RESEARCH REPORT
SWINE HEALTH

Title: Sequencing and characterization of genes encoding capsid proteins VP4, VP6 and VP7 from field porcine rotavirus strains**, NPB #12-099**

Investigator: Jianfa Bai

Institution: Kansas State University

Date Submitted: July 1, 2014

Industry Summary

The lack of genome sequence information largely limited our ability to generate accurate molecular diagnostics and to identify vaccine candidate for porcine rotaviruses, especially for group B and C strains. This project aimed to use more advanced sequencing technology to obtain full genome sequences of contemporary porcine rotavirus strains. Our data clearly illustrated the shift of porcine rotavirus from predominant group A (37%) to the current more prevalent group C (59%) strains. Within group A strains, the most important immunogenic VP7 protein has also shifted from predominant G5 and G11 some 20 years ago to the current G9 (46%) and G4 (44%) genotypes. Sequence information on VP4 of both group B and group C strains revealed rather divergent sub-clusters by phylogenetic analysis. These observations indicated that there are needs for the development of molecular diagnostics to include all three major groups, especially for group C, and G4 and G9 group A porcine rotaviruses as they become the predominant group or genotypes in the swine production systems. Our data also suggest that new vaccine candidate selection should consider group C, and G4 and G9 group A strains, and new vaccines should protect pigs against the infections from the new group or genotypes of porcine rotaviruses. For more information, contact Dr. Dick Hesse: dhesse@vet.ksu.edu; or Dr. Jianfa Bai: jbai@vet.ksu.edu.

We appreciate the following professionals' participation in the project, and their help in collecting samples: Drs. Lisa Tokach, Megan Potter, & Steve Henry, Abilene Animal Hospital, Kansas; Dr. Jeffrey Demint, Bern-Sabetha Veterinary Clinic, Bern, Kansas; Drs. Aaron Lower, Dyneah Classen, Lucas Funk, and Joe Conner, Carthage Veterinary Clinic, Carthage, Illinois; Drs. Chris Rademacher & Ed Kluber, Murphy Brown, Ames, Iowa; Dr. Ben Hause, Newport Laboratory, Worthington, Minnesota, and Drs. Mark Engle & Larry Coleman, PIC North America, Hendersonville, Tennessee.

Keywords

Porcine rotavirus, whole genome sequencing, next-generation sequencing, single-primer amplification, rotavirus group A, rotavirus group B, rotavirus group C, prevalence.

Scientific Abstract

The Veterinary Diagnostic Laboratory (VDL) at Kansas State University in Manhattan, KS, received over 200 swine fecal samples of suspected rotavirus infections between 2012 and 2014. The single-primer amplification technique was employed to create cDNA of the double-stranded RNA genome segments. This sequence-independent method enabled us to amplify any rotavirus segments present in the fecal samples in an unbiased manner. This, coupled with next-generation

sequencing technology, is a clear advantage over traditional sequencing because of its ability to generate full-length sequences of all eleven genome segments, even in samples that have mixed rotavirus infections. One hundred and fifty-four samples were successfully sequenced, and 97% of those yielded full or partial sequences for all eleven rotavirus gene segments, enabling the full genome constellations to be recorded. Twenty samples yielded sequences from two rotavirus groups (A&B, B&C, or A&C), and one sample produced sequences from all three common rotavirus groups (A, B, and C). The prevalence of porcine group A rotavirus has shifted, as nearly 60% of all the sequenced samples contained group C rotavirus, with group A at 37% and group B at 18%. The VP7, VP4, VP6, and NSP1 gene segments displayed the greatest diversity, based on the number of putative and/or confirmed genotypes identified in our sample set. The prevalence of particular porcine sero- or genotypes has also shifted during the period of time that researchers have been tracking them, and a total of 30 new putative group B and C rotavirus genotypes were identified.

Introduction

Each year rotavirus causes more than 600,000 childhood deaths worldwide, and the cost to manage rotavirus infections in the US reaches $1 billion. Porcine rotavirus not only causes gastroenteritis in piglets, but is also considered, along with bovine rotavirus, a reservoir for emerging human strains. The most effective way of preventing rotaviral infection is through vaccination. Existing human and animal vaccines can only provide limited protection due to genetic diversity of the strains and the emergence of new strains. Similar to influenza A virus, the segmented genome of rotavirus enables the virus to undergo reassortment of its genome with other strains fairly frequently. Application of vaccines generated from strains with narrow genetic background further imposed the selection pressure for new strains to emerge. The most critical criterion to a successful vaccine is that it covers the most prevalent and predicted emerging strains, and genetic information from contemporary strains is the key for new vaccine development. Sensitivity and specificity are two key criteria for diagnostic tests. An extremely specific test, i.e., based on a single strain, may not be sensitive enough; a very sensitive test, on the other hand, may over-estimate disease condition, as not all strains are clinically important. A good diagnostic test has to balance its sensitivity and specificity, and design of a good diagnostic test relies on availability of genomic information. Limited genetic sequence data, especially of group B and C strains, is available for contemporary swine rotaviruses, which becomes the bottleneck for new vaccine and diagnostic method development; and what is available suggests greater genetic and antigenic diversity than previously realized. Understanding this diversity is critical in the design of new vaccines and diagnostic strategies. Through this NPB grant, we have generated full or partial genomes of more than 150 porcine rotavirus strains, and such information will be critical to aid new diagnostic method development, and to bridge the gap between prevailing strains in the field and candidate selections for new vaccine development.

Objectives

Our specific deliverables listed below will lead to greater understanding of genetic/antigenic diversity of group A, B & C rotaviruses, and will provide guidance to develop new vaccines and diagnostics;

a. Provide detailed genetic information on the diversity of contemporary group A, B and C strains;
b. Provide information on prevalence of contemporary rotavirus strains in swine production systems;
c. Identify new genotypes, and provide detailed genetic information as compared to known genotypes;
d. Provide information to develop new diagnostic method development, and new vaccine candidate selection.
e. Provide improved methodology for sequencing segments encoding capsid proteins from group A, B, and C rotavirus.

Materials & Methods

The original project design included a combination of three strategies to obtain sequences of the porcine rotavirus genes (segments 4, 6, and 9) encoding the VP4, VP6, and VP7 proteins. We anticipated that bovine rotavirus primers from a previous project would amplify some porcine rotavirus segments, but we recognized that the genotypes would most likely be different between bovine and porcine rotaviruses. To aid us in sequencing the unknown, we planned to use the single primer amplification technique (SPAT) coupled with porcine rotavirus segment-specific internal primers to obtain full-length gene sequences. However, did not need to design or use segment-specific primers to preferentially amplify segments of interest because a next-generation sequencing machine was made available to us (MiSeq, Illumina). For approximately the same cost as traditional "Sanger" sequencing of three genome segments, we were able to sequence all eleven genome segments simultaneously. All rotavirus segments that were present in any given sample were amplified and sequenced,

regardless of the rotavirus group to which they belong. In total, the Kansas State University Veterinary Diagnostic Laboratory (KSVDL) received more than 200 porcine fecal samples for rotavirus sequencing.

Total RNA was extracted with either the Direct-zol™ RNA MiniPrep kit from Zymo Research (Irvine, CA), or Trizol LS reagent from Invitrogen (Carlsbad, CA) coupled with the RNeasy MinElute kit from Qiagen (Valencia, CA). A primer (5'-CCCGTCGACGAATTCTTT-3'-NH_2) adapted from Lambden et al., (1992) with terminal 3'-NH_2 blocking group to prevent primer concatenation was ligated to the ends of segments. The ligation cocktail, which was adapted from Maan et al. (2007) by Jasperson (2011), consisted of 50 mM HEPES/NaOH, 18.33 mM $MgCl_2$, 0.01% BSA, 1 mM ATP, 3.33 mM DTT, 10% DMSO, 20% PEG_{6000}, 20 units Ribolock RNase Inhibitor from Thermo Scientific (Waltham, MA), 67.5 units T4 RNA Ligase from New England Biolabs (Ipswitch, MA), 200 ng primer (above), and 100-2,000 ng purified RNA in a total volume of 90 µl. The mixture was incubated at 37°C overnight, and the ligated RNA was purified with the QiaQuick Gel Extraction kit from Qiagen. The ligated primer served as template for a complementary primer (5'-AAAGAATTCGTCGACGGG-3') to anneal for cDNA synthesis, which was carried out using the SuperScript® III First-Strand Synthesis SuperMix kit from Life Technologies (Grand Island, NY). The cDNA was amplified by polymerase chain reaction (PCR) using the LA Taq® DNA Polymerase kit from TaKaRa/Clontech (Mountain View, CA) and the complementary primer (above). The sizes of the PCR products were observed by analysis on QiAxcel machine. Those samples which produced bands at or near the expected sizes of rotavirus genome segments (~600 bp to ~3,300 bp) were purified using the QIAquick PCR Purification Kit from Qiagen, and were included in the library preparation for sequencing on the MiSeq sequencer.

Briefly, next-generation sequencing libraries were prepared using the Nextera® XT DNA Sample Preparation Kit and MiSeq® v2 Reagent Kit from Illumina (San Diego, CA). The libraries were sequenced on a MiSeq sequencer using 300 cycles, which produced overlapping 2x150 bp sequencing reads. Those reads were assembled into full-length rotavirus segments and those segments were aligned with CLC Genomics Workbench software from CLC Bio/Qiagen using the slow/accurate algorithm with gap open penalty of 15, gap extension penalty of 5, and no gap close penalty to allow shorter reference sequences to properly align. MEGA (version 5.2.2, www.megasoftware.net) was used to construct phylogenetic trees from the alignment data. The phylogenetic trees used the neighbor-joining statistical method and the Jukes-Cantor substitution model with uniform rates among sites. Gaps were treated as pairwise deletions. Trees were observed with traditional or radiation-style branching. Each sample was genotyped for each gene segment based on comparison with reference sequences. Where no or few reference sequences were available (for group B and C rotavirus), putative genotypes were assigned based on percent nucleotide identity, and visualized by the phylogenetic dendrograms. The presence of distinct branching on the dendrogram, along with a nucleotide identity cutoff of approximately 85% was employed when necessary to differentiate genotypes.

I. Results

Group and Strain Diversity

Of the more than 200 porcine fecal samples received, 154 were successfully sequenced by next- generation sequencing methods. More than 97% of those yielded sequences for each of the eleven genome segments, allowing the full genome constellation to be determined. The number of porcine genotypes detected by this project for any gene segment ranged from 1 to 7 (Table 1). Genotype assignments for group A rotavirus are well-established by the Rotavirus Classification Working Group (Matthijnssens et al., 2008), and, perhaps consequently, our samples yielded a higher total number of group A genotypes for all 11 segments (33) than for the other rotavirus groups. Our samples yielded slightly fewer total genotypes for groups B and C (26 and 28, respectively), however, groups were not evenly sampled (see Rotavirus Group Prevalence, below). The genome segments that encode the VP7 and VP4 proteins tied for the highest sum of genotypes across all three groups (13) and NSP1 and VP6 tied for the next highest sum (9).

	Rotavirus A	**Rotavirus B**	**Rotavirus C**	**Sum**
VP7	7	2	4	13
VP4	6	3	4	13
VP6	2	2	5	9
VP1	2	3	1	6
VP2	2	3	2	7
VP3	2	3	1	6
NSP1	3	2	4	9
NSP2	2	2	2	6
NSP3	3	2	2	7
NSP4	2	2	1	5

NSP5	2	2	2	6
Sum	33	26	28	

Table 1. Number of porcine genotypes identified in samples sequenced for this project. Genotype assignments were taken from literature whenever possible, and additional genotypes were assigned when nucleotide identity was less than ~85% to known genotypes.

Group and Strain Prevalence

Group Prevalence. Twenty-four percent (37) of samples were identified as containing only group A rotavirus, 13% (20) contained only rotavirus B, and 49% (76) contained only rotavirus C. The remaining ~14% comprised specimens that had more than one group of rotavirus present: 8% (13) yielded sequences from rotavirus groups A and C, 4% (6) from groups A and B, 1% (1) from groups B and C, and 1% (1) from all three rotavirus groups A, B, and C. With 59% (91) of sequenced samples yielding rotavirus group C sequences, it was found to be the most prevalent strain of porcine rotavirus in samples for this project. Rotavirus group A was the second most prevalent with 37% (57) of samples, and group B was the third most prevalent with 18% (28) of samples.

Group A Genotype Prevalence. For the gene segment encoding VP7, genotypes G9 and G4 were the most prevalent, as they were identified in 46% (26) and 44% (25) of all sequenced samples, respectively (

Figure 1). Genotype G3 was the third most prevalent genotype with 16% (9); G6 was in 5% (3), G11 was in 4% (2), and GX and G10 were each in 2% (1) of group A sequenced samples. Note that the percentages may not add up to 100% because some samples yielded more than one genotype for a given gene segment, indicating a mixed infection. For VP4, genotype P[7] was the most prevalent, as it was identified in 39% (22) of all group A sequenced samples (Figure 2). Genotype P[23] was in 32% (18), P[13] was in 26% (15), P[6] was in 7% (4), P[5] was in 5% (3), and P[11] was in 2% (1) of group A sequenced samples. For VP6, genotype I5 was highly prevalent, as it was found in 93% (53) of group A sequenced samples, with genotype I2 found in 11% (6) of samples. For VP1, R1 was highly prevalent with 93% (52) of samples, and R2 was found in 7% (4). For VP2, C1 was highly prevalent with 95% (56) of samples, and C2 was found in 7% (4) of samples. For VP3, M1 was highly prevalent with 95% (53) of samples, and M2 was found in 6% (3) of samples. For NSP1, A8 was highly prevalent with 95% (53) of samples, and A3 and A13 were found in 4% (2) and 2% (1) of samples, respectively. For NSP2, N1 was highly prevalent with 95% of samples, and N2 was found in 7% (4) of samples. For NSP3, T1 was the most prevalent with 77% (44) of samples. T7 was the second most prevalent with 23% (13) of samples, and T6 was found in 5% (3) of samples. For NSP4, E1 was highly prevalent with 95% (54) samples, and E2 was found in 9% (5) of samples. For NSP5, H1 was highly prevalent with 95% (54) of samples, and H3 was found in 7% (4) of samples.

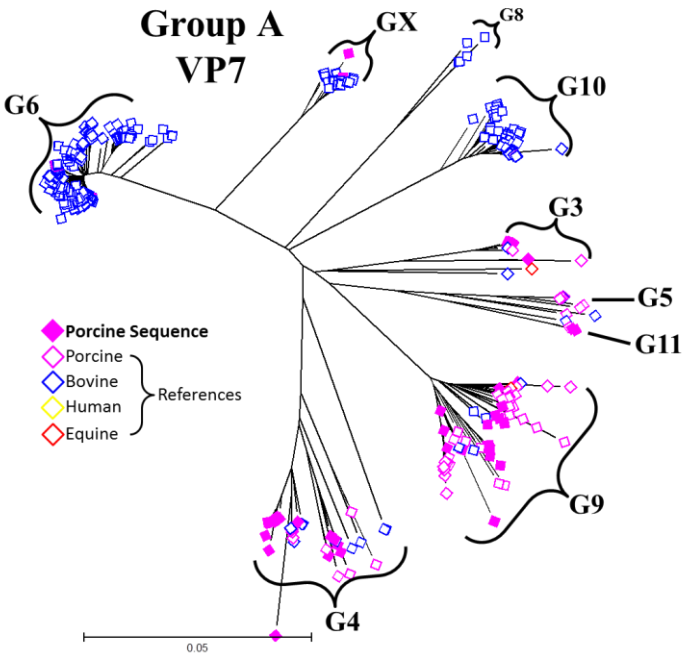

Figure 1. Phylogenetic dendrogram depicting the relationships between group A porcine rotavirus VP7 gene sequences from this study (solid pink diamonds) and reference sequences (empty diamonds).

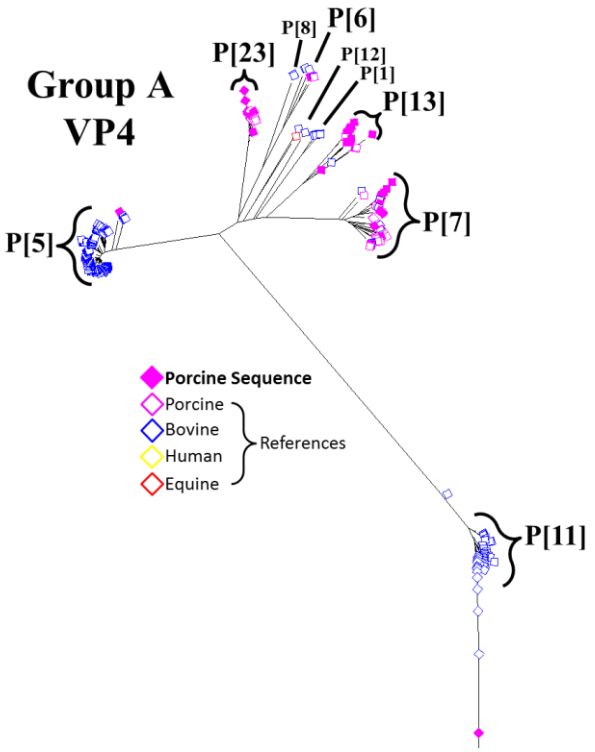

Figure 2. Phylogenetic dendrogram depicting the relationships between group A porcine rotavirus VP4 gene sequences from this study (solid pink diamonds) and reference sequences (empty diamonds).

Group B Genotype Prevalence. Genotype assignments have previously been assigned for only one of the eleven genome segments (VP7) of group B rotavirus (Marthaler et al, 2012). For the gene segment encoding VP7, G16 was identified in 100% (25) of sequenced samples, and G14 was found in 4% (1) of samples (Figure 3). The variables X, Y, and Z were used in place of numbers to indicate novel genotypes for the other ten genome segments of group B rotavirus. For VP4, three new genotypes were identified (Figure 4). P[X] and P[Y] were found in 68% (17) and 52% (13) of sequenced samples, and P[Z] was found in 8% (2) of samples. For VP6, IX was highly prevalent, as it was found in 88% (23) of samples, and IY was found in 12% (3) of samples. For VP1, RX was highly prevalent with 89% (24) of samples, and RY and RZ were found in 7% (2) and 4% (1) of samples, respectively. For VP2, CX was highly prevalent with 85% (23) of samples, and CY and CZ were found in 11% (3) and 4% (1) of samples, respectively. For VP3, MX was highly prevalent with 92% (23) of samples, and MY and MZ were each found in 4% (1) of samples. For NSP1, AX was highly prevalent with 88% (23) of samples, and previously sequenced genotype AY was found in 12% (3) of samples. For NSP2, NX was highly prevalent with 85% (23) of samples, and previously sequenced genotype NY was found in 15% (4) of samples. For NSP3, TX was highly prevalent with 88% (22) of samples, and TY was found in 12% (3) of samples. For NSP4, EX was highly prevalent with 88% (22) of samples, and EY was found in 12% (3) of samples. For NSP5, HX was highly prevalent with 85% (22) of samples, and previously sequenced HY was found in 15% (4) of samples.

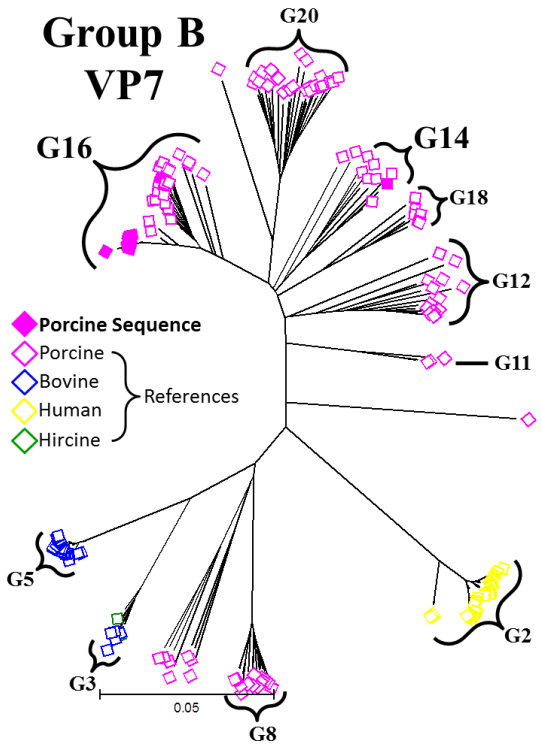

Figure 3. Phylogenetic dendrogram depicting the relationships between group B porcine rotavirus VP7 gene sequences from this study (solid pink diamonds) and reference sequences (empty diamonds).

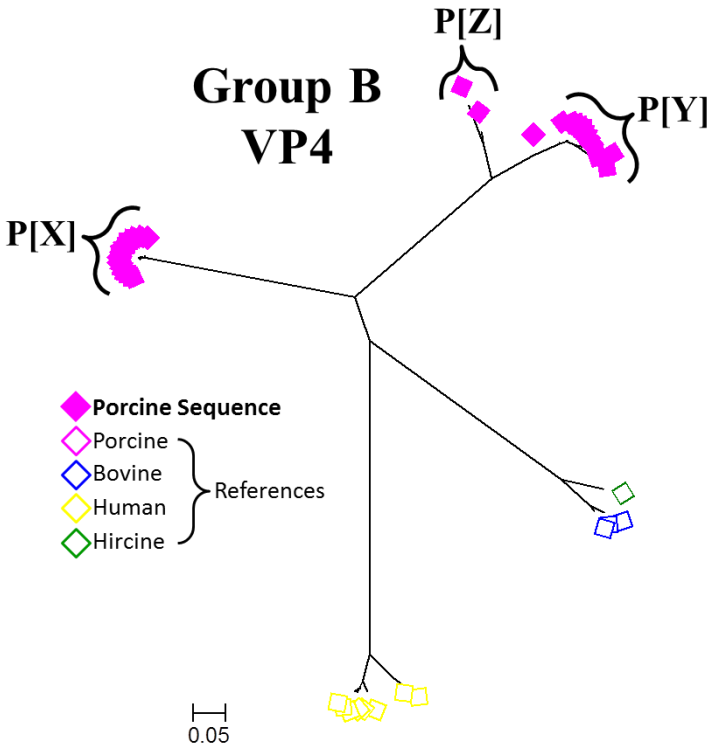

Figure 4. Phylogenetic dendrogram depicting the relationships between group B porcine rotavirus VP4 gene sequences from this study (solid pink diamonds) and reference sequences (empty diamonds).

Group C Genotype Prevalence. In the case of the group C rotavirus gene encoding VP7, a total of four porcine genotypes have been previously identified (G1 [Cowden], G3, G5, and G6) (Martella et al., 2007). No new VP7 genotypes were identified in this study. Genotype G6 was the most prevalent, as it was found in 70% (61) of samples (Figure 5). Genotype G5 was the second most prevalent with 44% (38) of samples, and G3 and Cowden-like G1 were found in 5% (4) and 2% (2) of samples, respectively. The genotype numbering system for the current study adheres to the system proposed by Soma et al. (2013); whereas any additional porcine clades for the other ten genome segments were assigned to numbers 4 and up, and the previously published Cowden-like clade remains genotype 1. For VP4, previously sequenced genotype P[6] and novel genotype P[5] were the most prevalent, with 62% (55) and 49% (44) of samples, while novel genotype P[4] and Cowden-like P[1] were less prevalent with 5% (4) and 1% (1) of samples (Figure 6). For VP6, at least one representative from each genotype had been previously sequenced (although unpublished). Genotypes I7 and I6 were the most prevalent with 52% (47) and 51% (46) of samples. Genotypes I5 and Cowden-like I1 were each found in 6% (5) of samples and I4 was found in 3% (3) of samples. For VP1, the Cowden-like genotype R1 was found in 100% (85) of sequenced samples. For VP2, novel genotype C4 was highly prevalent with 92% (81) of samples, and Cowden-like C1 was found in 11% (10) of samples. For VP3, the Cowden-like genotype M1 was found in 100% (84) of sequenced samples. For NSP1, novel genotype A5 was highly prevalent with 84% of samples. Novel genotypes A4 and A6 were less prevalent with 9% (8) and 7% (6) of samples, and Cowden-like A1 was found in 2% (2) of samples. For NSP2, Cowden-like N1 was highly prevalent with 84% (73) of samples, and novel genotype N4 was found in 16% (14) of samples. For NSP3, Cowden-like T1 was highly prevalent with 91% (80) of samples, and novel genotype T4 was found in 9% (8) of samples. For NSP4, Cowden-like E1 was found in 100% (88) of sequenced samples. For NSP5, Cowden-like H1 was highly prevalent with 91% of samples, and novel H4 was found in 9% of samples.

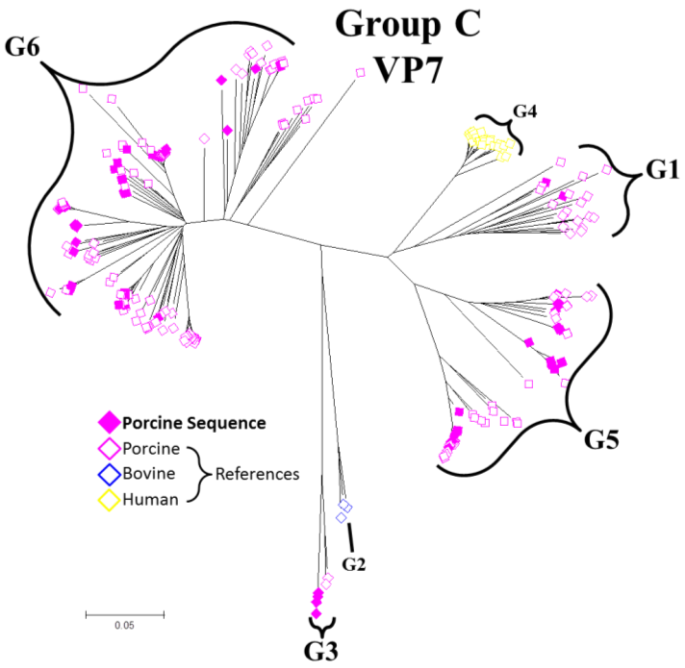

Figure 5. Phylogenetic dendrogram depicting the relationships between group C porcine rotavirus VP7 gene sequences from this study (solid pink diamonds) and reference sequences (empty diamonds).

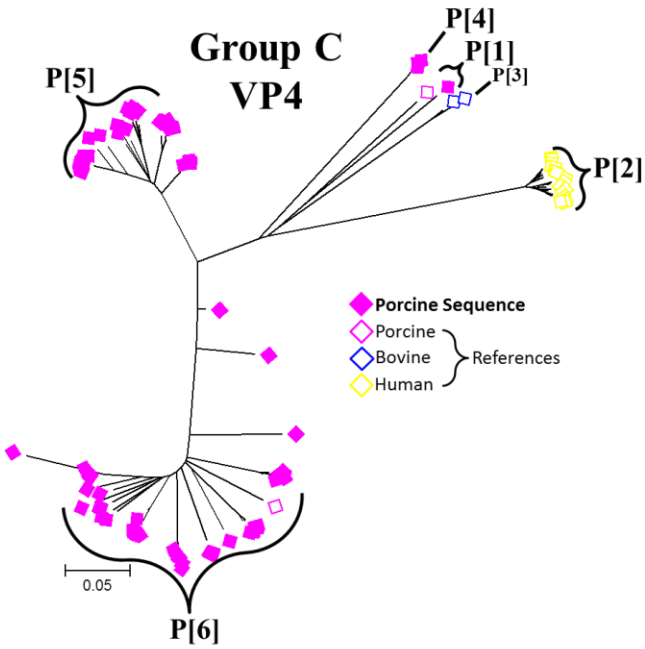

Figure 6. Phylogenetic dendrogram depicting the relationships between group C porcine rotavirus VP4 gene sequences from this study (solid pink diamonds) and reference sequences (empty diamonds).

New vs. Known Genotypes

All group A genotypes found in samples from this project have been previously classified, except for one. Group A VP7 genotype "GX", which was first identified by Hesse et al. (2012) in cattle, is on average 81% identical at the nucleotide level in the VP7 coding region to established genotype G6, both of which consist of predominantly bovine rotavirus sequences. Literature (Matthijnssens et al., 2008) suggests that the cutoff value to be considered a "new" group A VP7 genotype is 80% identity. We suggest that genotype "GX" should be considered a new genotype based on its representation as a robust branch on the group A VP7 phylogenetic tree and its similarity to genotype G6 being close to the recommended cutoff value.

Very few porcine group B and C rotavirus sequences were previously available for analysis on the NCBI database (Table 2). The gene segment encoding the VP7 protein had been previously studied the most in groups B and C and, consequently, had the highest number of sequences available for comparison (Group B: 106, Group C: 89). Notably, group B rotavirus had no published porcine sequences for seven of the eleven genome segments. After the sequences resulting from this study have been published or otherwise shared among veterinary diagnosticians, the number of available sequences will be greatly increased (Table 3).

	Rotavirus B	**Rotavirus C**
VP7	106	89
VP4	0	2
VP6	0	30
VP1	0	1
VP2	0	1
VP3	0	1
NSP1	15	1
NSP2	19	1
NSP3	0	1
NSP4	0	4
NSP5	18	1

Table 2. Numbers of full-length or near full-length porcine rotavirus groups B and C sequences that were available in the NCBI database prior to this study.

	Rotavirus B	**Rotavirus C**
VP7	132	194
VP4	32	106
VP6	26	136
VP1	27	86
VP2	27	92
VP3	25	85
NSP1	41	92
NSP2	46	88

NSP3	25	89
NSP4	25	92
NSP5	44	90

Table 3. Numbers of full-length or near full-length porcine rotavirus groups B and C sequences that will be available after this study.

A total of 21 genotypes over ten genome segments were considered "new" for porcine rotavirus B: 3 VP4, 2 VP6, 3 VP1, 3 VP2, 3 VP3, and 1 NSP1, 1 NSP2, 2 NSP3, 2 NSP4, and 1 NSP5. A total of 9 new genotypes over seven genome segments were identified for porcine group C rotavirus: 2 VP4, 1 VP2, 3 NSP1, 1 NSP2, 1 NSP3, and 1 NSP5.

Discussion

Diversity

At the onset of the current project, we anticipated that group A rotavirus would be the most prevalent, based on data from a previous study on bovine rotavirus (Hesse, 2012). However, our final count shows that group C was more prevalent in our porcine sample set (59% of samples). Not only are the genotypes shifting in prevalence (below), but also are the very groups of rotavirus. While the U.S. cattle industry appears to be currently affected by mainly group A rotaviruses (Hesse, 2012), it must be taken into consideration that the swine industry may be in more direct need of a group C rotavirus vaccine.

The high number of distinct genotypes for the gene segments encoding the VP4 and VP7 proteins was expected, based on the function of the corresponding proteins. VP4 and VP7 are both present in the outermost shell of the virus protein coat, and, along with VP6, provoke an adaptive immune response in the host. Because of this increased selection pressure, they are expected to evolve at a higher rate than other gene segments to maintain an evolutionary advantage over the host. Similarly, the nonstructural protein 1 (NSP1) interacts with host interferons which trigger an immune response, so its high level of diversity was also anticipated. Since we do not have much information on the direction of flow of rotavirus groups B and C to and from other species, it is possible that some porcine genotypes may have evolved first in other host species and then migrated into swine, although the converse is equally likely. In any case, the continuous spread of infection within and across host species is certainly responsible for the constantly increasing diversity of viruses.

Prevalence

Group A rotavirus strain prevalence has shifted dramatically over the period of time in which we have been monitoring strains in pigs. Twenty years ago, VP7 genotypes G5 and G11 were the most prevalent among swine (Gouvea et al., 1994), but no instances of G5 and only two instances of G11 were detected in the present study. In 2007, Khamrin et al., summarized previous investigations, stating that G3, G4, G5, and G1 in association with P[6] and P[7] had been the most common genotypes in swine. Contrastingly, in the current study P[7]-**G9** and **P[23]**-G4 were the most common genotypes (present in 26% [15] and 25% [14] of group A samples, respectively). Since there were previously very few sequences for porcine group B rotavirus, it is largely unknown whether the prevalence has shifted over time. Since the group B rotavirus samples in the current study were largely homogeneous for the VP7 gene segment, it would be of great benefit to the swine industry for sequence-independent whole genome sequencing to be carried out on a subset of those samples for which there has been a non-G16 VP7 sequence published. For group C rotavirus, we can at least compare to the Cowden porcine reference strain from Bremont et al. (1993). For five of the eleven gene segments (VP7, VP4, VP6, VP2, and NSP1), the prevalent genotype was not Cowden-like. For the other six gene segments, the Cowden-like gene genotype persisted as the most prevalent in the current study. This shift in prevalence further supports (along with genotypic diversity) that viral gene segments encoding the most biologically important proteins have been evolving the fastest.

New Genotypes vs. Known

A total of 30 putative new genotypes were identified in this study (21 from group B and 9 from group C). For group B rotavirus, Marthaler et al. (2012) characterized 20 VP7 G genotypes, 16 of them exclusively porcine, based on a slightly more stringent 80% nucleotide cutoff. In the yet-unknown group B rotavirus VP4 gene

segment, three putative genotypes were identified: P[X], P[Y], and P[Z]. While P[Z] is similar to P[Y] (~90% nucleotide identity) and may not be considered a different genotype by some standards, genotypes P[X] and P[Y] are strikingly dissimilar at only 64% nucleotide identity. The divergence between P[X] and P[Y]/P[Z] is nearly as great as between some distantly-related group A VP4 genotypes, suggesting that there is just as much diversity – and possibly just as many extant genotypes in group B rotavirus. Because all of the group B rotavirus samples in the present study were found to have the same G16 VP7 genotype (and one with dual G16/G14 genotypes), it is likely that we are greatly underestimating the full genetic diversity of group B rotavirus. In the group C rotavirus VP7 gene segment, at least one representative from each of the five porcine genotypes had been previously sequenced and genotypes were assigned numbers by Martella et al. (2007) based on an 85% nucleotide identity cutoff. For the VP4 gene segment, only two porcine strains had been previously sequenced (one Cowden-like P[1], and one P[6]). Those two genotypes were found to have approximately 83% nucleotide identity with each other, while novel genotypes P[4] and P[5] shared no more than 75% nucleotide identity with any of the other genotypes.

Improving Diagnostic Methods and Vaccines

The methodology used in this project is repeatable, and may be beneficial when a group-unknown rotavirus infection is suspected. Currently, other rotavirus group B and C researchers are using traditional sequencing which yields one sequence at a time, requires multiple sequencing runs for the larger rotavirus segments (for example, outer capsid protein gene VP4 at ~2.3 kbp), and can be impossible in the case of mixed infections. This method of single-primer ligation/amplification coupled with next-generation sequencing (NGS) is a clear improvement with its generation of a complete genome constellation (or multiple genome constellations, in the case of mixed infections) in one sequencing run. For researchers that wish to continue using traditional sequencing,.

The plethora of gene sequences yielded from this study will be the basis for developing new diagnostic tests that will be able to not only detect the presence of porcine rotavirus, but will be able to identify the group. Real-time PCR primers and probes, as well as gene-specific sequencing primers will easily be designed from these new sequences based on conserved regions in the sequence alignments. Sequencing primers may not be as valuable for diagnostic protocols, but in order to keep real-time PCR tests up to date, we must continue to track rotavirus sequence divergence and group/genotype prevalence. This collection of porcine rotavirus sequences will constitute an invaluable resource for designing new primers that target contemporary strains. Traditional sequencing will continue to be necessary whenever primer-independent whole-genome sequencing is unavailable or too costly. The information from this study on current group and strain prevalence will guide new vaccine candidate selection, and virus-like-particles for vaccine development will be able to be synthesized based on the sequences gained from this study.

References:

Bremont, M; Juste-lesage, P; Chabanne-vautherot, D; Charpilienne, A; Cohen, J. 1992. Sequences of the four larger proteins of a porcine group C rotavirus and comparison with the equivalent group A rotavirus proteins. *Virology* **186** (2): 684-692.

Gouvea, V; Santos, N; Timenetsky, MC. 1994. Identification of bovine and porcine rotavirus G types by PCR. *Journal of Clinical Microbiology* **32** (5): 1338-1340.

Hesse, D; Bai, J; Poulsen, E; Li, H; Anderson, J; Breazeale, B. 2012. Final report for: Genetic Diversity of Bovine Group A Rotaviruses. Pfizer, Inc.

Jasperson, D. 2011. Personal communication. USDA, ARS. Center for Grain and Animal Health Research (CGAHR), Arthropod-Borne Animal Diseases Research Unit (ABADRU), Manhattan, KS.

Khamrin, P; Maneekarn, N; Peerakome, S; Chan-it, W; Yagyu, F; Okitsu, S; Ushijima, H. 2007. Novel porcine rotavirus of genotype P[27] shares new phylogenetic lineage with G2 porcine rotavirus strain. *Virology* **361**: 243-252.

Lambden, PR; Cooke, SJ; Caul, EO; Clarke, IN. 1992. Cloning of Noncultivatable Human Rotavirus by Single Primer Amplification. *Journal of Virology* **66** (3): 1817-1822.

Maan, S; Rao, S; Maan, NS; Anthony, SJ; Attoui, H; Samuel, AR; Mertens, PPC. 2007. Rapid cDNA synthesis and sequencing techniques for the genetic study of bluetongue and other dsRNA viruses. *Journal of Virological Methods* **143**: 132-139.

Martella, V; Banyai, K, Lorusso, E; Decaro, N; Bellacicco, A; Desario, C; Corrente, M; Greco, G; Moschidou, P; Tempesta, M; Arista, S; Ciarlet, M; Lavazza, A; Buonavoglia, C. 2007. Genetic heterogeneity in the VP7 of group C rotaviruses. *Virology* **367**: 358-366.

Marthaler, D; Rossow, K; Gramer, M; Collins, J; Goyal, S; Tsunemitsu, H; Kuga, K; Suzuki, T; Ciarlet, M; Matthijnssens, J. 2012. Detection of substantial porcine group B *Rotavirus* genetic diversity in the United States, resulting in a modified classification proposal for G genotypes. *Virology* **433**: 85-96.

Matthijnssens, J; Ciarlet, M; Rahman, M; Attoui, H; Banyai, K; Estes, M; Gentsch, JR; Iturriza-Gomara, M; Kirkwood, CD; Martella, V; Mertens, PPC; Nakagomi, O; Patton, JT; Ruggeri, FM; Saif, LJ; Santos, N; Steyer, A; Taniguchi, K; Desselberger, U; Van Ranst, M. 2008. Recommendations for the classification of group A Rotaviruses using all 11 genomic RNA segments. *Archives of Virology* **153**: 1621-1629.

Soma, J; Tsunemitsu, H; Miyamoto, T; Suzuki, G; Sasaki, T; Suzuki, T. 2013. Whole-genome analysis of two bovine rotavirus C strains: Shintoku and Toyama. *Journal of General Virology* **94**: 128-135.