et al., 1999). Genetic heterogeneity was also noted at these junction sites, differing by a single base to a couple of nucleotides, showing there is slight flexibility within the disjoining and reattachment of the viral RdRp during this step (Meng et al., 1996b; Nelsen et al., 1999). On top of this, surveying Northern analyses of different PRRSV strains shows that each strain has different quantities of each sgRNA transcript size, and often differ in their TRS motifs (Gauger et al., 2012; Guo et al., 2013b; Wang et al., 2008). Thus, the mechanism(s) regulating PRRSV sgRNA synthesis when comparing different viral strains appear to be more complex than appears when examining one viral strain in depth.

**Heteroclite RNAs**

Defective interfering (DI) RNAs are a normally observed by-product of positive-sense RNA virus replication, particularly under high multiplicity of infection (m.o.i.) culturing conditions (Masters, 2006; Molenkamp et al., 2000a; Pattnaik et al., 1992; Xiao et al., 2011). DI RNAs are generated through nonhomologous recombination between viral genomes resulting in random internal deletions but still encode the replication elements essential for generation of
defective progeny virus including the genes encoding for the polymerase, essential replicase proteins, and capsid protein(s) (Yuan et al., 2000). Unlike DIIs that have been described in other nidoviruses, a group of “heteroclite” sgRNAs (heteroclite = deviating from common forms or rules) were identified within PRRSV replicative products of unusual structure but containing large internal deletions (Yuan et al., 2000). Heteroclite sgRNAs species were identified within infected cells, purified virions, porcine alveolar macrophages infected with field isolates, under natural infection conditions, within plaque-purified viral infections, and are assumed to result from homologous recombination at atypical nucleotide stretches (Fig. 2B) (Yuan et al., 2000, 2004). The essential replicative products such as the viral RdRp were found to be absent within the heteroclite RNAs, discriminating them from prototypical DI genomes. In addition, heteroclites do not appear to interfere with ongoing genomic RNA and sgRNA transcription (Yuan et al., 2000, 2004). Sequencing analysis showed a short site of two to seven conserved nucleotides between the 5’ and 3’ joining regions, but these aberrant TRS motifs varied in sequence (Yuan et al., 2004). When RT-PCR products of culture supernatants of strain VR-23332 using a 5’ and 3’ primer pair were analyzed, at least 9 bands were discriminated (S1–S9). The result of this analysis and subsequent nucleotide sequencing showed that many similar-sized heteroclites were present in each band, but each band represented several individual heteroclites with different TRS motifs. The 5’ region of the heteroclites encoded ORF1a proteins including one or more of the papain-like proteases and joined within the downstream coding region either in-frame or within alternative reading frames, perhaps producing aberrant proteins (Yuan et al., 2000, 2004). These RNA species persist in a range of experimental culturing conditions including low m.o.i. passage and plaque purification (Yuan et al., 2000), and were found to be packaged within the PRRSV virion (Yuan et al., 2004). All identified heteroclites included at least the first 476 nt of the Type 2 prototype, strain VR-2332, and later studies pinpointed 35 nucleotides in nsp1 that bound to the N protein and therefore was an important element in viral packaging (Baig and Zakhartchouk, 2011). Additional DI RNA that possesses many similar genetic features has been identified, but each contains a smaller deletion (nsp2–9) and encodes all structural proteins (Xiao et al., 2011). Currently there is no known function for heteroclite sgRNAs, but they have been proposed as a packaging vector to study the effect of viral factors, or the effect of exogenous elements on viral replication, translation, progeny phenotype, or immune response/modulation (Yuan et al., 2000). They may also allow for recombination events to occur in the background of an ongoing PRRSV co-infection of two or more dissimilar strains. In this way, new viral sequences may be allowed to coexist with and interact with nascent viruses via recombination in the background, occasionally leading to new viral species with enhanced properties. Hypothetical at this juncture, this concept should be evaluated by defined recombination studies.

PRRSV recombination

The mechanism of PRRSV recombination is ill described (Fig. 3). Homologous recombination was first described for nidoviruses using mouse hepatitis virus (MHV), a coronavirus, and the authors posited that less than full-length RNA intermediates might be generated during viral RNA replication. These early studies led to the proposal that MHV replication proceeded in a discontinuous and nonprocessive manner, perhaps at sites of secondary and tertiary structure thus generating free RNA intermediates, which could be used in RNA recombination via a copy-choice mechanism (Lai, 1992; Lai et al., 1985; Makino et al., 1986). In the absence of selection pressure, MHV RNA recombination was found to be random, but that only certain recombinants are selected over passage in tissue culture, leading to the conclusion that there were “hotspots” for recombination (Banner and Lai, 1991). Recombination was detected during both negative and positive strand RNA transcription, and took place not only between two different viral

Fig. 3. Intergenomic homologous recombination between two different co-infecting PRRSV strains (green and blue bars) of the same genotype during negative strand synthesis (crosshatched bars). Early studies on coronaviruses suggested that intergenicomic recombination could also occur during replication of positive strand RNA. The PRRSV genome is represented at the top of the figure. See legend of Fig. 2 for further detail.
strains but also between one replicating viral RNA and transfected non-replicating MHV RNA fragments. Furthermore, the recombinants could be detected after viral growth in cell culture and in animals, and successful recombination occurred more frequently within a hypervariable region which was also subject to deletion (Liao and Lai, 1992). These results were shown to be representative of other coronaviruses, most notably infectious bronchitis virus (IBV), where numerous reports detail the ability of the virus to recombine in the field (Kottier et al., 1995; Toro et al., 2012; Wang et al., 1993). In the case of IBV, a pathogen of poultry, recombination has been shown to be robust, perhaps due to housing poultry in large flocks (Cavanagh and Davis, 1993). Most of these early conclusions mirror what has been shown to occur in PRRSV. Kapur et al. (1996) provided strong statistical evidence for intragenic recombination or gene conversion in ORFs 2, 3, 4, 5 and 7, but not in ORF 6. The first laboratory examination of PRRSV recombination was done using two different PRRSV strains to infect MA-104 cells. Differential primer pairs were used in RT-PCR studies to examine cloned cell culture progeny for recombination events over an 1182 nucleotide span encompassing part of ORF3 to ORF5 of Type 2 PRRSV. Five clones were selected for sequence analysis, which revealed that four clones each represented a single unique crossover, and one clone appeared to be a triple crossover recombinant. RNase treatment of the cell supernatant before RT-PCR analysis proved that the recombinant RNAs were protected from degradation and therefore represented packaged viral RNAs. Finally, the investigators showed that the recombinants could be detected for up to three passages, but eventually were overtaken by one parental strain that had increased replication kinetics in the cultured cell line. Recombination frequencies of up to 10% were estimated and recombinants could also be found in animals (Murtaugh et al., 2001; Yuan et al., 1999). A similar experiment was completed with Type 1 PRRSV, reporting frequencies of only 0.1 and 2.5% RNA recombination occurring within a 621 bp fragment, but also noting that recombination events are correlated with the size of the fragment analyzed (van Vugt et al., 2001).

There are many reported algorithmically detected instances of recombination occurring between PRRSV field strains of the same genotype, as well as defined coinfection studies in swine (Li et al., 2009; Liu et al., 2011; Martin-Valls et al., 2014; Shi et al., 2013a). Recombination hotspots have been observed to take place within the 3’ end structural genes, as well as in ns2, and ns9 (Li et al., 2009; Liu et al., 2011). A more thorough analysis has shown that multiple breakpoints of recombination were detected by genetic algorithm recombination detection (GARD) software all along the genome of both Type 1 and Type 2 isolates. GARD analysis of 25 Type 1 genomes produced 11 statistical breakpoints. Similarly, 55 Type 2 genomes led to the identification of 9 breakpoints (Martin-Valls et al., 2014).

Rare evidence exists for nonhomologous recombination between the PRRSV genome and other RNA segments. A survey of 3188 field viruses (18 Type 1) sequenced in the G5 region (597–618 bp) by the University of Minnesota Veterinary Diagnostic Laboratory showed that a key segment coding for G5 hypervariable region 1 was subject to insertions and deletions (Faaberg, 2007). This region contains 219 different patterns for an 8 amino acid stretch in hypervariable region 1, inducing non-neutralizing antibodies (Ostrowski et al., 2002; Plagemann et al., 2002), but was also seen to have additions (up to 6 amino acids) or deletions (1 amino acid). Two particular field isolates, encoding an extra 6 amino acids in the hypervariable region (NGGMRKATANSSNS), were found to be identical in nucleotide sequence (GGGGGGAUGAGGACCCG) in that 6 amino acid stretch to PRRSV ORF1 sequence, as well as many other swine host transcripts and other pathogens (unpublished data).

Although breakpoints and recombinants are valuable tools to understand viral evolution, there is a paucity of research directed toward understanding the viral and host machinery that PRRSV utilizes to successfully carry out recombination. Except for the illustration that mutations in the SDD motif of nsp9 of PRRSV results in viral replication without sgRNA transcription, and the HEL and NendoU domains being critical for both genomic and sgRNA replication in EAV and PRRSV, no detailed mechanism for arterivirus homologous and nonhomologous recombination has been put forward.

Phenotype and genotype diversity

The molecular evolution of PRRSV has been examined by many investigators in detail and will not be addressed in this review (Frossard et al., 2013; Shi et al., 2013a, 2010a, 2010b, 2013b; Stadejek et al., 2013). The cause of such rapid evolution may be primarily due to the lack of PRRSV RdRp proofreading and tremendous viral recombination, resulting in an extraordinary diverse composition of isolates with varying pathogenicity. From the emergence of PRRS in the United States in 1987 (Keffaber, 1989), it was apparent that there were several circulating viruses besides the USA prototype virus, VR-2332 (Collins et al., 1992). The disease was first recognized as mostly a reproductive disease, causing anorexia, late term abortions, and delayed return to estrus in sows. The infection of sows also led to increased preweaning mortality in young pigs that survived. Historically, interstitial pneumonitis, lymphomononuclear encephalitis, and lymphomonomonuclear myocarditis in piglets and focal vasculitis in the brain of the sow were seen. In nursing, growing, and finishing pigs, mild flu-like symptoms are evident, with pronounced hyperpnea, fever, and interstitial pneumonitis (Collins et al., 1992). In Europe, similar disease phenotypes were observed, with the additional finding that sows sometimes had blue ears, but antigenic differences were seen between European virus isolates from different countries and these were radically dissimilar from United States and Canadian isolates (Albina et al., 1992; Drew et al., 1995; Hopper et al., 1992; Paton et al., 1991; Plana et al., 1992; Wensvoort et al., 1992, 1991). Similar findings based on herd clinical symptoms as well as seroconversion were also reported in the USA and Canada (Dea et al., 1992; Morrison et al., 1992). The first nucleotide sequencing efforts directed at the PRRSV 3’–end of the genome confirmed these antigenic findings, revealing approximately 10% nucleotide differences between Type 2 isolates and 40% between Type 1 and Type 2 isolates (Drew et al., 1997; Kapur et al., 1996; Mardassi et al., 1994b; Meng et al., 1994; Meulenbergs et al., 1994; Pesch et al., 2005). Further evolution now places the divergence within both genotypes at >15% when comparing whole genomes (Han et al., 2006; Van Doorselaere et al., 2012). Investigators soon found that these differences were also reflected in the degree of pathogenicity caused by different viral isolates (Halbur et al., 1995, 1996).

Key events in PRRSV diversity were the emergence of an “atypical” or “acute” variant that appeared in Iowa, USA in the mid 1990s (Meng et al., 1996a; Mengeling et al., 1998), the sudden appearance in 2001 of a novel class of PRRSV named MN184 in Minnesota, USA (Han et al., 2006), the notable Type 2 highly-pathogenic PRRSV (HP-PRRSV) in 2006 in China and subsequently most of Asia (Tian et al., 2007), and the recent demonstration of enhanced pathogenicity of the Type 1 Lena strain (Karnyuchuk et al., 2010). Some PRRSV isolates have been shown to be neurovirus- lent (Rossov et al., 1999; Tian et al., 2007), and most recently the ability of particular strains to depress the swine immune response has been shown to vary (Brockmeier et al., 2012; Guo et al., 2013a, 2013b; Wang et al., 2013). Since the beginning, instances of increased virulence have emerged episodically in different regions
of the world. In each case, however, the appearance of the novel isolates is sudden and the result of a significantly divergent PRRSV genome sometimes accompanied by insertions or deletions in most often nsp2 ([Brockmeier et al., 2012]), and often containing detectable recombination breakpoints (Shi et al., 2013a). The remarkable phenotypic and genetic diversity is most likely the result of the innate attribute of PRRSV persistence, replication, and recombination to yield an extraordinarily flexible viral genome, attempting to circumvent attempts to eradicate the pathogen from swine by the host response, vaccination or other means.

Disclaimer

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

Acknowledgments

This work builds upon the doctoral thesis of Matthew Kappes, Department of Veterinary Microbiology and Preventive Medicine, Iowa State University, Ames, IA, 2014, entitled “Identification and characterization of a novel structural protein of porcine reproductive and respiratory syndrome virus, the replicase nonstructural protein 2”. Project 3625-32000-108-00D of the USDA Agricultural Research Service provided support for Dr. Faaberg. Dr. Kappes was supported by BI Vetmedica, Inc., (Grant no. 991-3625-791), The National Pork Board (Grant no. 13-196) and by the USDA Agricultural Research Service. USDA is an equal opportunity provider and employer.

References


van Dinten, L.C., Rensen, S., Gorbalenya, A.E., Snijder, E.J. 1999. Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. J. Virol. 73, 2027–2037.


