Detection and genetic characterization of porcine sapovirus from pigs with diarrhea

Huigang Shen[1, $], Jianfeng Zhang[1, 2, $], Phillip C. Gauger[1], Eric R Burrough[1], Jianqiang Zhang[1],

Karen Harmon, Leyi Wang[3], Ying Zheng[1], Thomas Petznick[4], Ganwu Li[1*]

[1]Department of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary

Medicine, Iowa State University, Ames, IA 50011, USA

[2]Institute of Animal Health, Guangdong Academy of Agricultural Sciences, Key Laboratory of

Livestock Disease Prevention of Guangdong Province, Scientific Observation and Experiment

Station of Veterinary Drugs and Diagnostic Techniques of Guangdong Province, Guangzhou,

Guangdong 510640, China

[3]Department of Veterinary Clinical Medicine and the Veterinary Diagnostic Laboratory, College

of Veterinary Medicine, University of Illinois, Urbana, IL 61802, USA

[4]Arkcare, 70929 574th Ave, Fairbury, NE 68352, USA

[$]These two authors contributed to this work equally.

*Correspondence

Ganwu Li, Department of Veterinary Diagnostic and Production Animal Medicine, College of

Veterinary Medicine, Iowa State University, Ames, IA. Email: liganwu@iastate.edu

Abstract

Porcine Sapovirus (SaV) was first identified by electron microscopy in the United States in 1980 and has since been reported from both asymptomatic and diarrheic pigs usually in mixed infection with other enteric pathogens. SaV as the sole etiological agent of diarrhea in naturally infected pigs has not previously been reported in the United States. Here, we used four independent lines of evidence including metagenomics analysis, real-time RT-PCR (rRT-PCR), histopathology, and in situ hybridization to confirm porcine SaV genogroup III (GIII) as the sole cause of enteritis and diarrhea in pigs. A highly sensitive and specific rRT-PCR was established to detect porcine SaV GIII. Examination of 184 fecal samples from the outbreak farm showed that pigs with clinical diarrhea had significantly lower Ct values (15.9 ± 0.59) compared to clinically unaffected pigs (35.8 ± 0.71). Further survey of 336 fecal samples from different states in the United States demonstrated that samples from pigs with clinical diarrhea had a comparable positive rate (45.3%) with those from non-clinical pigs (43.1%). However, the SaV-positive pigs with clinical diarrhea had significantly higher viral loads (Ct $= 26.0 \pm 0.5$) than those positive but clinically healthy pigs (Ct $= 33.2 \pm 0.9$). Phylogenetic analysis of 20 field SaVs revealed that all belonged to SaV GIII and recombination analysis indicated that intra-genogroup recombination occurred within the field isolates of SaV GIII. These results suggest that porcine SaV GIII plays an important etiologic role in swine enteritis and diarrhea and rRT-PCR is a reliable method to detect porcine SaV. Our findings provide significant insights to better understand the epidemiology and pathogenicity of porcine SaV.

INTRODUCTION

Sapovirus (SaV) is a member of the genus *Sapovirus* within the family *Caliciviridae*. It is a positive sense, single-stranded RNA virus with a genome of approximately 7.1 to 7.7 kb in length and a poly A tail at the 3′ end (Alhatlani, Vashist, and Goodfellow, 2015; Oka et al., 2016). The SaV genome contains two open reading frames (ORFs). The ORF1 encodes a large polyprotein and the major capsid protein VP1 and the ORF2 encodes the minor structural protein VP2 (Oka et al., 2016). The ORF1-encoded polyprotein undergoes protease processing to produce nonstructural (NS) proteins including NS1, NS2, NS3 (putative NTPase), NS4, NS5 (VPg), and NS6-NS7 (protease and RNA dependent RNA polymerase, RdRp) (Oka et al., 2015). SaVs are classified into different genogroups based on the VP1 sequences (Farkas et al., 2004; Scheuer et al., 2013). A pairwise identity of VP1 amino acid sequence (< 57%) was used to define a different SaV genogroup (Oka et al., 2016; Yinda et al., 2017). At present, SaVs are classified into 19 genogroups and at least 52 genotypes based on complete VP1 sequences using a pairwise distance cut-off value of ≤ 0.488 to distinguish different genogroups and ≤ 0.169 to distinguish different genotypes (Nagai et al., 2020; Oka et al., 2015). Among them, 17 genogroups (GII, GIII, and GV–GXIX) have been identified in animals and four (GI, GII, GIV, and GV) in humans (Scheuer et al., 2013) with genogroups II (GII) and V (GV) detected in both animals and humans (Oka et al., 2015). Thus far, at least eight genogroups (GIII, GV, GVI, GVII, GVIII, GIX, GX, and GXI) of SaVs have been associated with pigs (Nagai et al., 2020; Oka et al., 2016; Scheuer et al., 2013). Our recent genomic characterization study reveals that highly divergent porcine SaV strains co-circulate in the field in the US (Wang et al., 2020). In addition to nucleotide mutation, insertion and deletion in genome and recombination might contribute to virus evolution (Wang et al., 2020).

66   SaVs can cause gastroenteritis in humans and animals (Oka et al., 2015). Porcine SaV was first

67   identified by electron microscopy in the United States in 1980 in a mixed infection with other

68   viruses (Saif et al., 1980) and since then, SaVs have been reported from both diarrheic and

69   asymptomatic pigs (das Merces Hernandez et al., 2014; Wang et al., 2020; Wang et al., 2006).

70   Experimental infection of cell culture-adapted SaVs in gnotobiotic piglets successfully induced

71   enteritis and diarrhea (Flynn, Saif, and Moorhead, 1988; Guo et al., 2001). However, in the field,

72   SaVs are often identified as a mixed infection with other diarrheic pathogens (Dufkova et al.,

73   2013; Mijovski et al., 2010; Scheuer et al., 2013), while SaV as the sole etiological agent of

74   diarrhea in naturally infected pigs has, until this time, never been reported in the United States.

75   **MATERIALS AND METHODS**

76   **Clinical samples and history**

77   In February of 2019, the Iowa State University Veterinary Diagnostic Laboratory (ISU VDL)

78   received a diagnostic submission from a farm experiencing an ongoing problem with piglet

79   diarrhea in the lactation phase for more than 2 years (Case #1 in Table 1). The pigs reportedly

80   exhibited a pasty to occasional semiliquid diarrhea starting around 10 days of age. The diarrhea

81   was generally described as self-limiting, but the pigs would typically lose 1-2 lbs of expected

82   weaning weight. Fecal and intestinal tissue samples were received at the ISU VDL as part of this

83   diagnostic submission. Histopathological examination of intestinal samples from five clinically

84   affected piglets revealed that 5/5 small intestines had moderate villous atrophy with lymphocytic

85   infiltration in the lamina propria suggestive of an enteric viral infection (Fig. 1A-D). Porcine

86   epidemic diarrhea virus (PEDV), porcine deltacoronavirus (PDCoV), transmissible

87   gastroenteritis virus (TGEV), and rotaviruses (genogroups A-C) were not detected in the

4

88 intestinal contents using real-time RT-PCR (rRT-PCR) assays. Additionally, there was no

89 significant growth of common bacterial pathogens from the affected small intestines.

90 As part of a follow-up investigation, a total of 184 fecal swab samples were collected from pigs

91 with clinical diarrhea and healthy pigs of varying ages. These swabs included 34 samples from

92 pigs between 8 and 12 days of age, 100 samples from two-week-old pigs, and 50 samples from

93 pigs between 12 and 16 weeks of age. All samples were collected from the original submitting

94 farm and were subjected to SaV GIII rRT-PCR.

95 In addition, 234 fecal samples from swine with diarrhea and 102 from clinically healthy pigs that

96 were submitted to the ISU VDL for routine diagnostics by unrelated farms from different states

97 across the United States were also tested by SaV GIII rRT-PCR.

98 **Next generation sequencing (NGS)**

99 Total nucleic acid of clinical samples was extracted using MagMAX Pathogen RNA/DNA Kit

100 with KingFisher™ Flex System (Thermo Fisher Scientific) (Zhang et al., 2017). Double stranded

101 cDNA was synthesized using NEXTflex™ Rapid RNA-Seq Kit (Bioo Scientific Corp, TX).

102 Sequencing library was prepared using Nextera XT DNA library preparation kit (Illumina, CA)

103 with dual indexing. The pooled libraries were sequenced on an Illumina MiSeq platform at the

104 NGS unit in the ISU VDL, with 500-Cycle v2 Reagent Kit (Illumina) to generate 250 base-pair

105 paired-end reads by following standard Illumina protocols. Raw reads of each sample were

106 demultiplexed automatically on the MiSeq with the default settings.

107 **Bioinformatics analysis**

108 Raw sequencing reads were pre-processed using Trimmomatic v0.36 to remove adapters and

109 trim low quality ends (Bolger, Lohse, and Usadel, 2014). Raw reads and pre-processed reads

110 were subjected to sequencing quality analysis with FastQC

111 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to ensure the efficiency of

112 cleaning. Cleaned reads were fed to a comprehensive reference-assisted virus genome assembly

113 pipeline (Chen et al., 2018; Zhang et al., 2017) with modifications. The cleaned reads were

114 aligned to host reference genome using BWA-MEM (Li and Durbin, 2009). The non-host reads

115 were classified using Kraken v1.0 (Wood and Salzberg, 2014) and the unclassified reads were

116 further classified using Kaiju v1.6.2 (Menzel, Ng, and Krogh, 2016). KronaTools-2.7 (Ondov,

117 Bergman, and Phillippy, 2011) was used to generate the interactive html charts for hierarchical

118 classification results. Reads of interest were extracted and used for assembly by using ABySS

119 v1.3.9 (Simpson et al., 2009). The resulting contigs were manually curated to remove

120 contaminated contigs and trim chimeric contigs in SeqMan Pro (DNASTAR® Lasergene 11

121 Core Suite) and finally refined to obtain the genome sequence.

122 **Histopathology and Fluorescent in situ hybridization (FISH)**

123 Sections of duodenum, jejunum, and ileum were fixed in 10% neutral buffered formalin and

124 processed by standard paraffin embedding technique and then stained with hematoxylin and

125 eosin (HE). Histopathological evaluation was performed by diagnostic pathologists at the ISU

126 VDL. Sections from selected paraffin blocks with lesions of atrophic enteritis were then used for

127 FISH development with an Alexa555-labeled probe Psap5122 (Table 2) and an in-house

128 hybridization buffer as previously described (Burrough et al., 2013) with slight modification (20

129 mM of Tris, 0.9 M of NaCl, 0.1% sodium dodecyl sulfate [SDS] buffer, 20% formamide, 10%

130 dextran sulfate [pH 7.2]). In situ hybridization was performed using an automated system

131 (Discovery Ultra, Roche Diagnostics, Indianapolis, IN) with the manufacturer's standard

132 protocol and commercial reagents with the exception of the hybridization buffer as noted above.

133 Hybridization was carried out at 52 °C for 12 hours.

Development of real-time RT-PCR (rRT-PCR) for porcine SaV GIII

134

135 An rRT-PCR assay specific for porcine SaV GIII was developed with the primers PSapV-F and

136 PSapV-R, and probe PSapV-P (Table 2), which were commercially synthesized (Integrated DNA

137 Technologies, Coralville, IA). The real-time RT-PCR was performed in a reaction mixture of 25

138 μl containing 12.5 μl 2× AgPath-ID RT-PCR Buffer (Applied Biosystems), 1 μl 25× RT-PCR

139 Enzyme Mix, 1 μl (final concentration 0.4 μM) of each of the primers, 0.5 μl (final concentration

140 0.2 μM) of probe, 4 μl nuclease-free water, and 5 μl extracted RNA. The amplification was

141 performed at 48 °C for 10 minutes, 95 °C for 10 minutes, followed by 40 cycles at 95 °C for 15

142 s, 55 °C for 20 s, and 60 °C for 45 s. For each amplification plot a cycle threshold (Ct) value was

143 calculated representing the cycle number at which the reporter signal was above threshold.

144 The sensitivity of the real-time PCR assay was evaluated by using serial 10-fold dilutions of the

145 *in vitro* transcribed RNA, which was produced by Transcript Aid T7 High Yield Transcription

146 Kit (Thermo Fisher Scientific, Waltham, MA) with primers Sap-T7-F1 and Sap5193R (Table 2).

147 The specificity of SaV GIII rRT-PCR was evaluated using oligo DNA fragments with length

148 between 107 and 117 nt designed for SaV GV, GVI, GVII, GVIII, GIX, GX and GXI according

149 to their corresponding sequences at the same region that SaV GIII rRT-PCR targeted (Table 2).

150 All oligo DNAs were synthesized in the DNA Facility at Iowa State University. In addition, the

151 specificity was evaluated with other viral pathogens related to swine diarrhea including virus

152 isolates of PEDV, TGEV, PDCoV and rotavirus A, and fecal samples positive for rotavirus B

153 and C, and bacterial pathogens *Salmonella* spp. and *E. coli*.

Sequence comparison and analysis

154

155 Sequence comparison was performed with MegAlign (DNASTAR® Lasergene 11 Core Suite).

156 Sequences representative of different SaV genogroups were obtained from GenBank and aligned

157 with the 20 SaV genome sequences in the present study. Phylogenetic trees were constructed

158 using the neighbor-joining method in MEGA 7.0 (Kumar, Stecher, and Tamura, 2016) with p-

159 distance as the substitution model. Bootstrap analysis was carried out with 1,000 replicates.

160 Recombination analysis was performed in the region from ORF1 55 nt to ORF2 522 nt using

161 SimPlot software v. 3.5.1 (Lole et al., 1999).

162

163 **RESULTS**

164 **Next-generation sequencing (NGS)**

165 NGS and bioinformatics analysis were performed on the fecal and tissue samples from the

166 outbreak farm (case 1 in Table 1). The hierarchical classification results were shown as an html

167 pie-chart by Kraken and Krona Tools. Surprisingly, the classification results showed that 100%

168 of the viral reads belong to SaV. The complete genome sequence of porcine SaV was then

169 assembled and designated as NE9550B. No other viral pathogens were identified in the

170 classification; these results were consistent with the real-time PCR results in which all other

171 common viruses related to porcine enteric disease were negative (Table 1). NGS analysis also

172 identified SaV as the sole viral pathogen from fecal samples of 8 to 10-day-old diarrheic piglets

173 (case # 2) (Table 1).

174

175 **Histopathology and fluorescent in situ hybridization (FISH)**

176 Histopathological examination revealed lesions consistent with viral infection in the small

177 intestine. Specifically, there was moderate segmental villous blunting, atrophy, and contraction

178 (Fig. 1A-B) with marked, locally extensive villous vascular congestion (Fig. 1C). Low to

179 moderate numbers of lymphocytes, macrophages and fewer neutrophils were observed in the

180 lamina propria (Fig. 1B-C) and multifocal crypts contained low to moderate numbers of

181 neutrophils and cellular debris. Apical enterocytes over affected villi ranged from vacuolated to

182 flattened or attenuated (Fig. 1C) and some villi lacked apical enterocytes with exposed lamina

183 propria (Fig. 1D). In the colon, multifocal glands contained low to moderate numbers of

184 neutrophils admixed with cellular debris and low numbers of lymphocytes and macrophages

185 were identified in the lamina propria (data not shown).

186 For slides with histopathological lesions that also originated from SaV rRT-PCR positive

187 intestine tissue samples, additional sections of small intestine were further subjected to FISH.

188 Fluorescently-labeled enterocytes were identified along the attenuated villi indicating that the

189 lesion was caused by SaV infection (Fig. 1E-F).

190 **Development of real-time RT-PCR (rRT-PCR)**

191 Standard curves were established using the 10-fold serially diluted RNA control *in vitro*

192 transcribed from a DNA fragment amplified by SaV GIII primers Sap-T7-F1 and Sap5193R

193 (Table 2). RNA control was serially diluted from 2.4×10^1 to 2.4×10^9 copies per reaction (in 5

194 μl) and amplified in triplicate. The threshold cycle (Ct) values were plotted against the copy

195 numbers. The SaV GIII rRT-PCR was able to detect the RNA control over a linear range. The

196 standard curve obtained showed a slope of -3.80, an intercept of 43.33, and a regression

197 coefficient (R2) of 0·9998. The detection limit of the assay using RNA control as template was 5

198 copy numbers per reaction. No PCR amplification signals were detected in the SaV GV, GVI,

199 GVII, GVIII, GIX, GX and GXI oligos, PEDV, TGEV, PDCoV and rotavirus A isolates,

200 rotavirus B and C fecal samples, or the *Salmonella* spp. and *E. coli* cultures by the SaV GIII rRT-

201 PCR (Ct > = 40).

202

rRT-PCR examination of porcine sapovirus in clinical and non-clinical pigs from the outbreak farm

The intestine and fecal samples of the original case from the outbreak farm were analyzed using the SaV rRT-PCR assay. SaV was detected with high viral amounts in both intestine and fecal samples, with Ct values of 18.6 in the tissue sample and 15.4 in the fecal sample, respectively. A follow-up investigation was performed with samples from the outbreak farm. In total, 34 fecal samples from pigs 8 to 12 days old with clinical signs were examined and 85.5% of samples were found positive with an average Ct value of 21.5 ± 3.38 (Fig. 2). Another 100 fecal swab samples from two-week-old pigs with and without clinical signs were compared. The 50 samples from pigs with clinical signs had significantly lower Ct values (15.9 ± 0.59) compared to the 50 samples from non-clinical pigs (35.8 ± 0.71) ($p < 0.01$). All fecal samples from pigs 12 to 16 weeks old without clinical signs were negative with Ct values of equal to or higher than 40 (Fig. 2).

Determine the incidence of porcine sapovirus in swine herds of the United States

Totally, 234 fecal samples from pigs with diarrhea from different states across the United States (U.S.) were examined using SaV rRT-PCR. Among them, 106 samples (45.3%) were positive with Ct values of 26.0 ± 0.5; in contrast, 128 samples were negative with Ct values equal to or higher than 40. In addition, 102 fecal samples from clinically healthy pigs from different states across the United States were also examined with SaV rRT-PCR. Among them, 44 samples (43.1%) were positive by SaV rRT-PCR with Ct values of 33.2 ± 0.9, which were significantly higher than those from pigs with diarrhea (26.0 ± 0.5). The remaining 58 samples were negative by SaV rRT-PCR.

Sequence comparison, phylogenetic analysis, and recombination analysis

NGS analysis was performed with the fecal samples submitted to ISU VDL and totally 20

genome sequences from field porcine SaVs were assembled. The neighbor-joining (NJ)

phylogenetic tree based on genome sequences showed that all 20 SaVs were clustered with GIII

reference isolate sequences (Fig. 3). The genome sequences of the six GIII SaVs that were from

the outbreak farm showed 99.8% to 100% nucleotide (nt) identity, indicating that SaV is stable at

the farm level. Whereas, the genome sequences of the SaV from the outbreak farm showed only

82.0-98.0% nt identities to all other 14 GIII SaVs (Table S1).

SimPlot analysis was performed with SaV strains MW316746_IN17168-3A_2020,

MW316747_IL16624_2020, and MW316757_IA4517-2_2019, in the region between ORF1 55

nt and ORF2 522 nt, to determine recombination events within the genogroup GIII (Fig. 4). The

strain IN17168-3A showed higher nt identity (89.8%) to the IL16624 strain in the NS-VP1

coding region (ORF1 55-5235nt), but much lower nt identity (74.8%) to IA4517-2. While in the

VP2 coding region (ORF2 61-522nt), strain IN17168-3A showed higher nt identity (89.2%) to

the IA4517-2 strain, but lower nt identity (82.0%) to IL16624. These results suggested that the

strain IN17168-3A could be a recombinant from IL16624 and IA4517-2.

DISCUSSION

Porcine SaV is widely detected in pig herds throughout the world (Liu et al., 2012; Reuter et al.,

2010; Scheuer et al., 2013). Although experimental inoculation of isolated SaVs into gnotobiotic

pigs successfully induced enteritis (Flynn, Saif, and Moorhead, 1988; Guo et al., 2001), porcine

SaV is usually detected from asymptomatic pigs (Collins et al., 2009) or diarrheic piglets with

mixed infections with other viral or bacterial enteric pathogens (Dufkova et al., 2013; Mijovski

248   et al., 2010; Scheuer et al., 2013). In this study, a metagenomics analysis detected a porcine

249   sapovirus of genogroup III from fecal and intestine samples of piglets with diarrhea. PEDV,

250   PDCoV, TGEV, or rotavirus were not detected using rRT-PCR assays and additionally, there was

251   no significant bacterial pathogens isolated from the intestine. The newly developed rRT-PCR

252   revealed large amounts of porcine sapovirus RNA in the small intestine and the FISH assay

253   conducted on tissue with microscopic lesions consistent with a viral enteritis further confirmed

254   that GIII SaV was the cause of the observed clinical scenario. To the best of our knowledge, this

255   is the first confirmation that SaV can serve as a sole etiological agent of enteritis and diarrhea in

256   naturally infected piglets in the United States.

257   In this study, a significantly sensitive (5 copy numbers per reaction) and highly specific rRT-PCR

258   was successfully developed to detect porcine GIII sapovirus, the most prevalent genogroup in

259   swine worldwide. Examination of 184 fecal samples from pigs from the outbreak farm showed

260   that those pigs with clinical diarrhea had significantly lower Ct values (15.9 ± 0.59) than those

261   from clinically healthy pigs (35.8 ± 0.71) with 1024-fold difference in quantity. We further

262   investigated the prevalence of SaV in fecal samples of 234 pigs with diarrhea and 102 clinically

263   healthy pigs from different states across the United States. Our results showed that pigs with

264   clinical diarrhea had a comparable positive rate (45.3%) with non-clinical pigs (43.1%), which is

265   consistent with previous reports. However, the SaV-positive pigs with clinical diarrhea had

266   significantly higher virus loads (Ct=26.0 ± 0.5) compared to rRT-PCR positive but clinically

267   healthy pigs (Ct=33.2 ± 0.9). Most porcine SaVs have so far been reported from asymptomatic

268   pigs and several studies reported no significant differences in the positive rates of porcine SaV

269   between clinical and non-clinical pigs (Barry, Alfieri, and Alfieri, 2008; Collins et al., 2009; das

270   Merces Hernandez et al., 2014; Dufkova et al., 2013; Mijovski et al., 2010; Reuter et al., 2010).

271 Further, it has been suggested that testing for the presence of SaV may not provide significant

272 diagnostic value (Salamunova et al., 2018). Our results suggest that rRT-PCR is a reliable

273 method to detect porcine SaV and that porcine GIII sapovirus likely plays an important role in

274 suckling pig enteritis. This PCR provides an additional diagnostic tool that can be used in cases

275 of suckling pig diarrhea, particularly in those cases where other common agents are not detected.

276 The FISH probe developed herein is another highly specific tool that can be used for direct

277 detection of porcine SaV within intestinal lesions. It should be noted that other genogroups of

278 SaVs could not be excluded in the samples in our study since only porcine GIII SaV was

279 screened.

280 The age susceptibility of pigs to porcine SaV infection was investigated by different groups

281 (Barry, Alfieri, and Alfieri, 2008; Martinez et al., 2006; Valente et al., 2016). Post-weaning pigs

282 (4 to 8 weeks of age) have been suggested to have a higher positive frequency of SaV ($p<0.05$)

283 than grower-to-finish and breeder pigs (Valente et al., 2016) and the lowest prevalence of porcine

284 GIII SaVs was found in nursing pigs (Wang et al., 2006). In our study, 85.3% of suckling piglets

285 (8 to 12 days) were detected positive (Ct=15.9±0.59) with porcine SaV infection, a decent

286 percentage of pigs at 2-weeks-old were still positive (Ct=15.9), while none of pigs at 12-16

287 weeks-old were rRT-PCR positive. Our results suggest that the suckling and post-weaning pigs

288 are most susceptible to porcine SaV infection, which is consistent with the previous studies.

289 The genome sequences of 20 SaVs obtained in this study all belong to GIII, which is the

290 predominant genogroup of porcine sapovirus worldwide (Keum et al., 2009; L'Homme et al.,

291 2010; Nakamura et al., 2010; Wang et al., 2006; Yu et al., 2008). The genome sequences of some

292 GIII strains shared only 85.8% nt identities. Both intra- and inter-genogroup recombinant strains

293 have been reported for porcine SaV (Kuroda et al., 2017; Li et al., 2017; Wang et al., 2020; Wang

294   et al., 2005); however, most of the recombinant studies were performed with reference strains

295   and few recombination studies have been investigated among field strains. In the present study,

296   strain IN17168-3A shared high AA identity (98.7%) with strain IL16624 in the NS-VP1 region,

297   whereas a high AA identity (92.8%) with strain IA4517-2 in the VP2 region. The results

298   suggested that intra-genogroup recombination may occur between field GIII strains. To the best

299   of our knowledge, this is the first report of intra-genogroup recombination within GIII SaV

300   among field strains circulating in swine. NGS analysis also identified coinfections of several (up

301   to three) different SaV strains in the same fecal sample (data not shown), and coinfections of

302   different strains and/or different genogroups in the same pig could facilitate viral recombination

303   and thus increase genetic diversity.

304   In summary, using metagenomics analysis, GIII porcine SaV was detected from fecal and

305   intestine samples of piglets with enteritis that were negative for PEDV, PDCoV, TGEV, rotavirus,

306   or other bacterial enteric pathogens. Independent evidence of viral-induced microscopic lesions

307   and *in situ* hybridization confirmed that porcine SaV was the cause of the observed clinical

308   scenario, indicating it is the sole etiological agent. A highly sensitive and specific rRT-PCR

309   method was established for the detection of porcine GIII SaV. An epidemiological survey

310   demonstrated that the clinical pigs had significantly higher viral loads of SaV compared to the

311   non-clinical pigs. Further recombination analysis indicated that intra-genogroup recombination

312   occurs within the SaV GIII field isolates. Our findings provide significant insights for a better

313   understanding of the epidemiology and pathogenicity of GIII SaV in swine.

314

317

Conflict of Interest

The authors declare that they have no conflict of interest

320

Ethical Approval

Ethical statement is not applicable.

323

Data Availability Statement

All data generated or analyzed in this study are included in the article

326

Reference

Alhatlani, B., Vashist, S., and Goodfellow, I. (2015). Functions of the 5' and 3' ends of
     calicivirus genomes. *Virus Res* **206,** 134-43.

Barry, A. F., Alfieri, A. F., and Alfieri, A. A. (2008). High genetic diversity in RdRp gene of
     Brazilian porcine sapovirus strains. *Vet Microbiol* **131**(1-2)**,** 185-91.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina
     sequence data. *Bioinformatics* **30**(15)**,** 2114-2120.

Burrough, E. R., Wilberts, B. L., Bower, L. P., Jergens, A. E., and Schwartz, K. J. (2013).
     Fluorescent in situ hybridization for detection of "Brachyspira hampsonii" in porcine
     colonic tissues. *J Vet Diagn Invest* **25**(3)**,** 407-12.

Chen, Q., Wang, L., Zheng, Y., Zhang, J., Guo, B., Yoon, K. J., Gauger, P. C., Harmon, K. M.,
     Main, R. G., and Li, G. (2018). Metagenomic analysis of the RNA fraction of the fecal

339   virome indicates high diversity in pigs infected by porcine endemic diarrhea virus in the

340   United States. *Virol J* **15**(1)**,** 95.

341 Collins, P. J., Martella, V., Buonavoglia, C., and O'Shea, H. (2009). Detection and

342   characterization of porcine sapoviruses from asymptomatic animals in Irish farms. *Vet*

343   *Microbiol* **139**(1-2)**,** 176-82.

344 das Merces Hernandez, J., Stangarlin, D. C., Siqueira, J. A., de Souza Oliveira, D., Portal, T. M.,

345   Barry, A. F., Dias, F. A., de Matos, J. C., Mascarenhas, J. D., and Gabbay, Y. B. (2014).

346   Genetic diversity of porcine sapoviruses in pigs from the Amazon region of Brazil. *Arch*

347   *Virol* **159**(5)**,** 927-33.

348 Dufkova, L., Scigalkova, I., Moutelikova, R., Malenovska, H., and Prodelalova, J. (2013).

349   Genetic diversity of porcine sapoviruses, kobuviruses, and astroviruses in asymptomatic

350   pigs: an emerging new sapovirus GIII genotype. *Arch Virol* **158**(3)**,** 549-58.

351 Farkas, T., Zhong, W. M., Jing, Y., Huang, P. W., Espinosa, S. M., Martinez, N., Morrow, A. L.,

352   Ruiz-Palacios, G. M., Pickering, L. K., and Jiang, X. (2004). Genetic diversity among

353   sapoviruses. *Arch Virol* **149**(7)**,** 1309-23.

354 Flynn, W. T., Saif, L. J., and Moorhead, P. D. (1988). Pathogenesis of porcine enteric

355   calicivirus-like virus in four-day-old gnotobiotic pigs. *Am J Vet Res* **49**(6)**,** 819-25.

356 Guo, M., Hayes, J., Cho, K. O., Parwani, A. V., Lucas, L. M., and Saif, L. J. (2001).

357   Comparative pathogenesis of tissue culture-adapted and wild-type Cowden porcine

358   enteric calicivirus (PEC) in gnotobiotic pigs and induction of diarrhea by intravenous

359   inoculation of wild-type PEC. *J Virol* **75**(19)**,** 9239-51.

360 Keum, H. O., Moon, H. J., Park, S. J., Kim, H. K., Rho, S. M., and Park, B. K. (2009). Porcine

361   noroviruses and sapoviruses on Korean swine farms. *Arch Virol* **154**(11)**,** 1765-74.

362   Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics

363   Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**(7)**,** 1870-4.

364   Kuroda, M., Masuda, T., Ito, M., Naoi, Y., Doan, Y. H., Haga, K., Tsuchiaka, S., Kishimoto, M.,

365   Sano, K., Omatsu, T., Katayama, Y., Oba, M., Aoki, H., Ichimaru, T., Sunaga, F.,

366   Mukono, I., Yamasato, H., Shirai, J., Katayama, K., Mizutani, T., Oka, T., and Nagai, M.

367   (2017). Genetic diversity and intergenogroup recombination events of sapoviruses

368   detected from feces of pigs in Japan. *Infect Genet Evol* **55,** 209-217.

369   L'Homme, Y., Brassard, J., Ouardani, M., and Gagne, M. J. (2010). Characterization of novel

370   porcine sapoviruses. *Arch Virol* **155**(6)**,** 839-46.

371   Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler

372   transform. *Bioinformatics* **25**(14)**,** 1754-60.

373   Li, J., Shen, Q., Zhang, W., Zhao, T., Li, Y., Jiang, J., Yu, X., Guo, Z., Cui, L., and Hua, X.

374   (2017). Genomic organization and recombination analysis of a porcine sapovirus

375   identified from a piglet with diarrhea in China. *Virol J* **14**(1)**,** 57.

376   Liu, G. H., Li, R. C., Huang, Z. B., Yang, J., Xiao, C. T., Li, J., Li, M. X., Yan, Y. Q., and Yu,

377   X. L. (2012). RT-PCR test for detecting porcine sapovirus in weanling piglets in Hunan

378   Province, China. *Trop Anim Health Prod* **44**(7)**,** 1335-9.

379   Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G.,

380   Ingersoll, R., Sheppard, H. W., and Ray, S. C. (1999). Full-length human

381   immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India,

382   with evidence of intersubtype recombination. *J Virol* **73**(1)**,** 152-60.

383   Martinez, M. A., Alcala, A. C., Carruyo, G., Botero, L., Liprandi, F., and Ludert, J. E. (2006).

384   Molecular detection of porcine enteric caliciviruses in Venezuelan farms. *Vet Microbiol*

385   **116**(1-3)**,** 77-84.

386   Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for

387   metagenomics with Kaiju. *Nat Commun* **7,** 11257.

388   Mijovski, J. Z., Poljsak-Prijatelj, M., Steyer, A., Barlic-Maganja, D., and Koren, S. (2010).

389   Detection and molecular characterisation of noroviruses and sapoviruses in asymptomatic

390   swine and cattle in Slovenian farms. *Infect Genet Evol* **10**(3)**,** 413-20.

391   Nagai, M., Wang, Q., Oka, T., and Saif, L. J. (2020). Porcine sapoviruses: Pathogenesis,

392   epidemiology, genetic diversity, and diagnosis. *Virus Res* **286,** 198025.

393   Nakamura, K., Saga, Y., Iwai, M., Obara, M., Horimoto, E., Hasegawa, S., Kurata, T., Okumura,

394   H., Nagoshi, M., and Takizawa, T. (2010). Frequent detection of noroviruses and

395   sapoviruses in swine and high genetic diversity of porcine sapovirus in Japan during

396   Fiscal Year 2008. *J Clin Microbiol* **48**(4)**,** 1215-22.

397   Oka, T., Lu, Z., Phan, T., Delwart, E. L., Saif, L. J., and Wang, Q. (2016). Genetic

398   Characterization and Classification of Human and Animal Sapoviruses. *PLoS One* **11**(5)**,**

399   e0156373.

400   Oka, T., Wang, Q., Katayama, K., and Saif, L. J. (2015). Comprehensive review of human

401   sapoviruses. *Clin Microbiol Rev* **28**(1)**,** 32-53.

402   Ondov, B. D., Bergman, N. H., and Phillippy, A. M. (2011). Interactive metagenomic

403   visualization in a Web browser. *BMC Bioinformatics* **12,** 385.

404   Reuter, G., Zimsek-Mijovski, J., Poljsak-Prijatelj, M., Di Bartolo, I., Ruggeri, F. M., Kantala, T.,

405   Maunula, L., Kiss, I., Kecskemeti, S., Halaihel, N., Buesa, J., Johnsen, C., Hjulsager, C.

406   K., Larsen, L. E., Koopmans, M., and Bottiger, B. (2010). Incidence, diversity, and

407   molecular epidemiology of sapoviruses in swine across Europe. *J Clin Microbiol* **48**(2)**,**

408   363-8.

409   Saif, L. J., Bohl, E. H., Theil, K. W., Cross, R. F., and House, J. A. (1980). Rotavirus-like,

410   calicivirus-like, and 23-nm virus-like particles associated with diarrhea in young pigs. *J*

411   *Clin Microbiol* **12**(1)**,** 105-11.

412   Salamunova, S., Jackova, A., Mandelik, R., Novotny, J., Vlasakova, M., and Vilcek, S. (2018).

413   Molecular detection of enteric viruses and the genetic characterization of porcine

414   astroviruses and sapoviruses in domestic pigs from Slovakian farms. *BMC Vet Res* **14**(1)**,**

415   313.

416   Scheuer, K. A., Oka, T., Hoet, A. E., Gebreyes, W. A., Molla, B. Z., Saif, L. J., and Wang, Q.

417   (2013). Prevalence of porcine noroviruses, molecular characterization of emerging

418   porcine sapoviruses from finisher swine in the United States, and unified classification

419   scheme for sapoviruses. *J Clin Microbiol* **51**(7)**,** 2344-53.

420   Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., and Birol, I. (2009).

421   ABySS: a parallel assembler for short read sequence data. *Genome Res* **19**(6)**,** 1117-23.

422   Valente, C. S., Alfieri, A. F., Barry, A. F., Leme, R. A., Lorenzetti, E., and Alfieri, A. A. (2016).

423   Age distribution of porcine sapovirus asymptomatic infection and molecular evidence of

424   genogroups GIII and GIX? circulation in distinct Brazilian pig production systems. *Trop*

425   *Anim Health Prod* **48**(1)**,** 21-7.

426   Wang, L., Marthaler, D., Fredrickson, R., Gauger, P. C., Zhang, J., Burrough, E. R., Petznick, T.,

427   and Li, G. (2020). Genetically divergent porcine sapovirus identified in pigs, United

428   States. *Transbound Emerg Dis* **67**(1)**,** 18-28.

429   Wang, Q. H., Han, M. G., Funk, J. A., Bowman, G., Janies, D. A., and Saif, L. J. (2005). Genetic

430   diversity and recombination of porcine sapoviruses. *J Clin Microbiol* **43**(12)**,** 5963-72.

431   Wang, Q. H., Souza, M., Funk, J. A., Zhang, W., and Saif, L. J. (2006). Prevalence of

432   noroviruses and sapoviruses in swine of various ages determined by reverse transcription-

433   PCR and microwell hybridization assays. *J Clin Microbiol* **44**(6)**,** 2057-62.

434   Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification

435   using exact alignments. *Genome Biol* **15**(3)**,** R46.

436   Yinda, C. K., Conceicao-Neto, N., Zeller, M., Heylen, E., Maes, P., Ghogomu, S. M., Van Ranst,

437   M., and Matthijnssens, J. (2017). Novel highly divergent sapoviruses detected by

438   metagenomics analysis in straw-colored fruit bats in Cameroon. *Emerg Microbes Infect*

439   **6**(5)**,** e38.

440   Yu, J. N., Kim, M. Y., Kim, D. G., Kim, S. E., Lee, J. B., Park, S. Y., Song, C. S., Shin, H. C.,

441   Seo, K. H., and Choi, I. S. (2008). Prevalence of hepatitis E virus and sapovirus in post-

442   weaning pigs and identification of their genetic diversity. *Arch Virol* **153**(4)**,** 739-42.

443   Zhang, J., Zheng, Y., Xia, X. Q., Chen, Q., Bade, S. A., Yoon, K. J., Harmon, K. M., Gauger, P.

444   C., Main, R. G., and Li, G. (2017). High-throughput whole genome sequencing of

445   Porcine reproductive and respiratory syndrome virus from cell culture materials and

446   clinical specimens using next-generation sequencing technology. *J Vet Diagn Invest*

447   **29**(1)**,** 41-50.

448

449

450

451 **Table 1. Presence of viral pathogens in fecal samples of two NGS cases**

452

Case #	Sequence ID	Age	SaV GIII	PEDV*	PDCoV	TGEV	Rota A	Rota B	Rota C
1	MW316743_NE9550B	2 w	(+)	(-)	(-)	(-)	(-)	(-)	(-)
2	MW316759_IA0365	8-10 d	(+)	(-)	(-)	(-)	(-)	(-)	(-)

453 *, all viruses other than SaV GIII were tested by real-time PCR; (+), positive; (-), negative, i.e. Ct >= 36.

454

455

456   Table 2. Primers, probe and oligo sequences used in this study

Name	Sequence	Used for
PSapV-F	AACGCRGTGGCAACGTACAA	rRT-PCR
PSapV-R	GCCTCCATCACGAACACTTC	rRT-PCR
PSapV-P	FAM-TGGCTCYTCATCTTCATTGGTGGGAGC-TAMSp	rRT-PCR
Sap-GV-115	AATTACGAGCAGGCCACTGCCTGTTACAACAGCTGGTTTATTGGC GGGAGCCAACCTGAGGTGCCCACTACCAGTGAAGGCTATGGGCT ATTAGTGTTTGAGATGGAGGGCAATC	Specificity assay
Sap-GVI-112	GATTACACAGAGGCATTTGACGTCATGCGCAACTTGTCATCTACC CAACCCGATGGCGGGGCTGTTGTGTACACAATGGAGGGGCCCAA GCCCTCTCGTGCGTCGAATGGGT	Specificity assay
Sap-GVII-107	GAATACTCTGAGGCCATCGAGGTGATAACCAACATTTCATCAGGC ACGCCCGAGGGCGAGGCAATTGTGTATAAAATGGAGGGTCCTGA GGGGCCAAAGCCCACAGA	Specificity assay
Sap-GVIII-116	CACCAATTATGCTCAAGCTGCAGCCACCTACAATGCGTGGTATAT TGGTGGCGTGGAACCACAATTGGGGAGCCTCGCCAGTGAAGGTT CAGCTCAAGTAGTGTTTGAGATGGAGG	Specificity assay
Sap-GIX-117	CTGGTGTGGAATACTCAGAGGCAATTAACGTGCTAACATCTATTT CATCCAGACCGCCTGAGGGTGAGGCAATAGTGTATGTGATGGAG GGTCCAAACGGCCCTAAGGGCGCTCAGC	Specificity assay
Sap-GX-116	CTGGTGTGGAATACTCAGAGGCAATACACGTATTGACATCAATTT CATCCAAACCGCCCGAGGGCGAGGCAATAGTGTATGTGATGGAG GGCCCAAAGGGTCCCGTTGAGGGCGAG	Specificity assay
Sap-GXI-116	CTGGTGTGGAATACTCAGAGGCAATACACGTATTGACATCAATTT CATCCAAACCGCCCGAGGGCGAGGCAATAGTGTATGTGATGGAG GGCCCAAAGGGTCCCGTTGAGGGCGAG	Specificity assay
Psap5122probe	Alex555N-GCCTCCATCACGAACACTTCTGGCTC	FISH
Sap-T7-F1	TAATACGACTCACTATAGGGAACGCAGTGGCAACGTACAA	Create RNA standards
Sap5193R	TTGAGTACCCTCTGGGTTGCT	Create RNA standards

457

458

459

Table S1. Pairwise identities of SaV GIII strain MW316743_NE9550B from the outbreak farm compared with other strains

Sequence ID	Genome	NS		VP1		VP2	
	nt	CDS	AA	CDS	AA	CDS	AA
MK965904_NE9550	100.0	100.0	100.0	100	100.0	100.0	100.0
MW316752_NE7211-1	99.9	99.9	99.9	99.9	99.7	99.7	99.4
MW316753_NE7211-2	99.8	99.8	99.8	99.9	99.8	99.8	99.4
MW316754_NE7211-3	99.8	99.8	99.8	99.9	99.8	99.7	98.9
MW316755_NE7211-4	99.8	99.8	99.9	99.7	99.8	99.8	99.4
MW316759_IA0365	98.0	97.7	99.5	98.1	99.5	97.2	100
MW316747_IL16624	97.7	97.6	99.5	97.2	98.1	97.5	99.4
MW316758_IL6678-1	97.2	96.7	99.2	97.4	99.0	97.2	100.0
MK965905_MO13472	96.0	95.3	99.2	96.2	98.8	95.9	99.4
MW316749_MI64533	95.4	95.0	99.3	95.1	97.8	93.9	98.9
MW316751_KS66065	95.4	95.0	99.2	94.9	98.0	93.7	98.9
MW316750_NE67454	91.7	90.1	98.4	92.6	95.9	92.7	96.6
MK965902_IA27912-A	90.8	91.6	98.5	85.7	87.8	86.0	86.3
MW316748_OH19789	90.7	86.9	96.9	96.1	98.8	96.0	98.9
MW316744_MO1188	90.5	86.6	96.8	96.1	99.0	95.5	97.7
MW316745_IN17168-1	90.5	91.7	98.7	84.5	87.4	84.1	82.9
MW316746_IN17168-3A	90.3	91.4	98.6	84.3	87.8	84.8	84.0
MW316757_IA4517-2	82.1	78.0	86.2	85.5	85.6	84.6	83.4
MW316756_IA4517-1	82.0	78.1	85.6	84.3	85.4	85.8	88.0
KT922089_GIII	92.4	93.6	98.8	87.7	90.5	85.1	84.0
LC215878_GIII	92.0	90.3	98.0	93.1	96.6	93.1	94.9
KF204570_GIII	91.9	89.8	97.9	94.3	96.9	93.4	96.0
KX688107_GIII	91.8	90.2	98.0	92.7	96.8	92.9	97.7
KT922087_GIII	87.1	85.2	95.6	87.5	88.8	85.5	83.4
LC215881_GIII	86.1	84.4	95.0	85.2	85.7	85.1	85.1
FJ387164_GIII	85.8	85.6	96.1	88.4	90.3	89.9	92.6
AB521771_GV	58.1	53.1	41.1	56.4	43.8	44.5	29.1
AY974192_GVI	55.4	46.7	32.8	49.9	37.5	41.0	25.1
LC215890_GVII	55.3	47.0	31.0	48.5	37.0	38.8	29.1
LC215895_GVIII	52.6	53.3	40.7	54.8	40.7	40.0	25.7
KC309418_GIX	NA	NA	NA	48.4	38.2	48.6	32.0
LC215897_GX	50.1	47.6	31.6	48.9	41.8	41.2	29.1
LC215899_GXI	54.7	46.6	31.8	49.6	38.9	41.2	25.7

nt, nucleotides; CDS, coding sequence; AA, amino acid sequence; NA, not available in GenBank.

Figure legends

466 **Fig. 1.** Microscopic lesions and fluorescent in situ hybridization (FISH) in the small intestine of

467 pigs naturally infected with porcine sapovirus GIII. (A) There is marked villous atrophy and rare

468 fusion. HE. 40 ×. (B) Mild lymphocytic infiltration in the lamina propria. HE. 200 ×. (C)

469 Variable villous vascular congestion. HE. 200 ×. (D) Mildly vacuolated villus enterocytes that

470 are often detached and exposing the lamina propria. HE. 200 ×. (E) Scattered FISH-positive

471 villus enterocytes (in the red circle) are observed within sections with villus attenuation; FISH

472 was performed with Alexa555-labeled sapovirus probe; observed with CY3 Filter. 200 ×. (F) No

473 fluorescence was observed in the FISH-positive cells when observed with a FITC Filter. 200 ×.

474

475 **Fig. 2.** Sapovirus GIII rRT-PCR results from the outbreak farm. The line indicates the average Ct

476 value, and the scattered plots indicate the individual values.

477

478 **Fig. 3.** Phylogenetic analysis based on SaV genome sequences. The phylogenetic tree was

479 constructed by the neighbor-joining (NJ) method. Significant bootstrap values are indicated as a

480 percentage for 1000 replicates; bootstraps higher than 50 are displayed along the relative

481 branches. ▲, the SaVs from the outbreak farm; ◆, the recombinant SaV strain; ■ and ■,

482 potential parent strains involved in recombination; ●, other strains obtained in this study.

483

484 **Figure 4.** Recombination analysis. SimPlot analysis for the region from ORF1 55 nt to ORF2

485 522 nt of the nucleotide sequences: MW316747_IL16624_2020 (green line),

486 MW316757_IA4517-2_2019 (blue line), and MW316746_IN17168-3A_2020 as the query

487 sequence. The position indicates the breakpoints, 5175 nt (ORF1 5235 nt) and 6768 nt (ORF2 61

488 nt), in the sequence alignment.