

1 **Whole genome sequencing of porcine reproductive and respiratory syndrome virus**
2 **(PRRSV) from field clinical samples improves the genomic surveillance of the virus**

3
4
5 Christian Lalonde¹, Chantale Provost^{1,2}, Carl A. Gagnon^{1,2}

6
7
8 ¹Swine and Poultry Infectious Diseases Research Center (CRIPA) and ²Molecular diagnostic
9 laboratory (MDL); Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe,
10 Québec, Canada.

11
12
13 Corresponding author: Carl A. Gagnon, Swine and Poultry Infectious Diseases Research Center
14 (CRIPA), Faculté de médecine vétérinaire (FMV), Université de Montréal (UdeM), 3200 rue
15 Sicotte, office 3963, St-Hyacinthe, Québec, Canada, J2S 2M2. carl.a.gagnon@umontreal.ca .
16 Phone: 450-773-8521 (ext: 8681).

17
18
19 Running title: Incidence of PRRSV coinfection and recombination by WGS

20
21 Keywords: animal viral disease, swine virus, porcine reproductive and respiratory syndrome
22 virus, PRRSV, next-generation sequencing, NGS, whole genome sequencing, WGS,
23 recombinant, coinfection, classification

24 **Abstract**

25 Porcine reproductive and respiratory syndrome virus (PRRSV) is a major economic concern
26 worldwide. There are currently large data sets available about the ORF5 gene of the virus, with
27 thousands of sequences available, but little data is currently available on the full-length genome
28 of PRRSV. We hypothesized that whole genome sequencing (WGS) of PRRSV genome would
29 allow a better epidemiological monitoring compared to ORF5 gene sequencing. PRRSV PCR
30 positive sera, oral fluids and tissue clinical samples submitted to the diagnostic laboratory for
31 routine surveillance or diagnosis of PRRSV infection in Québec, Canada, swine herds were used.
32 The PRRSV RT-qPCR Cq values of the processed samples varied between 11.5 and 34.34.
33 PRRSV strain genomes were isolated using a poly(A)-tail method and were sequenced with an
34 MiSeq Illumina sequencer. Ninety-two full length PRRSV genomes were obtained from 88
35 clinical samples, out of 132 tested samples, resulting in a PRRSV WGS success rate of 66.67%.
36 Three important deletions in the ORF1a were found in most wildtype (i.e. not vaccine-like)
37 strains. The importance of these deletions remains undetermined. Two different full-length
38 PRRSV genomes were found in four different samples (three sera and one pool of tissues),
39 suggesting a 4.55% PRRSV strain coinfection prevalence in swine. Moreover, six PRRSV whole
40 genomes (6.52% of PRRSV strains) were found to cluster differently compared to ORF5
41 classification method. Overall, WGS of PRRSV enables a better strains classification and/or
42 interpretation of results in 9.10% of clinical samples compared to ORF5 sequencing, as well as
43 allowing interesting research avenues.

44

45

46 **Introduction**

47

48 Porcine reproductive and respiratory syndrome (PRRS) is a major economic concern
49 worldwide, costing 663 million dollars yearly to the American swine industry(1) and over 150
50 million dollars yearly to the Canadian swine industry(2). Part of this cost is due to reproductive
51 troubles in sow (characterized by late abortion, increased incidence of stillbirth, mummified
52 fetuses and weaker newborn piglets)(3-6) and retarded growth in piglets, mainly caused by
53 respiratory problems (characterized by interstitial pneumonia)(5-10) and increased animal
54 susceptibility to other pathogens(11-13). Beside this direct loss in productivity, a lot of efforts
55 and resources are allocated to control and contain this infectious disease. Among those efforts,
56 vaccination and epidemiological surveillance are the most prominent (14, 15).

57

58 The etiological agents of PRRS are one of two viruses of the *Arteriviridae* family and the
59 *Porartevirus* genus that are aptly named porcine reproductive and respiratory syndrome virus
60 (PRRSV) type 1 and type 2 (officially *Betaarterivirus suis 1* and *Betaarterivirus suis 2*). Both
61 PRRSV are enveloped viruses with a diameter of around 50-65 nm(16). They contain a single
62 stranded positive-sense RNA genome, of around 15 kb in length, capped at the 5' end and
63 polyadenylated at the 3' end and contains at least 11 known open reading frames (ORF)(17, 18).
64 The first two ORFs (ORF1a and ORF1b), which constitute about 75% of the viral genomes and
65 contain at least three ribosome shifting sites, encodes replicase polyproteins that are post-
66 translationally cleaved into at least 16 distinct non-structural proteins (nsp) (18-20). The
67 remaining 25% of the genomes code for at least 8 known structural proteins (GP2, GP3, GP4,

68 GP5, M, N, E and GP5a). Both PRRSV species share a nucleotide identity of around 60% (22,
69 23).

70 Many genotypes of both PRRSV exist and the virulence of each strain is highly variable.
71 PRRSV strains found in Canadian swine herds belong to type 2 PRRSV. To our knowledge,
72 indigenous PRRSV type 1 strains have never been reported in Canada but was found on one
73 occasion in European imported piglets kept in quarantine. Nine distinct monophyletic lineages of
74 PRRSV type 2 have been described worldwide. These lineages are genetically distinct, with a
75 nucleotide (nt) identity between lineage under 89%. Of these 9 lineages, lineage 1, 2, 5 and 8 are
76 the most prominent in Canada, with lineage 9 showing up sporadically. The vast majority of
77 lineages 5 and 8 strains are thought to be vaccine related. These lineages have been determined
78 using Bayesian phylogenies of the ORF5 gene encoding the membrane structural glycoprotein
79 protein GP5 of the virion. Interestingly, recent studies have investigated the genomic diversity
80 and relatedness of PRRSV Canadian strains, but all those reports are using PRRSV ORF5
81 nucleotide sequences into their genomic analyses. However, there is much less information and
82 data available about the whole viral genome of PRRSV strains both in Canada and around the
83 globe.

84

85 The inter-strain genetic variability of PRRSV is very high. It is believed that this enables
86 the virus to better evade the immune system and potentially diminish vaccine efficacy. In an
87 ongoing effort in Quebec, Canada, as many PRRSV ORF5 sequences as possible are added to an
88 ever-growing database, with over 4695 ORF5 sequences to date. The goal of this initiative was
89 to better understand the epidemiological links between various PRRSV strains versus outbreaks
90 and determine the proximity of a given strain to commercially available vaccinal strains as a

91 strategy to fight PRRS disease. This database is also used to find genetically similar strains,
92 enabling veterinarians to better predict the severity and outcome of a current outbreak based on
93 the severity and outcome of past outbreaks with ORF5 genetically similar strains. The Molecular
94 diagnostic laboratory of the Diagnostic service (Faculté de médecine vétérinaire of Université de
95 Montréal), is the depository of that Quebec ORF5 sequences databank. The PRRSV ORF5 gene
96 was selected for the molecular epidemiological surveillance because it was considered
97 hypervariable and encodes the GP5 protein that acts as the main target of neutralizing antibodies.

98

99 However, the ORF5 is only 603 nucleotides long, representing only 4% of the viral
100 genome and it has been shown that other genes are also hypervariable and that pathogenicity of
101 PRRSV is determined by multiple genes. The GP5 antigenicity property (i.e. its recognition by
102 neutralizing antibodies) has contributed to the selection of ORF5 for PRRSV surveillance despite
103 the fact that other viral proteins are involved in the virion recognition by neutralizing
104 antibodies(40). Given the high prevalence of PRRSV infections, concomitant infections with
105 different strains are expected to occur quite frequently as it is known to happen already in other
106 swine respiratory viruses. It was previously reported that PRRSV viral recombination occurs
107 prominently during coinfection with two different strains. All of these facts suggest that in some
108 clinical cases of PRRSV infections, PRRSV recombinant strains could be misclassified, affecting
109 the interpretation of the data at hand and the subsequent intervention by veterinarian
110 practitioners. We thus hypothesized that whole genome sequencing of PRRSV strains from
111 clinical samples would enable a better classification of PRRSV strains compared to the current
112 surveillance method of ORF5 sequencing. Therefore, whole genome sequencing of PRRSV
113 strains could lead to more appropriate interventions by veterinarians and swine producers, in

114 addition to improving our understanding of the pathogenicity and the epidemiology of this
115 important swine pathogen.

116

117 **Material and Methods**

118

119 **Swine samples.** Convenience swine samples that tested positive for PRRSV type 2 by a RT-
120 qPCR diagnostic assay were selected for whole genome sequencing (WGS) of the entire PRRSV
121 viral genome. The samples were submitted to and tested by the «Molecular diagnostic
122 laboratory (MDL) of Faculté de médecine vétérinaire (FMV), Université de Montréal (UdeM)»
123 for the identification of PRRSV after an outbreak of the disease in swine herds and to a lesser
124 extent to conduct surveillance of the virus in swine herds. The PRRSV RT-qPCR diagnostic
125 assay conducted by MDL (own the American Association of Veterinary Laboratory
126 Diagnosticians, AAVLD, accreditation) was an in-house assay (protocol: PON-MOL-029). The
127 swine samples originated from different herds and types of production throughout the province
128 of Quebec, Canada. Those clinical samples were submitted between December 2015 and
129 November 2018 (most of them were collected in 2017 [n=62] and 2018 [n=66]). A total of 132
130 samples were used and included: 70 sera (composed of 43 pooled sera), 2 oral fluids (OF), 32
131 lungs and 28 pools of tissues (PoT; mainly lungs with several other types of tissues such as
132 lymph nodes, spleen, liver, intestine, etc.), with PRRSV RT-qPCR diagnostic assay Cq values
133 between 11.5 and 34.34. The PRRSV viral genome was considered complete if at least 98% of
134 the coding sequence was obtained.

135

136 **Genome extraction and purification.** One hundred mg of lung or PoT were grinded using a
137 Beadbeater apparatus (model: Mini-Beadbeater-96; BioSpec Products Inc., OK, USA) in PBS,
138 then centrifuged at full speed for one minute and supernatant was use for viral extraction. Two
139 hundred ul of sera and OF were centrifuged for 5 minutes at 10 000 g and supernatant was used

140 for viral extraction. Viral RNA was extracted using Quick-RNA Viral Kit (#R1035, Zymo
141 Research, CA, USA) as described in the company's protocol. Thereafter, RNA was eluted using
142 50 ul of nuclease free water (Corning, NY, USA). Total elution volume was used to isolate RNA
143 with poly(A)-tails with a magnetic beads purification method, the NEBNext® Poly(A) mRNA
144 Magnetic Isolation Module, as described by the company's protocol, and poly(A)-tails RNA was
145 resuspended in 15 ul Tris Buffer (#E7490; New England BioLabs, ON, Canada). Then, first
146 strand cDNA was synthesised using the Non-directional Reaction Step Up protocol of the
147 NEBNext® RNA First Strand Synthesis Module (#E7525) (New England BioLabs, ON,
148 Canada), starting with 10 ul of isolated poly(A)-tails RNA. Immediately after first strand cDNA
149 synthesis, the second DNA strand was synthesized using NEBNext® Ultra II Non-Directional
150 RNA Second Strand Synthesis Module, as described by the manufacturer protocol with a minor
151 modification at the incubation step in the thermal cycler for 2 hours (instead of 1.5 hours) at
152 16°C (#E61111; New England BioLabs, ON, Canada). dsDNA was then purified using AxyPrep
153 Mag™ PCR Clean-up Kits (Axygen – Corning, NY, USA) using 1.8X of beads and 70%
154 ethanol. The purified dsDNA was diluted in 30 ul of 10mM Tris-HCl at pH 8.0 and stored at -
155 20°C for later use.

156

157 **PRRSV whole genome sequencing.** dsDNA was quantified using Qubit™ dsDNA HS Assay
158 Kit and a Qubit™ fluorometer (ThermoFisher Scientific, MA, USA). Sequencing libraries were
159 prepared using Nextera XT DNA Library Preparation Kit (Illumina, CA, USA). Sequencing
160 libraries construction were still performed even if the dsDNA quantification results were below
161 the Qubit™ dsDNA HS Assay Kit threshold of detection (0.2 ng). Briefly, 0.2 to 0.3 ng of
162 dsDNA were tagged with 10 µl Tagment DNA Buffer (TD) and 5 µl Amplicon Tagment Mix

163 (ATM) at 55°C for 5 minutes, using a Thermocycler TProfessional Basic 96 (Biometra GmbH,
164 37079 Göttingen, Germany). The transposomes were inactivated with 5 µl of Neutralize
165 Tagment Buffer (NT). Sequencing libraries were then amplified using index adapter and Nextera
166 PCR Master Mix (NMP) following those PCR steps: 72°C for 3 minutes, 95°C for 30 seconds;
167 then 14 cycles of 95°C for 10 seconds, 55°C for 30 seconds, 72°C for 30 seconds; a final
168 elongation step at 72°C for 5 minutes; and a hold at 10°C until the next step. Sequencing
169 libraries were then purified using AxyPrep MagTM PCR Clean-up Kits (Axygen – Corning, NY,
170 USA) as described by the Nextera XT protocol. Sequencing libraries quality were assessed using
171 Agilent High Sensitivity DNA Kit with a Bioanalyzer (Agilent, CA, USA). Sequencing libraries
172 were normalised using LNB1 beads (Nextera XT protocol) or the manual normalisation protocol
173 if the concentrations of the libraries were on the lower end of the Bioanalyzer curves.
174 Sequencing libraries were sequenced in a v3 600-cycle cartridge (#MS-102-3003; Illumina)
175 using an MiSeq Illumina benchtop sequencer (located at the “Veterinary High Throughput
176 Sequencing Laboratory (VHTSL)”, at FMV) and PhiX was included at around 1% of the total
177 sequencing libraries, as a control to establish the sequencing run efficacy (Illumina, CA, USA).

178

179 **Bioinformatic analyses.** At first, reads were trimmed for adaptors and quality by the MiSeq
180 software during FastQ generation. Using CLC Genomic Workbench software (version 12.0.3,
181 Qiagen, CA, USA), reads from a sample were assembled using the software’s built-in *de novo*
182 assembly module, including another step of adaptor and quality trimming. Contigs obtained were
183 checked against a database made using reference type 2 PRRSV genomes available online to
184 identify the contigs that corresponded to full-length PRRSV sequences. These full-length
185 genomes obtained by *de novo* assembly were then added to our PRRSV reference database.

186 Secondly, reads from each sample were mapped against the same database. For each sample, the
187 reference genome onto which the highest number of reads mapped and offered a uniform
188 coverage was used as the reference genome for resequencing analysis using the software's map
189 to reference module. When more than one reference genome seemed to offer an appropriate
190 depth and coverage, both were used separately for resequencing analysis, and the resulting
191 sequences compared to one another. Samples from which two PRRSV genomes could be
192 assembled and possessed a nucleotide (nt) identity of $< 96\%$ were considered to contain more
193 than one PRRSV strain. The $< 96\%$ nt identity cut-off value was determined based on the
194 sequencing error rate of $< 1\%$ and the fact that we have ambiguity in less than 1% of the
195 positions for each complete sequence. Thus, the worst case scenario where both PRRSV
196 coinfecting strains could each have 1% error rate and 1% uncertainty could lead to identical
197 strains with a nt identity of $\geq 96\%$. Therefore, this nt identity value was established as the cut-off
198 value. PRRSV recombinant status was determined by comparing the location of a given strain's
199 individual ORFs on their respective phylogenetic trees. Bioinformatics and data analyses were
200 performed using the CLC Genomic Workbench software (version 12.0.3, Qiagen, CA, USA).
201 Phylogenetic trees were constructed using a maximum likelihood model with a type 1 PRRSV
202 reference strain (Lelystad virus) used as an outgroup and with a bootstrap setting of 1000. The
203 trees were constructed with the Geneious Prime software (Biomatter, version 2019).

204

205 **Accession Numbers**

206 All entire viral genomes of successfully sequenced PRRSV strains were submitted to GenBank
207 databank, with accession numbers MN865482 to MN865573 (Suppl. Table 1).

208

209 **Results**

210 PRRSV WGS were obtained from a total of 88 swine clinical samples (an overall 66.67%
211 PRRSV WGS success rate). In those clinical samples, the PRRSV WGS was successful in two
212 OF, 13 PoT, 21 lungs and 52 sera. WGS made a difference over ORF5 sequencing in eight of
213 those samples (Table 1). Ninety-two PRRSV entire genomes were obtained from those clinical
214 samples (Suppl. Table 1). The additional 4 PRRSV genomes were the consequences of the
215 identification of 4 PRRSV coinfections, samples in which two different PRRSV strains with a nt
216 identity below 96% were identified (listed in Table 1). Therefore, a total of 4.55% PRRSV
217 coinfections were found among PRRSV WGS successful cases. The inter-strains nt identities of
218 the 92 PRRSV entire viral genome sequences were found to vary between 79.57% and 99.94%
219 (data not shown). At the viral gene level, the nsp2 and nsp7 encoded genes nt identities were
220 found to be the least conserved between the 92 PRRSV Quebec strains, with 66.04% and 64.38%
221 nt identities, respectively, indicating that ORF1a genetic diversity seemingly a little higher
222 compared to the others ORF. Noteworthy, the overall genetic diversity was scattered throughout
223 all the ORFs (Fig. 1, Table 2). The highest nt identities were found for the nsp4, E, N and M
224 gene encoded proteins with 92.41%, 86.94%, 86.13% and 86.10%, respectively.

225

226 Three deletions were identified by WGS in the ORF1a of most wild-type strains (Fig. 1).
227 The first one, located in the nsp2 coding region at nt positions 2159 to 2491 and is 333 nt long.
228 The second one, also in the nsp2 coding region, was found around nt positions 2717 to 2773 with
229 a span of 57 nt. Both deletions occurred together and were found in 63 of the 92 PRRSV Québec
230 sequenced strains. The third deletion was found in the nsp7 coding region, at position 7106 to
231 7228 and was 123 nt long. The nt position of the deletions are given relative to the coding

232 sequence of the MLV vaccine strain used as a reference. All the strains that harbor the nsp7
233 deletion also have the nsp2 deletions but some strains that do not possess the nsp7 deletion have
234 the nsp2 deletions. Only 47 of the 92 sequenced strains have the nsp7 deletions.

235

236 Based on ORF5 sequencing classification method, all 92 PRRSV strains were found to
237 belong to lineage 1, 5 and 8 (Fig 2b). Interestingly, 18 PRRSV strains were considered to be
238 MLV vaccine-like strains (i.e. with nt identities > 96%) if the tree was built using ORF5 while
239 17 strains were considered to be MLV vaccine-like strains while using the whole genome.
240 Noteworthy, one discrepancy was found in regard to MLV vaccine-like strain relatedness
241 classification. In fact, PRRSV strain 2072533 was classified as a MLV vaccine-like strain based
242 on the ORF5 method whereas it was not genetically related to MLV vaccine-like strain based on
243 WGS results, with nt identities of 98.51% and 84.01%, respectively (Fig. 2). Moreover, PRRSV
244 strain 2072533 was identified to be a recombinant virus which explains the discrepancy between
245 the ORF5 and WGS classification methods (Table 1). No difference was found for the ATP
246 vaccine-like and Foster vaccine-like classification status (Fig 2a compared to 2b).
247 Consequently, misclassification by ORF5 sequencing of PRRSV vaccine-like strains was
248 revealed by WGS in 3.23% of samples.

249

250 As illustrated in Figure 2, PRRSV whole genome sequences phylogenetic tree has
251 allowed to construct a more robust classification, based on the bootstrap values, compared to the
252 ORF5 sequencing method. More interestingly, six PRRSV strains (identified by colored arrows
253 in Fig. 2) were found to cluster differently within the WGS phylogenetic tree compared to the
254 ORF5 phylogenetic tree. A summary of those classification differences is presented in Table 1.

255 One of the reasons that could explain this difference is recombination events, at some point in
256 time, between two PRRSV strains during coinfection in swine. As an example of this
257 phenomenon, the 2072533 PRRSV strain nucleotide identity with a presumed parental PRRSV
258 strain (the MLV vaccine strain) is shown at Figure 3 in which the recombination position was
259 identified around nt position 12,100. The nt identity between these two PRRSV strains for the
260 first 12100 nt of the 5' end (i.e. encompassing ORF1a, ORF1b, ORF2a and ORF2, is 80.46%
261 whereas it is 98.63% (data not shown) for the 3' end region of the genome (the last 3,000 nt of
262 the genome). For strain 1952821 #1, the first half of the genome (mainly the ORF1a) shares a
263 85.53% nt identity with the MLV vaccine strain, whereas the second half (mostly ORF1b to
264 ORF7) shares a 96.55% nt identity with the MLV vaccine strain (data not shown), suggesting
265 again a recombination between a wild-type and vaccine-like strain. For the other possible
266 recombinant strains, the specific recombination point could not be determined, either because the
267 parental strains are unknown or because those strains may have undergone several recombination
268 events over time. Overall, 6.52% possible recombinant viruses over a total of 92 PRRSV strains
269 were found among PRRSV WGS successful cases. Noteworthy, 2 of the 6 possible recombinant
270 viruses were found in PRRSV coinfecting samples. This is however to be expected given that
271 coinfection is a prerequisite for recombination events to occur, taking into account that more
272 than one third of produced virions during coinfections can be PRRSV recombinant viruses(42).

273

274 In four of the 88 successfully sequenced PRRSV positive samples, two different PRRSV
275 strains were found (Table 1), representing 4.55% coinfections among PRRSV WGS successful
276 cases as indicated above. The nt identities of these four pairs of PRRSV were 88.8%, 92.36%
277 91.57% and 81.83% for samples 1890826, 1952821, 1968868 and 2153073, respectively (data

278 not shown). A graph of the nucleotide conservation status of the PRRSV strains in sample
279 2153073 is shown in Figure 4 as an example. Interestingly, the 2153073 #2 strain possesses the
280 nsp2 and nsp7 deletions like Quebec wildtype strains whereas it's coinfecting strain, 2153073
281 #1, is more related to ATP vaccine-like strains (sharing a 99.02% nt identity). All coinfections
282 were identified from one individual animal (i.e. sample coming from a single animal and not
283 from a pool of different animals). Interestingly, coinfections were identified in three sera samples
284 (5.77% of WGS successful sera) and one PoT sample (7.14% of WGS successful PoT) (Suppl.
285 Table 1 and Table 1).

286 **Discussion**

287 The first observation that can be easily made from our type 2 PRRSV complete genome
288 analyses is that many of the PRRSV strains currently circulating in Quebec appear to have three
289 relatively large deletions (Fig. 1). The first two of these occur in the region coding for the nsp2
290 protein and the third one in the region coding for the nsp7 protein. While the exact implication of
291 these deletions is unknown, several investigations have been conducted on type 2 PRRSV strains
292 possessing various nsp2 deletions, especially with Asian PRRSV strains. The nsp2 protein is
293 known to play a role in both immune system modulation and pathogenicity. For example, it is
294 believed that some amino acids in the nsp2 protein could play a role in evading the ubiquitin-
295 ISG15 system of the innate immune response. There are also many reports of highly pathogenic
296 or epidemic strains recognizable by nsp2 deletions. The importance of nsp2 in the pathogenicity
297 of PRRSV is what makes these deletions so interesting, and why they warrant further
298 investigations. Noteworthy, these specific nsp2 deletions can also be found in some PRRSV
299 strains that were reported in the United States (*i.e.*: GenBank #KT258000 and #EF532803),
300 meaning that they are not markers of a specific North American country but are probably
301 specific to lineage 1 or at least to a subset of lineage 1 PRRSV strains. The involvement of the
302 third deletion, in the region coding for the nsp7 protein into the pathogenicity of PRRSV, is
303 much more elusive. PRRSV pathogenicity studies involving nsp7 deletions and mutations
304 experiment are scarce, and usually involve pinpoint mutations or very short deletions(50).
305 However, at least one study has reported that nucleotide mutations at some nsp7 specific site
306 could lead to changes in PRRSV replication efficiency *in vitro*. Noteworthy, to our knowledge,
307 this specific nsp7 deletion does not seem to have been previously reported by other investigators.
308 Therefore, this deletion certainly requires further investigation.

309

310 Overall, the whole genome diversity of PRRSV type 2 strains in Quebec is quite high, as
311 expected based on ORF5 diversity, with up to 20.43% difference in nucleotide identities between
312 strains. This diversity is higher in regions coding for proteins considered more or less accessory
313 and lower for those considered essential, like the polymerase nsp9 or the helicase nsp10, as well
314 as the structural proteins (Table 2). Interestingly, the most conserved nt encoded protein gene
315 was found to be the serine protease nsp4 (Table 2), which is the main PRRSV protease
316 responsible to cleave most of the replicase polyprotein into it's smaller functional parts(52).

317

318 Type 2 PRRSV WGS classification stays similar for most PRRSV strains compared to
319 ORF5 classification method which includes 9 lineages. Although, few PRRSV strains appeared
320 to be of an uncertain lineage when looking at the whole genome sequences. This could be
321 because WGS may have improve the depth of the genomic analyses compared to ORF5 method
322 (i.e. 15,000 nt based analyses compared to 603 nt based analyses, respectively). Two examples
323 of this are the strains 2108698 and 2137004. While both clustered inside the lineage 8 based on
324 ORF5 phylogenetic tree, they could just as easily be grouped with the reference strain MN30100
325 of lineage 9 based on whole genome sequencing. Another example of this phenomenon is the
326 reference strain IAF-klop. We know it belongs to lineage 1, but the whole genome seems to
327 cluster more closely to the lineages 5-9. Another possible explanation of this is that this specific
328 strain has been passaged repeatedly *in vitro* in MARC-145 cell line. Lastly, based on whole
329 genome sequencing, both strains from the 1890896 sample seem to belong to an uncertain
330 lineage. The limited availability of PRRSV whole genome sequences data and the relatively
331 frequent occurrence of viral genome recombination events raise the need to come up with a

332 genomic classification system that would take into account the higher variability of some non-
333 structural proteins and the incidence of viral genome recombination events. Such recombination
334 events pose a real challenge to the current monophyletic lineage classification system.

335 A more adequate classification of the strains could be made using both the nsp2 and
336 ORF5 gene sequences in a manner akin to influenza subtyping(54). The nsp2 coding region
337 seems to routinely contain deletions which can serve as markers of sub-lineage as well as being
338 implicated in pathogenicity, making it a good candidate for typing of different strains (45, 46,
339 49). WGS remains however the method of choice compared to nsp2/ORF5 sequencing for a few
340 reasons. Firstly, the nsp2 coding-region is over 3000 nt long and highly variable, making it's
341 amplification by PCR for Sanger sequencing quite difficult. Secondly, while adding a second
342 gene from the 5' end of the genome would greatly improve odds of catching recombinants,
343 simply doing a second Sanger sequencing instead of whole genome sequencing would not detect
344 coinfections in swine samples.

345 The 4.55% occurrence of PRRSV coinfections amongst PRRSV WGS positive cases was
346 in line with what has been previously reported for swine influenza virus (SIV). In fact, more than
347 one subtype (such as H1N1 and H3N2) of SIV are found in at least 2% of the tested samples, not
348 accounting at all for coinfections with strains from the same subtype(41).

349 The six possible recombinant PRRSV strains (a prevalence of 6.52%) that cluster
350 differently, based on WGS analyses, are also of high importance. In the field, veterinarian
351 practitioners and swine producers are mostly interested in knowing if a strain within a PRRSV
352 positive sample is genetically related to a PRRSV vaccine-like or wild-type strains, and if it is
353 related to previous infections on that farm or to other known outbreaks. When a PRRSV strain
354 changes clustering with WGS, due to a possible recombination event, this can have an important

355 impact on the clinical interpretation as it can change its relatedness to a known high or low
356 pathogenic strain.

357 From a clinical point of view, the 2072533 strain would have been identified by
358 veterinarian practitioners, as a MLV vaccine-like strain based on the ORF5 gene method whereas
359 in fact, it is not a MLV vaccine-like strain based on WGS. Overall, PRRSV WGS has changed
360 the vaccine-like strain status of one PRRSV strain, i.e. 3.23% of the vaccine-like related cases
361 (data not shown).

362 While these analyses have been conducted on Quebec PRRSV strains, there is nothing to
363 suggest that these findings could not extend to most if not all regions currently affected by
364 PRRSV type 2 outbreaks as both intra and inter-lineage recombination events have been
365 described and as coinfections can also occur with type 2 PRRSV strains from the same or
366 different lineages. These types of recombination events are known to have occurred elsewhere.

367 The main limit of our current study is that since Canada is exempt of type 1 PRRSV, we
368 cannot say if these finding also extends to type 1 PRRSV strains. While type 2 PRRSV is the
369 most frequent PRRSV found in both China and the US, type 1 is also known to occur and data on
370 the latest is currently limited(57). It is also currently unknown if recombination between both
371 viral species causing PRRS is possible in a clinical setting. As both species are now becoming
372 endemic to regions where the other one used to predominate, this could be something to
373 investigate in the future. It should be noted however that the lack of type 1 PRRSV strains do not
374 seem to have hindered the Quebec diversity of type 2 PRRSV strains.

375 To our knowledge, it is the first time that a study has investigated the entire viral genome
376 of several Canadian PRRSV type 2 strains. In addition to allowing the description of 3
377 substantial deletions in the ORF1a region of many contemporary lineage 1 strains, WGS appears

378 to enable a better interpretation of clinical cases in about 9.10% of PRRSV cases due to either
379 PRRSV coinfecting strains or the presence of PRRSV recombinant strains. Again, to our
380 knowledge, it is the first time that both the proportion of misclassifications and the incidence of
381 coinfections of type 2 PRRSV strains using whole genome sequencing data have been quantified.
382 In those cases, this additional information (that more than 1 strain was present in a sample or that
383 the strain is not related to the same strains than those we would have thought with ORF5
384 sequencing method) could lead directly to more appropriate preventive and palliative measures
385 being applied by the veterinarians and swine producers, which could have a positive impact in
386 the overall health of swine herds around the world.

387

388 **Acknowledgments**

389

390 This research was financially supported by the Ministère de l’Agriculture, des Pêcheries
391 et de l’Alimentation du Québec (MAPAQ) Innov’Action program. Carl A. Gagnon was
392 financially supported by a Natural Sciences and Engineering Research Council of Canada
393 (NSERC) discovery grant and a Canadian Swine Research and Development Cluster (CSRDC)
394 grant. Christian Lalonde was a recipient of a scholarship from the CRIPA, a research network
395 financially supported by the Fonds de recherche du Québec – Nature et technologies (FRQNT).
396 The authors declared no potential conflicts of interest with respect to the research, authorship,
397 and/or publication of this article.

398

399 **References**

- 400 1. Holtkamp DJ, Kliebenstein JB, Neumann E, Zimmerman JJ, Rotto H, Yoder TK, Wang C,
401 Yeske P, Mowrer CL, Haley CA. 2013. Assessment of the economic impact of porcine
402 reproductive and respiratory syndrome virus on United States pork producers. *Journal of*
403 *Swine Health and Production* 21:72.
- 404 2. Mussel A. 2011. A Risk, Benefit, Strength, Weakness, Opportunity and Threat Analysis
405 for the Control and Possible Eradication of Porcine Reproductive and Respiratory
406 Syndrome (PRRS) Virus Within the Canadian Swine Herd. *Canadian Swine Health*
407 *Board*:1-122.
- 408 3. Christianson WT. 1992. Stillbirths, mummies, abortions, and early embryonic death.
409 *Veterinary Clinics of North America: Food Animal Practice* 8:623-639.
- 410 4. Pol J, van Dijk J, Wensvoort G. 1991. Pathological, ultrastructural, and changes caused by
411 Lelystad virus in experimentally induced infections of mystery swine disease (synonym
412 porcine epidemic abortion and respiratory syndrome (PEARS)). *Vet Q* 13:137-143.
- 413 5. Loula T. 1991. Mystery pig disease. *Agri-practice* 12:23-34.
- 414 6. Bilodeau R, Dea S, Sauvageau R, Martineau G. 1991. 'Porcine reproductive and respiratory
415 syndrome' in Quebec. *The Veterinary Record* 129:102.
- 416 7. Albina E. 1997. [Porcine reproductive and respiratory syndrome: ten years of experience
417 (1986-1996) with this undesirable viral infection]. *Vet Res* 28:305-52.
- 418 8. Collins JE, Benfield DA, Christianson WT, Harris L, Hennings JC, Shaw DP, Goyal SM,
419 McCullough S, Morrison RB, Joo HS. 1992. Isolation of swine infertility and respiratory
420 syndrome virus (isolate ATCC VR-2332) in North America and experimental reproduction
421 of the disease in gnotobiotic pigs. *Journal of Veterinary Diagnostic Investigation* 4:117-
422 126.
- 423 9. Halbur P, Paul P, Frey M, Landgraf J, Eernisse K, Meng X-J, Andrews J, Lum M, Rathje J.
424 1996. Comparison of the antigen distribution of two US porcine reproductive and
425 respiratory syndrome virus isolates with that of the Lelystad virus. *Veterinary Pathology*
426 33:159-170.
- 427 10. Rossow KD, Morrison RB, Goyal SM, Singh GS, Collins JE. 1994. Lymph node lesions in
428 neonatal pigs congenitally exposed to porcine reproductive and respiratory syndrome virus.
429 *Journal of veterinary diagnostic investigation* 6:368-371.
- 430 11. Harms P, Sorden S, Halbur P, Bolin S, Lager K, Morozov I, Paul P. 2001. Experimental
431 reproduction of severe disease in CD/CD pigs coinfecting with PRRSV and type 2 porcine
432 circovirus.
- 433 12. Galina L, Pijoan C, Sitjar M, Christianson W, Rossow K, Collins J. 1994. Interaction
434 between *Streptococcus suis* serotype 2 and porcine reproductive and respiratory syndrome
435 virus in specific pathogen-free piglets. *The Veterinary Record* 134:60-64.
- 436 13. Cho JG, Dee SA, Deen J, Trincado C, Fano E, Jiang Y, Faaberg K, Murtaugh MP, Guedes
437 A, Collins JE. 2006. The impact of animal age, bacterial coinfection, and isolate
438 pathogenicity on the shedding of porcine reproductive and respiratory syndrome virus in

- 439 aerosols from experimentally infected pigs. *Canadian Journal of Veterinary Research*
440 70:297.
- 441 14. Linhares DC, Johnson C, Morrison RB. 2015. Economic analysis of vaccination strategies
442 for PRRS control. *PLoS one* 10:e0144265.
- 443 15. Nathues H, Alarcon P, Rushton J, Jolie R, Fiebig K, Jimenez M, Geurts V, Nathues C.
444 2017. Cost of porcine reproductive and respiratory syndrome virus at individual farm
445 level—an economic disease model. *Preventive veterinary medicine* 142:16-29.
- 446 16. Spilman MS, Welbon C, Nelson E, Dokland T. 2009. Cryo-electron tomography of porcine
447 reproductive and respiratory syndrome virus: organization of the nucleocapsid. *Journal of*
448 *general virology* 90:527-535.
- 449 17. Johnson CR, Griggs TF, Gnanandarajah J, Murtaugh MP. 2011. Novel structural protein in
450 porcine reproductive and respiratory syndrome virus encoded by an alternative ORF5
451 present in all arteriviruses. *Journal of General Virology* 92:1107-1116.
- 452 18. Meulenbergh JJ. 2000. PRRSV, the virus. *Veterinary research* 31:11-21.
- 453 19. Cao QM, Subramaniam S, Ni Y-Y, Cao D, Meng X-J. 2016. The non-structural protein
454 Nsp2TF of porcine reproductive and respiratory syndrome virus down-regulates the
455 expression of Swine Leukocyte Antigen class I. *Virology* 491:115-124.
- 456 20. Li Y, Treffers EE, Naphthine S, Tas A, Zhu L, Sun Z, Bell S, Mark BL, van Veelen PA, van
457 Hemert MJ. 2014. Transactivation of programmed ribosomal frameshifting by a viral
458 protein. *Proceedings of the National Academy of Sciences* 111:E2172-E2181.
- 459 21. Wu W-H, Fang Y, Rowland RR, Lawson SR, Christopher-Hennings J, Yoon K-J, Nelson
460 EA. 2005. The 2b protein as a minor structural component of PRRSV. *Virus research*
461 114:177-181.
- 462 22. Nelsen CJ, Murtaugh MP, Faaberg KS. 1999. Porcine reproductive and respiratory
463 syndrome virus comparison: divergent evolution on two continents. *Journal of virology*
464 73:270-280.
- 465 23. Murtaugh MP, Elam M, Kakach L. 1995. Comparison of the structural protein coding
466 sequences of the VR-2332 and Lelystad virus strains of the PRRS virus. *Archives of*
467 *virology* 140:1451-1460.
- 468 24. Allende R, Lewis T, Lu Z, Rock D, Kutish G, Ali A, Doster A, Osorio F. 1999. North
469 American and European porcine reproductive and respiratory syndrome viruses differ in
470 non-structural protein coding regions. *Journal of General Virology* 80:307-315.
- 471 25. Drew TW, Meulenbergh JJ, Sands JJ, Paton DJ. 1995. Production, characterization and
472 reactivity of monoclonal antibodies to porcine reproductive and respiratory syndrome
473 virus. *Journal of General Virology* 76:1361-1369.
- 474 26. Mardassi H, Mounir S, Dea S. 1994. Identification of major differences in the nucleocapsid
475 protein genes of a Quebec strain and European strains of porcine reproductive and
476 respiratory syndrome virus. *Journal of General Virology* 75:681-685.
- 477 27. Shi M, Lam TT-Y, Hon C-C, Murtaugh MP, Davies PR, Hui RK-H, Li J, Wong LT-W,
478 Yip C-W, Jiang J-W. 2010. Phylogeny-based evolutionary, demographical, and

- 479 geographical dissection of North American type 2 porcine reproductive and respiratory
480 syndrome viruses. *Journal of virology* 84:8700-8711.
- 481 28. Lambert M-È, Delisle B, Arsenault J, Poljak Z, D'Allaire S. 2019. Positioning Quebec
482 ORF5 sequences of porcine reproductive and respiratory syndrome virus (PRRSV) within
483 Canada and worldwide diversity. *Infection, Genetics and Evolution* 74:103999.
- 484 29. Shi M, Lemey P, Brar MS, Suchard MA, Murtaugh MP, Carman S, D'Allaire S, Delisle B,
485 Lambert M-È, Gagnon CA. 2013. The spread of type 2 porcine reproductive and
486 respiratory syndrome virus (PRRSV) in North America: a phylogeographic approach.
487 *Virology* 447:146-154.
- 488 30. Dewey C, Charbonneau G, Carman S, Hamel A, Nayar G, Friendship R, Eernisse K,
489 Swenson S. 2000. Lelystad-like strain of porcine reproductive and respiratory syndrome
490 virus (PRRSV) identified in Canadian swine. *The Canadian Veterinary Journal* 41:493.
- 491 31. Murtaugh MP, Stadejek T, Abrahante JE, Lam TT, Leung FC-C. 2010. The ever-
492 expanding diversity of porcine reproductive and respiratory syndrome virus. *Virus research*
493 154:18-30.
- 494 32. Petry D, Holl J, Weber J, Doster AR, Osorio FA, Johnson R. 2005. Biological responses to
495 porcine respiratory and reproductive syndrome virus in pigs of two genetic populations.
496 *Journal of animal science* 83:1494-1502.
- 497 33. Petry D, Lunney J, Boyd P, Kuhar D, Blankenship E, Johnson R. 2007. Differential
498 immunity in pigs with high and low responses to porcine reproductive and respiratory
499 syndrome virus infection. *Journal of Animal Science* 85:2075-2092.
- 500 34. Kim W-I, Lee D-S, Johnson W, Roof M, Cha S-H, Yoon K-J. 2007. Effect of genotypic
501 and biotypic differences among PRRS viruses on the serologic assessment of pigs for virus
502 infection. *Veterinary microbiology* 123:1-14.
- 503 35. Pirzadeh B, Gagnon CA, Dea S. 1998. Genomic and antigenic variations of porcine
504 reproductive and respiratory syndrome virus major envelope GP5 glycoprotein. *Canadian*
505 *Journal of Veterinary Research* 62:170.
- 506 36. Pirzadeh B, Dea S. 1998. Immune response in pigs vaccinated with plasmid DNA
507 encoding ORF5 of porcine reproductive and respiratory syndrome virus. *Journal of*
508 *General Virology* 79:989-999.
- 509 37. Gonin P, Pirzadeh B, Gagnon CA, Dea S. 1999. Seroneutralization of porcine reproductive
510 and respiratory syndrome virus correlates with antibody response to the GP5 major
511 envelope glycoprotein. *Journal of Veterinary Diagnostic Investigation* 11:20-26.
- 512 38. Han J, Rutherford MS, Faaberg KS. 2009. The porcine reproductive and respiratory
513 syndrome virus nsp2 cysteine protease domain possesses both trans- and cis-cleavage
514 activities. *Journal of virology* 83:9449-9463.
- 515 39. Music N, Gagnon CA. 2010. The role of porcine reproductive and respiratory syndrome
516 (PRRS) virus structural and non-structural proteins in virus pathogenesis. *Animal health*
517 *research reviews* 11:135-163.

- 518 40. Chen Z, Collin E, Peddireddi L, Clement T, Gauger P, Hause BM. 2017. Genetic diversity
519 in envelope genes of contemporary US porcine reproductive and respiratory syndrome
520 virus strains influences viral antigenicity. *Research in veterinary science* 115:432-441.
- 521 41. Choi Y, Goyal S, Kang S, Farnham M, Joo H. 2002. Detection and subtyping of swine
522 influenza H1N1, H1N2 and H3N2 viruses in clinical samples using two multiplex RT-PCR
523 assays. *Journal of virological methods* 102:53-59.
- 524 42. Liu D, Zhou R, Zhang J, Zhou L, Jiang Q, Guo X, Ge X, Yang H. 2011. Recombination
525 analyses between two strains of porcine reproductive and respiratory syndrome virus in
526 vivo. *Virus research* 155:473-486.
- 527 43. Yuan S, Nelsen CJ, Murtaugh MP, Schmitt BJ, Faaberg KS. 1999. Recombination between
528 North American strains of porcine reproductive and respiratory syndrome virus. *Virus
529 research* 61:87-98.
- 530 44. Zhou L, Zhang J, Zeng J, Yin S, Li Y, Zheng L, Guo X, Ge X, Yang H. 2009. The 30-
531 amino-acid deletion in the Nsp2 of highly pathogenic porcine reproductive and respiratory
532 syndrome virus emerging in China is not related to its virulence. *Journal of virology*
533 83:5156-5167.
- 534 45. Choi H-W, Nam E, Lee YJ, Noh Y-H, Lee S-C, Yoon I-J, Kim H-S, Kang S-Y, Choi Y-K,
535 Lee C. 2014. Genomic analysis and pathogenic characteristics of type 2 porcine
536 reproductive and respiratory syndrome virus nsp2 deletion strains isolated in Korea.
537 *Veterinary microbiology* 170:232-245.
- 538 46. Kim D-Y, Kaiser TJ, Horlen K, Keith ML, Taylor LP, Jolie R, Calvert JG, Rowland RR.
539 2009. Insertion and deletion in a non-essential region of the nonstructural protein 2 (nsp2)
540 of porcine reproductive and respiratory syndrome (PRRS) virus: effects on virulence and
541 immunogenicity. *Virus genes* 38:118-128.
- 542 47. Frias-Staheli N, Giannakopoulos NV, Kikkert M, Taylor SL, Bridgen A, Paragas J, Richt
543 JA, Rowland RR, Schmaljohn CS, Lenschow DJ. 2007. Ovarian tumor domain-containing
544 viral proteases evade ubiquitin-and ISG15-dependent innate immune responses. *Cell host
545 & microbe* 2:404-416.
- 546 48. Sun Z, Li Y, Ransburgh R, Snijder EJ, Fang Y. 2012. Nonstructural protein 2 of porcine
547 reproductive and respiratory syndrome virus inhibits the antiviral function of interferon-
548 stimulated gene 15. *Journal of virology* 86:3839-3850.
- 549 49. Yu Lx, Wang X, Yu H, Jiang Yf, Gao F, Tong W, Li Lw, Li Hc, Yang S, Chen Pf. 2018.
550 The emergence of a highly pathogenic porcine reproductive and respiratory syndrome virus
551 with additional 120aa deletion in Nsp2 region in Jiangxi, China. *Transboundary and
552 emerging diseases* 65:1740-1748.
- 553 50. Zhang M, Cao Z, Xie J, Zhu W, Zhou P, Gu H, Sun L, Su S, Zhang G. 2013. Mutagenesis
554 analysis of porcine reproductive and respiratory syndrome virus nonstructural protein 7.
555 *Virus genes* 47:467-477.
- 556 51. Fang Y, Snijder EJ. 2010. The PRRSV replicase: exploring the multifunctionality of an
557 intriguing set of nonstructural proteins. *Virus research* 154:61-76.

- 558 52. van Dinten LC, Rensen S, Gorbalenya AE, Snijder EJ. 1999. Proteolytic processing of the
559 open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine
560 protease and is essential for virus replication. *Journal of virology* 73:2027-2037.
- 561 53. Reyes YH, Provost C, Traesel CK, Jacques M, Gagnon CA. 2018. Actinobacillus
562 pleuropneumoniae culture supernatant antiviral effect against porcine reproductive and
563 respiratory syndrome virus occurs prior to the viral genome replication and transcription
564 through actin depolymerization. *Journal of medical microbiology* 67:249-264.
- 565 54. Memorandums M-L. 1980. A revision of the system of nomenclature for influenza viruses:
566 a WHO memorandum. *Bulletin of the World Health Organization* 58:585-591.
- 567 55. Wang A, Chen Q, Wang L, Madson D, Harmon K, Gauger P, Zhang J, Li G. 2019.
568 Recombination between Vaccine and Field Strains of Porcine Reproductive and
569 Respiratory Syndrome Virus. *Emerging infectious diseases* 25:2335.
- 570 56. Zhang X, Li Y, Xiao S, Yang X, Chen X, Wu P, Song J, Ma Z, Cai Z, Jiang M. 2019.
571 High-frequency mutation and recombination are responsible for the emergence of novel
572 porcine reproductive and respiratory syndrome virus in northwest China. *Archives of*
573 *virology* 164:2725-2733.
- 574 57. Wang A, Zhang J, Shen H, Zheng Y, Feng Q, Yim-Im W, Gauger PC, Harmon K, Zhu S,
575 An T-Q. 2019. Genetic diversity of porcine reproductive and respiratory syndrome virus 1
576 in the United States of America from 2010 to 2018. *Veterinary microbiology* 239:108486.
- 577

578 **Tables**

579 **Table 1: List of clinical samples containing more than one PRRSV strain (coinfection) and**
 580 **of strains that are classified differently by WGS compared to ORF5 method.**

Sample id number	Type of sample	Viruses whole genome sequencing characteristics	Impact on PRRSV classification
1898026	Serum	Coinfection ^a	Both strains clustered in different lineages
1952821	Pool of Tissues	Coinfection	Both strains clustered differently in the same lineage
1968868	Serum	Coinfection	Both strains clustered differently in the same lineage
2153073	Serum	Coinfection	Both strains clustered in different lineages
1425619	Lung	Recombinant ^b	Strain clustered differently in the same lineage compared to ORF5 method
1890826 #2 ^c	Serum	Recombinant	Strain clustered in a different lineage compared to ORF5 method
1927781	Serum	Recombinant	Strain clustered differently in the same lineage compared to ORF5 method
1952821 #1	Pool of Tissues	Recombinant	Strain clustered in a different lineage compared to ORF5 method
2072533	Serum	Recombinant	Strain clustered in a different lineage compared to ORF5 method
2087409	Lung	Recombinant	Strain clustered differently in the same lineage compared to ORF5 method

^aAt least two PRRSV different strains were identified in the same clinical sample and were possessing a nucleotide identity of < 96%.

^bA PRRSV recombinant status was assigned when a strain clustered more then two nodes away in phylogenetic trees based on PRRSV whole genome analysis compared to ORF5 classification method.

^cThe # after the sample identification (id) number represent one of the two PRRSV strains that was identified within the same clinical sample.

582 **Table 2: Nucleotide conservation status of the main PRRSV protein encoded genes.** The
 583 lowest inter-strains nucleotide identities between major PRRSV proteins are indicated.
 584

Corresponding proteins	Lowest inter-strain identities ^a	Corresponding proteins	Lowest inter-strain identities
nsp1	76.15%	nsp11	84.01%
nsp2	66.04%	nsp12	82.47%
nsp3	80.09%	GP2	84.70%
nsp4	92.41%	E	86.94%
nsp5	78.82%	GP3	82.35%
nsp6	77.08%	GP4	84.73%
nsp7	64.38%	GP5a	82.05%
nsp8	82.61%	GP5	82.95%
nsp9	85.75%	M	86.10%
nsp10	84.50%	N	86.13%

^aThe highest nucleotide identities obtained between strains were $\geq 99.8\%$ for all protein encoded genes.

585

586

587 **Figures**

588 **Figure 1: Nucleotide conservation status between Québec PRRSV strains.** All 92 PRRSV
589 sequenced strains are depicted here with the main ORFs shown. Three important and recurrent
590 deletions in some of the strains are indicated with an *. The scale to the left of the bar indicates
591 the percentage of strains that shared identity at a given location, with a full blue bar signifying
592 that all strains are identical at this position. The numbers on top of the blue bar indicate the
593 nucleotide position.

594

595

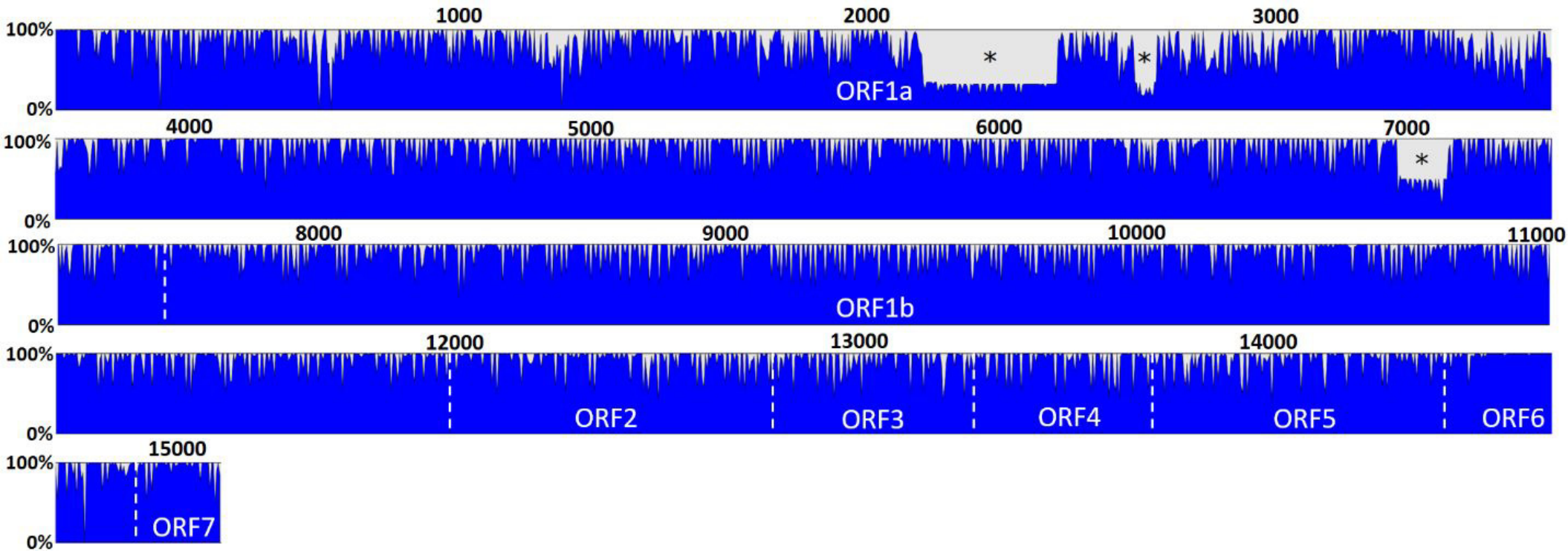
596 **Figure 2: PRRSV whole genome and ORF5 gene nucleotide phylogenetic trees comparison.**

597 A) PRRSV WGS nucleotide phylogenetic tree and B) PRRSV ORF5 nucleotide phylogenetic
598 tree. The brackets indicate the PRRSV type 2 lineages. MLV, ATP, Foster, PrimePac and
599 Prevacent corresponds to commercially available live attenuated vaccines. Strains 97-7895,
600 MN30100, VR2385 and MN184 are American reference strains. Strain PA8 is the first Canadian
601 PRRSV whole genome sequenced strain and strain IAF-Klop is a Québec reference strain. The
602 branches for the Lelystad virus strain (type 1 PRRSV reference strain which was used as an
603 outgroup) should be three times longer but were shortened to fit the graph. PRRSV coinfecting
604 strains, i.e. two different strains that were identified in the same clinical samples, are indicated
605 with a #1 or #2. Colored arrows identify potentially PRRSV recombinant strains that are
606 clustering differently when WGS and ORF5 phylogenetic trees are compared. The scale bars
607 represent a 5% difference in nucleotide identity.

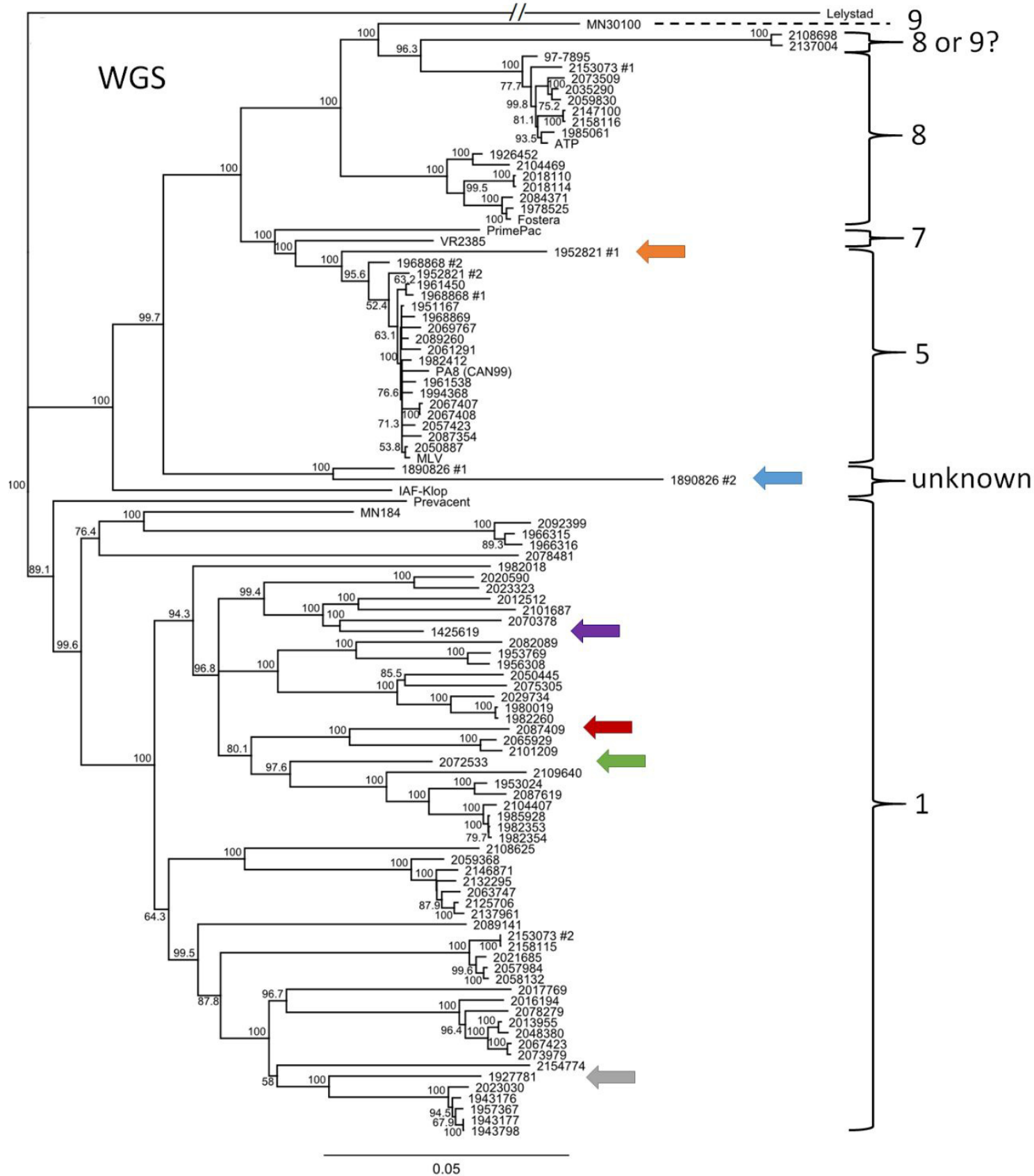
608

609 **Figure 3: Comparison of the PRRSV recombinant strain 2072533 with one of its possible**
610 **parental strain, the MLV vaccine strain.** The red arrow points to the region of the genome
611 where the recombination probably occurred. A full blue line indicates that both strains are
612 identical in that given nucleotide position and gray/white indicate that they differ at that position.
613 The numbers on top of the blue bar indicate the nucleotide position.
614

615 **Figure 4: Nucleotide conservation status of two PRRSV coinfecting strains found in sample**
616 **2153073.** Nucleotide identity between the two strains found in sample 2153073. A full blue line
617 indicates that both strains are identical in that given nucleotide position and gray/white indicate
618 that they differ at that position. The nucleotide identity between both PRRSV strains is 81.83%.
619 The numbers on top of the blue bar indicate the nucleotide position.



A



B