Infection, Genetics and Evolution 74 (2019) 103999

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Positioning Quebec ORF5 sequences of porcine reproductive and respiratory syndrome virus (PRRSV) within Canada and worldwide diversity

Marie-Ève Lambert[a,b,*], Benjamin Delisle[a,b], Julie Arsenault[a,b], Zvonimir Poljak[c], Sylvie D'Allaire[a,b]

[a] Laboratoire d'épidémiologie et de médecine porcine (LEMP), Faculty of Veterinary Medicine, Université de Montréal, St. Hyacinthe, Quebec, Canada
[b] Swine and Poultry Infectious Diseases Research Center (CRIPA), Faculty of Veterinary Medicine, Université de Montréal, St. Hyacinthe, Quebec, Canada
[c] Department of Population Medicine, Ontario Veterinary College, University of Guelph, Ontario, Canada

ARTICLE INFO

Keywords:
PRRS
Sequence
ORF5
Classification
Phylogeny
Genetic diversity
RFLP

ABSTRACT

Sequencing of ORF5 gene is widely used and considered essential for diagnostics and control of porcine reproductive and respiratory syndrome (PRRS) in Canada. The objective of this study was to position Quebec ORF5 sequences of PRRS virus within Canada and worldwide diversity. Overall, 76.8% of the 5204 sequences gathered from Quebec ($n = 5031$), Ontario ($n = 151$) and Manitoba ($n = 18$) were classified into one of 34 genetic clusters defined as groupings including ≥15 sequences and having ≥70% rapid bootstrap support value from a maximum likelihood (ML)-phylogeny. Following the addition of PRRSV 2 international reference dataset from Shi et al. (2010), the most predominant lineages in our dataset were wild-type 1 and vaccine-like 5.1 (MLV) and 8.9 (ATP). No strains or only a very few (1 or 2) were assigned to lineages 1.3–1.5, 3, 4, 5.2, 6, 7 or 9. Most wild-type clusters (97%) detected in a dataset from Canada did not include any sequence from the international reference dataset. It might reflect recent subpopulations that were absent at the time of Shi's publication. As an example, cluster #25 first appeared in 2007, but since then had expanded considerably and is now the most prevalent wild-type cluster found in Quebec. A total of 117 RFLP patterns were identified and those were poorly correlated with genetic clusters based on phylogeny. Factors modulating PRRSV diversity such as pig movement that occurred within and between provinces should be further investigated in a perspective of disease control.

1. Introduction

Porcine reproductive and respiratory syndrome (PRRS) is endemically present in most swine-producing countries. The causative agent is a single-stranded RNA enveloped virus of the Nidovirales order, *Arteriviridae* family (Adams et al., 2017). It is composed of 15,000 base pairs (bp) encoding at least ten functional open reading frames (ORFs) (Rahe and Murtaugh, 2017). The important heterogeneity observed among PRRS virus (PRRSV) strains, combined with the absence of complete protection following heterologous challenge, complicate disease management. Indeed, swine herds are always at risk of developing a clinical episode of the disease following the introduction of a new virus strain (Mateu and Diaz, 2008). The risk of introduction is influenced by many direct and indirect pathways of transmission of PRRSV between herds, including the introduction of infected animals or semen, transport vehicles, aerosols, flying insects, waterfowl and fomites (Albina, 1997; Arruda et al., 2015). High swine density as well as a multi-site production system increase the potential of virus spreading (Lambert et al., 2012b).

ORF5 sequencing of the virus is largely used by field veterinarians to first confirm the presence of a wild-type strain in the herd by opposition to vaccine-like strain, commercial vaccination being available and frequently used. Sequencing can also be useful to indicate a new viral introduction rather than recirculation of an endemic strain. Both scenarios can result in increased clinical signs, but would require either external or internal biosecurity enhancement at the herd level (Lambert et al., 2012b). Since 1998, sequences from the province of Quebec, Canada, are obtained from diagnostic testing and gathered by the Laboratoire d'épidémiologie et de médecine porcine (LEMP) of the Faculty of Veterinary Medicine of the Université de Montréal. Through the

Abbreviations: PRRSV, Porcine reproductive and respiratory syndrome virus; ORF, Open reading frame; ML, Maximum likelihood; RBS, Rapid bootstrap support; LEMP, Laboratoire d'épidémiologie et de médecine porcine; RFLP, Restriction fragment length polymorphism

* Corresponding author at: Laboratoire d'épidémiologie et de médecine porcine (LEMP), Faculty of Veterinary Medicine, Université de Montréal, St. Hyacinthe, Quebec, Canada.
E-mail address: marie-eve.lambert@umontreal.ca (M.-È. Lambert).

https://doi.org/10.1016/j.meegid.2019.103999
Received 21 March 2019; Received in revised form 6 August 2019; Accepted 8 August 2019
Available online 10 August 2019

years, the database (LEMP-DB) was used in research studies on PRRSV molecular epidemiology (Lambert et al., 2012a; Larochelle et al., 2003). Phylogenetic analyses were also performed to evaluate PRRSV diversity on data collected up to 2009 ($n = 1301$) and revealed several subgroups (Delisle et al., 2012). Since this latter study, approximately 4000 sequences were added to the database. Sequencing activities were also performed in other provinces of Canada. A study on 505 ORF5 sequences from Ontario obtained between 1999 and 2010 reported a diversity comparable to the United States (Brar et al., 2011). RFLP patterns, combined or not to phylogeny, were also frequently used in Ontario to identify particular genetic clusters (Arruda et al., 2015; Rosendal et al., 2014).

From a worldwide standpoint, the genetic diversity of PRRS strains can be divided into two species, 1 and 2 (Adams et al., 2017; Murtaugh et al., 2010). In Canada, only PRRSV 2 has been reported (Brar et al., 2011; Delisle et al., 2012; Rosendal et al., 2014). In 2010, a significant effort was performed to classify worldwide PRRSV 2 strains obtained from different countries, revealing nine principal lineages (Shi et al., 2010). Many lineages were absent from Canada at that time, but we cannot preclude a potential introduction since then. Moreover, since pig movements occurred both within and between provinces, monitoring viral subpopulations is particularly relevant. Further phylogenetic analyses are therefore necessary for an updated portrait of Canadian strains.

The objective of this study was to describe ORF5 PRRSV strain diversity found in Quebec, Ontario and Manitoba and more specifically to 1) identify genetic clusters based on phylogeny, 2) document RFLP patterns found among those clusters, and 3) assess the presence of previously reported worldwide lineages.

2. Materials and methods

2.1. Sequence collection

2.1.1. Quebec sequences

ORF5 PRRSV sequences were obtained from the LEMP-DB. All sequences from the province of Quebec, Canada, collected between January 1, 1998, and December 31, 2016 ($n = 5031$) through field submissions from regular veterinary services, active surveillance (i.e. area regional control and elimination (ARC&E) initiatives) or LEMP research activities were extracted. Field samples were submitted by veterinarians to the diagnostic laboratory of the Faculty of Veterinary Medicine (FVM) of the Université de Montréal or either of two other private laboratories. RNA extraction, reverse transcriptase PCR (RT-PCR) and ORF5 sequencing were performed according to the routine protocol of each laboratory.

2.1.2. Sequences from other Canadian provinces

Twenty swine veterinarians from across Canada, outside the province of Quebec, were invited by e-mail or phone to provide ORF5 sequences from field submissions along with sampling date and province identification. On December 31, 2016, a total of 169 sequences between 2005 and 2016 were obtained from Ontario ($n = 151$) and Manitoba ($n = 18$).

2.1.3. International sequences

A reference dataset of 841 sequences, with their respective lineage as described within a worldwide study on PRRSV 2, was obtained directly from the authors (Shi et al., 2010; Shi et al., 2013).

2.1.4. Reference vaccine sequences

An ORF5 sequence for four different PRRS commercial vaccines was obtained from GenBank: Ingelvac PRRS® MLV (Boehringer Ingelheim Vetmedica Inc., St. Joseph, Missouri, USA), Ingelvac PRRS® ATP (Boehringer Ingelheim Vetmedica Inc., St. Joseph, Missouri, USA), Fostera® PRRS (Zoetis, Florham Park, New Jersey, USA) and Prime Pac™ PRRS + (Merck Animal Health, Summit, NJ).

2.2. Classifying Canadian sequences

2.2.1. Phylogenetic analyses

All sequences from each dataset were examined for the presence of unusual characters other than known IUB symbols and for length. When present, unusual characters were replaced by N character. Sequences having ≥594 bp were selected for further analyses. Quebec and other Canadian provinces datasets as well as reference vaccine strains were merged, thereby forming a Canadian dataset of 5204 sequences for classification.

The dataset was sorted in increasing order of sampling date and a multiple alignment was performed using Clustal Omega with default settings (Sievers et al., 2011). Clustal Omega was chosen based on its efficiency at minimizing the number of gaps in dataset of different sizes with default open gap values as well as its capability to handle a large number of sequences in a timely manner (Lambert et al., 2019b).

A maximum likelihood (ML) phylogeny was inferred using RAxML (Pthreads AVX version 8.2.8, https://sco.h-its.org/exelixis/resource/download/NewManual.pdf) based on a GTR GAMMA evolutionary model. The software retains for its analysis only the first sequence listed among a particular group of sequences having all 100% similarity. A total of 1000 rapid bootstrap (RBS) were computed to determine branch confidence level. A mid-pointed root was placed on the tree for visual representation. Information regarding the tree branching pattern and RBS values were exported into a Newick partition file and data were organized into a relational database. Genetic clusters, which were defined as a group of ≥15 sequences linked to a branch of the tree with RBS value of ≥70%, were extracted using a Python script querying the database. When several clusters were embedded into each other, only the largest one or closest to the mid-point root (major cluster), was considered. More details regarding the automated clustering approach are available in a previously published paper (Lambert et al., 2019a). Clusters revealed were pictured using Archaeopteryx software on an unrooted ML-tree. Finally, 100% similar sequence(s) with at least one other sequence initially dropped to build the tree topology were added to the corresponding cluster to perform descriptive statistics using SAS version 9.4 software (SAS Institute Inc., Cary, NC, USA). Each cluster was then described according to the number of sequences, time (year of the first and last sequences submitted) and province of submission.

2.2.2. RFLP patterns

The RFLP patterns were determined using the following enzymes: *Mlu*I, *Hinc*II and *Sac*II (Wesley et al., 1998). Cut sites by the three enzymes were visually assessed and corrected when needed due to possible variation in length of sequences associated with insertion or deletion. Cut patterns were determined accordingly with the most up-to-date list of RFLP patterns kindly provided by the Minnesota Veterinary Diagnostic Laboratory of the University of Minnesota. The number of RFLP patterns and the most predominant pattern were described for each genetic cluster previously identified (Section 2.2.1).

2.3. Assessing the presence of worldwide lineages in Canada

Out of the 841 sequences of the international dataset, 771 sequences having ≥594 bp were selected and added to the Canadian dataset ($n = 5204$), the 70 sequences of < 594 bp were removed. Sequences from the merged dataset were first classified as described in Section 2.2.1. When at least one sequence from the international reference dataset fell into one of the genetic clusters identified by the system, then the whole cluster was associated to the same lineage/sublineage. When international reference sequences fell into an unclassified group, the presence of these lineages and sublineages in Canada was determined using a mid-pointed ML-tree built in Archaeopteryx and the common node linking all reference sequences from a specific lineage

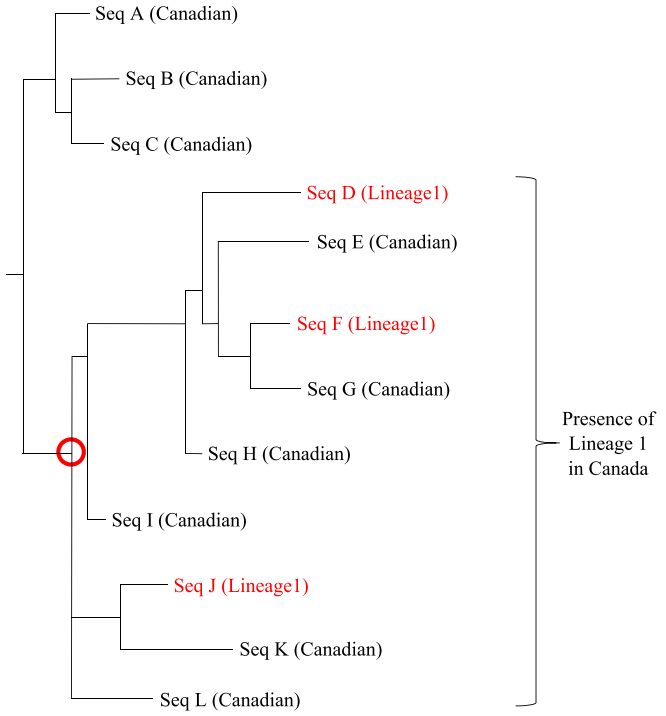

Fig. 1. Example to illustrate the methodology used to assess the presence of Shi's lineages and sublineages in a Canadian dataset. Sequences depicted in red are from the international reference dataset (Shi et al., 2010) whereas others in black are from the Canadian dataset. Red circle indicates the common node linking all reference sequences belonging to a particular lineage in Shi's international dataset (Shi et al., 2010). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

was identified (Fig. 1). The presence of this particular lineage or sublineage in Canada was defined by the presence of some Canadian sequences linked to that same common node. If not, the presence of that lineage or sublineage could not be determined. The number of sequences from each province included in the Canadian dataset and sharing common tree node with reference from international dataset was computed for each Shi's lineage or sublineage.

3. Results

3.1. Classification of Canadian strains

The Canadian dataset (5204 sequences) had a total of 253 N characters found in 164 sequences (3.2%) and these were dispersed on 149 different nucleotide positions. However, these N characters were found in a higher number of sequences at position 8, 38 and 44 with 13, 8 and 7 sequences, respectively. The number of sequences with only one N character at different positions represented 72% (118/164) of all sequences having an N character.

A total of 703 sequences having 100% pairwise genetic similarity with at least one sequence in the dataset ($n = 5204$) were removed for building the tree topology and then reinserted following the clustering step. The dataset had an average pairwise genetic distance of 0.158 (max = 0.400). Overall, 76.8% (4000/5204) of all sequences were classified into one of 34 genetic clusters (Table 1). Cluster membership illustrated on a maximum likelihood phylogenetic tree as well as a randomly selected proportional (10%) sample of sequences from each cluster are provided in supplementary material (Files 1–2). Each genetic cluster was described in terms of the total number of sequences and distribution of sequences among Canadian provinces of origin (Table 1). Based on phylogeny, only vaccine-like strains (cluster #22) were observed among sequences submitted from Manitoba. All wild-type clusters included at least one sequence from Quebec. Almost two thirds (64%, 89/139) of Ontario wild-type strains were included in five genetic clusters (cluster #16, 25, 32, 33, 34), others remained unclassified. Two clusters were predominantly formed by Ontario strains (cluster #32, 33). Information on the RFLP patterns within each genetic cluster is also provided. A total of 117 patterns were identified, including 33 that were not in the list of RFLP patterns provided by the University of Minnesota (November 2017). The mean (min-max) number of patterns per cluster was 8 (1–36). Only one genetic grouping (cluster #3) was represented by a single RFLP pattern (1-8-4), however, this latter pattern was also the predominant for 10 other wild-type clusters (Table 1). The most frequent RFLP patterns observed among the 4000 classified sequences were 1-8-4 (26%), 2-5-2 (19%), 1-8-2 (10%), 1-4-2 (8%), 1-4-4 (5%), 1-12-4 (4%), 1-8-3 (4%) and 1-4-3 (3%).

3.2. Assessing the presence of worldwide lineages

A total of 771 international reference sequences with ≥594 bp were included for this analysis. They originated from the United States (64.1%, $n = 494$), Canada (24.6%, $n = 190$ from Manitoba, Ontario and Quebec provinces), China (7.7%, $n = 59$), Japan (1.7%, $n = 13$), Taiwan (1.4%, $n = 11$), Korea (0.4%, $n = 3$) and Denmark (0.1%, n = 1). Representative of the nine lineages and some sublineages described by Shi et al. (2010) were present in this international dataset (Table 2). After merging these sequences to the Canadian dataset, the final unrooted ML-tree ($n = 5975$) is pictured in Fig. 2. Of the 25 lineages and sublineages present in the international dataset, half of them (n = 13) had reference(s) included in a genetic cluster identified by the current ML-phylogeny classification (≥15 sequences, ≥70% RBS); four were previously detected in the Canadian dataset (Table 1), whereas nine were only detectable after adding the international dataset. These nine clusters were linked to Shi's lineages/sublineages: 1.1, 1.2, 1.5, 3, 5.2, 6, 8.5, 8.7, and 9. Lineages and sublineages that were not included in a cluster were also evaluated using the methodology described in Fig. 1 to assess their presence in some Canadian provinces and results are provided in Table 2. Using that latter method, the presence of the other lineages or sublineages was identified: 1.4, 1.6, 1.7, 1.8, 2, 4 (Table 2). In Quebec, wild-type sublineages 1.6, 1.7, 1.8 and 1.9 were mostly observed, whereas 1.2 and 1.6 were the most frequent in Ontario. Lineages 4 and 9 were rarely observed in sequences from the three provinces investigated. Most sequences from Manitoba were among vaccine-like lineages (5, 8), wild-type strains being rare in the dataset submitted.

4. Discussion

A total of 34 clusters were formed by sequences obtained from samples sent by swine veterinarians between 1998 and 2016 (Supplementary material, Files 1 and 2), compared to the 18 groups previously identified on sequences gathered up to 2009 using a different classification method (Delisle et al., 2012). The 33 wild-type clusters were all related to Lineage 1 described by Shi et al. (2010). However, all of these but one were not linked by ML-phylogeny to a particular sublineage of the international dataset. This could either reflect a lack of sequences representative of the Canadian provinces in the international dataset or more recent viral subpopulations that were not captured in the study of Shi et al. (2010). For example, genetic cluster #25 (Table 1, first column), which was first detected in 2007, had expanded considerably and was the most prevalent wild-type cluster in Quebec in 2016. Other temporal trends were observed; whereas some genetic clusters (#1, 3, 13, 14, 18) were only observed before 2010, other had their first sequence submitted in 2010 or later (#8, 26, 29, 31, 33). The classification system used in the current study was thus useful to further characterize PRRSV sequences submitted by the three provinces. Swine industry would benefit from integrating genetic clusters results into ongoing surveillance activities to visualize

Table 1
Genetic clusters obtained from classification based on ML-phylogeny applied on Canadian sequences.

Cluster #[a]	Number of sequences $n = 5204$	Time range of sequence submissions[b]	Sequences per province			Number of RFLP Patterns[c]	Predominant RFLP (% of sequences)[d]	Cluster id system[e]
			Quebec $n = 5031$	Ontario $n = 151$	Manitoba $n = 18$			
1	16	2005–2009	16	0	0	5	1-4-4 (69)	CL-01
2	31	2006–2016	31	0	0	2	1-4-4 (94)	CL-02
3	18	2004–2006	18	0	0	1	1-8-4 (100)	CL-03
4	30	2006–2014	30	0	0	6	1-8-4 (53)	CL-04
5	28	2006–2012	28	0	0	5	1-8-4 (79)	CL-05
6	22	2000–2011	22	0	0	3	1-8-4 (77)	CL-06
7	16	2007–2011	16	0	0	4	1-3-4 (81)	CL-07
8	36	2010–2016	36	0	0	5	1-4-1 (83)	CL-08
9	100	2006–2015	100	0	0	9	1-8-3 (55)	CL-09
10	126	2006–2016	126	0	0	13	1-4-3 (48)	CL-11
11	65	2007–2016	65	0	0	10	1-8-4 (32)	CL-12
12	103	2003–2016	103	0	0	9	1-4-4 (63)	CL-14
13	23	2005–2009	23	0	0	4	1-12-4 (83)	CL-15
14	15	2003–2006	15	0	0	3	1-12-4 (80)	CL-16
15	16	2008–2012	16	0	0	3	1-8-4 (69)	CL-17
16	257	2002–2016	255	2	0	21	1-8-4 (41)	CL-18
17	94	2004–2016	94	0	0	14	1-23-4 (26)	CL-19
18	15	2004–2007	15	0	0	10	1-138-4 (20)	CL-20
19	28	1999–2010	28	0	0	8	1-3-3 (64)	CL-22
20	17	2003–2011	17	0	0	3	1-8-2 (82)	CL-23
21	22	1999–2013	22	0	0	5	1-4-4 (69)	CL-24
22[f]	1175	1998–2016	1141	12	18	36	2-5-2 (63)	CL-25
23	102	1999–2013	102	0	0	22	1-8-4 (26)	CL-26
24	16	2000–2011	16	0	0	8	1-117-4/ 1-40-4 (25)	CL-28
25	1372	2007–2016	1358	14	0	26	1-8-4 (53)	CL-29
26	20	2010–2015	20	0	0	2	1-8-3 (60)	CL-32
27	16	2002–2014	16	0	0	5	1-8-4 (44)	CL-33
28	20	2005–2016	20	0	0	3	1-16-4 (60)	CL-35
29	48	2013–2016	48	0	0	4	1-4-1 (54)	CL-36
30	16	2009–2016	16	0	0	7	1-4-3 (44)	CL-37
31	36	2013–2016	36	0	0	2	1-4-2 (89)	CL-38
32	42	2007–2016	9	33	0	5	1-3-2 (74)	CL-39
33	37	2014–2016	2	35	0	3	1-1-1 (62)	CL-40
34	22	2004–2015	17	5	0	5	1-3-4 (41)	CL-41
Unclassified	1204	1998–2016	1154	50	0	_	_	

[a] Sequential number attributed to each of the 34 clusters detected on the Canadian dataset (5204 sequences). Briefly, a maximum likelihood (ML) phylogeny was inferred using RAxML based on a GTR GAMMA evolutionary model and 1000 rapid bootstrap (RBS) were computed to determine branch confidence level. Then, major genetic clusters were defined based on two criteria (≥70% RBS, ≥15 sequences) applied on the tree (Lambert et al., 2019a).

[b] Date of sampling of the first and last sequence submitted for each genetic cluster.

[c] Number of different RFLP patterns observed in sequences for each genetic cluster.

[d] RLFP pattern most frequently observed in sequences for each genetic cluster. The percentage of sequences was computed as the number of sequences assigned to the predominant RFLP pattern over total number of sequences in the cluster.

[e] Correspondence to cluster id (1 to 38) obtained from the original classification system applied on 4958 sequences (Lambert et al., 2019a). A proportional (10%) random sample of sequences for each cluster id (1 to 38) as well as for the three newly formed clusters (39, 40, 41) is available in supplementary material. Each sequence provided was assigned to the cluster id (e.g. CL-01 to CL-41).

[f] This cluster only includes vaccine-like (MLV, ATP or Fostera-like) sequences from field samples ($n = 1171$), as well as the four commercial vaccine strains.

spatiotemporal movements of PRRSV subpopulations in space and time. Moreover, linking genetic clusters to additional epidemiological data could provide insight about pathways involved into PRRSV transmission at different geographical scales, which is essential in terms of disease control.

To the best of our knowledge, the number of RFLP patterns ($n = 117$) obtained is the highest reported in one paper for PRRSV ORF5 sequences and could be partly explained by the large temporal scale covered by the current project. This high number might limit their further use for surveillance, but more importantly RFLP patterns did not show a good correlation with phylogenetic clusters. These findings were in agreement with previous studies revealing a lack of correspondence between RFLP patterns and phylogeny (Brar et al., 2011; Larochelle et al., 2003). Phylogeny and RFLP patterns were based on ORF5 sequences, and not on the entire genome. Although previous studies reported similar phylogeny based on ORF5 and whole genome sequencing, misclassification can occur especially in case of recombinant strains, representing a limitation in our study (Martin-Valls et al., 2014; Zhou et al., 2018).

This study represents the first Canadian attempt to assess diversity of viral populations among Canadian provinces. Whereas some clusters were mainly composed of Quebec sequences (cluster #25), others were mostly formed by Ontario sequences (cluster #32, 33). Among the five wild-type clusters detected in Ontario, all contained sequences from Quebec. This suggests that provinces might have exchanged some viral populations or had a common source of infection. If information on submitting sites would have been available, the investigation of possible links between production sites located in different provinces and sharing the same cluster would have been relevant. Nevertheless, caution should be used in interpreting differences in their respective diversity as the size of datasets and years of sampling from Quebec (> 5000 sequences, 18 years) were not comparable to those from Ontario (151 sequences, 4 years) or Manitoba (18 sequences, 12 years). The limited number of sequences in the Ontario and Manitoba datasets, combined with our definition of a cluster which should include at least 15 sequences (Lambert et al., 2019a), precluded a complete description

Table 2
Assessment of the presence of Shi's lineages and sublineages within sequences from three provinces in Canada.

International reference dataset[a] (n = 771)			Nb of sequences from Canadian dataset sharing common tree node with reference[b]		
Lineage/ Sublineage	Nb sequences	Country of origin	Quebec	Ontario	Manitoba
1.1	19	USA, CAN		4	
1.2	16	USA, CAN	11	54	
1.3	7	USA			
1.4	12	USA	1		
1.5	20	USA			
1.6	27	USA, CAN	79	36	1
1.7	6	USA, CAN	19		
1.8	12	USA, CAN	20	3	
1.9	32	USA	20		
2a	14	USA, CAN		7	
2b	12	USA, CAN	8	7	2
3	16	Hong Kong, Taiwan			
4	12	Japan	2		
5.1 (MLV)	74	Korea, China, Japan, Denmark, USA, CAN	5	3	6
5.2	23	USA			
6	17	USA			
7 (Prime Pac)	10	USA			
8.1	7	USA			
8.2	5	USA			
8.3	5	USA			
8.5	26	USA			
8.6	5	USA			
8.7	51	China			
8.8	15	USA			
8.9 (ATP)	33	USA	2	0	6
9	164	USA, CAN	1	2	1

[a] Reference sequences (594, 597, 603 bp) from the different worldwide lineages and sublineages published by Shi et al. (2010) and provided by the authors upon request.

[b] The node linking all international sequences belonging to a particular lineage/sublineage was identified. Then sequences from either Quebec, Ontario or Manitoba also linked to that node were counted.

of diversity. These small datasets were partly attributable to confidentiality issues that limited sharing of sequences and information on submitting production sites. The representativeness of sequences obtained from other provinces is also difficult to assess for the same reason. Considering the latter limitations, the current study should be interpreted as an exploratory step. Further research at the Canadian level would be advisable to better evaluate differences of diversity between provinces as well as to investigate the frequency of inter-provincial transmission that could be potentially linked to pig movements.

When using Shi's classification, the most predominant lineages found in Canada were vaccine-like (5.1_MLV, 8.9_ATP) and wild-type lineage 1. No Canadian strains could be assigned to sublineages 1.3 to 1.5 indicating absence or scarce presence of these in Canada, even if their presence was documented in the United States (Table 2). These results are in line with the low level of live pig imported from the United States to Canadian provinces (Shi et al., 2013). International reference sequences from other sublineages (1.1, 1.2, 1.6–1.9), which were mostly from the United States, clustered tightly together and shared common node with strains from Quebec (1.7, 1.9), Ontario (1.1) or both provinces (1.2, 1.6, 1.8). These results are similar to those obtained by Shi et al. (2013) some years ago where sublineages 1.7 to 1.9 were mainly associated to Quebec-like strains whereas sublineages 1.1 to 1.6 were mainly associated to Ontario-like strains with limited exchange between provinces (Shi et al., 2013). However, this also suggests that ongoing classification would be necessary to better describe diversity in Canadian provinces and to further understand movement of viral populations between provinces and the United States. As an example, cluster 25 had emerged in 2007 and none of the international sequences regrouped with it.

The presence of lineage 2 in Canada also corroborated results from previous studies (Shi et al., 2013). However, results on lineage 2 were more difficult to interpret since the ML-phylogeny, using our dataset, splits reference strains (n = 26) in two different groups (2a, 2b) and both remained unclassified for unknown reason. Group 2b was observed in all three provinces whereas group 2a was only found in Ontario. Results on lineages 3–4, mostly from Asian countries, and on lineage 9 (USA) suggest that these could be sparse or almost non-existent in the provinces investigated as only four sequences from Canada were found in lineage 9. This corroborates findings by Shi et al. (2013). Sequences from our ML-phylogeny that were identified as lineage 8 regrouped Fostera-like and ATP-like vaccine strains (8.9), but also wild-type lineage 8 strains from China and the United States (8.7). However Canadian sequences of lineage 8 consisted of only vaccine-like strains. In addition, based on the current dataset and the methodology used, sequences belonging to the lineages/sublineages 5.2, 6 and 7 (Prime Pac) were not observed.

5. Conclusion

An important diversity was observed among > 5000 PRRSV ORF5 sequences gathered between 1998 and 2016. Classification revealed 33 wild-type clusters that were all assigned to worldwide lineage 1 of Shi's classification (Shi et al., 2010) and 5 were shared by Quebec and Ontario. However, differences in their respective diversity should be further assessed using more representative sequence samples from provinces outside Quebec. In addition, factors modulating PRRSV diversity such as pig movement occurring within and between provinces should be further investigated in a perspective of disease control.

Funding

This work was supported by the Swine Innovation Porc within the Swine Cluster 2 (project #1343), funded by Agriculture and Agri-Food Canada through the AgriInnovation Program, provincial producer organizations and industry partners. Financial support was also provided by the Éleveurs de Porcs du Québec and Zoetis Canada inc.

Declaration of Competing Interests

The authors declare that they have no competing interests.

Authors' contributions

MEL, JA, BD, ZP and SD designed the methodology; BD performed RFLP pattern analyses; MEL and BD performed phylogenetic analysis; MEL performed descriptive statistics and wrote the first draught of the manuscript; All authors revised and approved the final manuscript.

Acknowledgments

The authors would like to thank all Quebec swine veterinarians (AVIA) and Quebec diagnostic laboratories for their willingness to share data. Special thanks to swine veterinarians from other Canadian provinces that shared sequence data with the LEMP for the current project as well as to Dr. Shi and his research team for providing representative sequences of their lineages upon request.

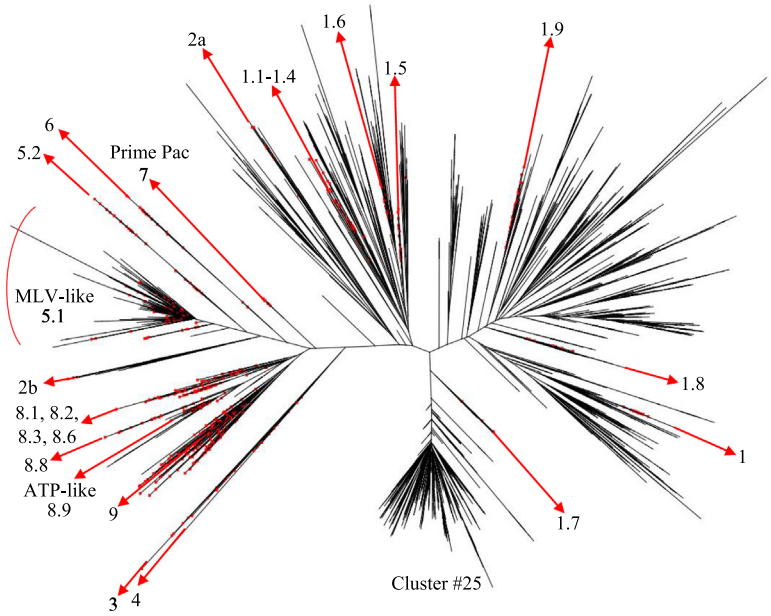

Fig. 2. Maximum likelihood tree of a Canadian dataset merged to international reference dataset (n = 5975). References from international dataset provided by other countries than Canada are depicted in red; others were from Canada. Arrows are pointing reference sequences belonging to different lineages/sublineages identified by Shi et al. (2010). The actual predominant wild-type cluster identified in Quebec is identified (cluster #25). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2019.103999.

References

Adams, M.J., Lefkowitz, E.J., King, A.M.Q., Harrach, B., Harrison, R.L., Knowles, N.J., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Mushegian, A.R., Nibert, M., Sabanadzovic, S., Sanfacon, H., Siddell, S.G., Simmonds, P., Varsani, A., Zerbini, F.M., Gorbalenya, A.E., Davison, A.J., 2017. Changes to taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses (2017). Arch. Virol. 162, 2505–2538. https://doi.org/10.1007/s00705-017-3358-5.

Albina, E., 1997. Porcine reproductive and respiratory syndrome: ten years of experience (1986-1996) with this undesirable virus infection. Vet. Res. 28, 305–352.

Arruda, A.G., Poljak, Z., Friendship, R., Carpenter, J., Hand, K., 2015. Descriptive analysis and spatial epidemiology of porcine reproductive and respiratory syndrome (PRRS) for swine sites participating in area regional control and elimination programs from 3 regions of Ontario. Can. J. Vet. Res. 79, 268–278.

Brar, M.S., Shi, M., Ge, L., Carman, S., Murtaugh, M.P., Leung, F.C.C., 2011. Porcine reproductive and respiratory syndrome virus in Ontario, Canada 1999 to 2010: genetic diversity and restriction fragment length polymorphisms. J. Gen. Virol. 92, 1391–1397. https://doi.org/10.1099/vir.0.030155-0.

Delisle, B., Gagnon, C.A., Lambert, M.E., D'Allaire, S., 2012. Porcine reproductive and respiratory syndrome virus diversity of eastern Canada swine herds in a large sequence dataset reveals two hypervariable regions under positive selection. Infect. Genet. Evol. 12, 1111–1119. https://doi.org/10.1016/j.meegid.2012.03.015.

Lambert, M.E., Arsenault, J., Audet, P., Delisle, B., D'Allaire, S., 2019a. Evaluating an automate clustering approach in a perspective of ongoing surveillance of porcine reproductive and respiratory syndrome virus (PRRSV) field strains. Infect. Genet. Evol. https://doi.org/10.1016/j.meegid.2019.04.014.

Lambert, M.E., Arsenault, J., Delisle, B., Audet, P., Poljak, Z., D'Allaire, S., 2019b. Impact of alignment algorithm on the estimation of pairwise genetic similarity of porcine reproductive and respiratory syndrome virus (PRRSV). BMC Vet. Res. 15, 1–10. https://doi.org/10.1186/s12917-019-1890-0.

Lambert, M.E., Arsenault, J., Poljak, Z., D'Allaire, S., 2012a. Correlation among genetic, Euclidean, temporal, and herd ownership distances of porcine reproductive and respiratory syndrome virus strains in Quebec, Canada. BMC Vet. Res. 8. https://doi.org/10.1186/1746-6148-8-76.

Lambert, M.E., Poljak, Z., Arsenault, J., D'Allaire, S., 2012b. Epidemiological investigations in regard to porcine reproductive and respiratory syndrome (PRRS) in Quebec, Canada. Part 1: biosecurity practices and their geographical distribution in two areas of different swine density. Prev. Vet. Med. 104, 74–83. https://doi.org/10.1016/j.prevetmed.2011.12.004.

Larochelle, R., D'Allaire, S., Magar, R., 2003. Molecular epidemiology of porcine reproductive and respiratory syndrome virus (PRRSV) in Quebec. Virus Res. 96, 3–14. https://doi.org/10.1016/S0168-1702(03)00168-0.

Martin-Valls, G.E., Kvisgaard, L.K., Tello, M., Darwich, L., Cortey, M., Burgara-Estrella, A.J., Hernandez, J., Larsen, L.E., Mateu, E., 2014. Analysis of ORF5 and full-length genome sequences of porcine reproductive and respiratory syndrome virus isolates of genotypes 1 and 2 retrieved worldwide provides evidence that recombination is a common phenomenon and may produce mosaic isolates. J. Virol. 88, 3170–3181. https://doi.org/10.1128/JVI.02858-13.

Mateu, E., Diaz, I., 2008. The challenge of PRRS immunology. Vet. J. 177, 345–351. https://doi.org/10.1016/j.tvjl.2007.05.022.

Murtaugh, M.P., Stadejek, T., Abrahante, J.E., Lam, T.T.Y., Leung, F.C.C., 2010. The ever-expanding diversity of porcine reproductive and respiratory syndrome virus. Virus Res. 154, 18–30. https://doi.org/10.1016/j.virusres.2010.08.015.

Rahe, M.C., Murtaugh, M.P., 2017. Effector mechanisms of humoral immunity to porcine reproductive and respiratory syndrome virus. Vet. Immunol. Immunopathol. 186, 15–18. https://doi.org/10.1016/j.vetimm.2017.02.002.

Rosendal, T., Dewey, C., Friendship, R., Wootton, S., Young, B., Poljak, Z., 2014. Spatial and temporal patterns of porcine reproductive and respiratory syndrome virus (PRRSV) genotypes in Ontario, Canada, 2004-2007. BMC Vet. Res. 10. https://doi.org/10.1186/1746-6148-10-83.

Shi, M., Lam, T.T.Y., Hon, C.C., Murtaugh, M.P., Davies, P.R., Hui, R.K.H., Li, J., Wong, L.T.W., Yip, C.W., Jiang, J.W., Leung, F.C.C., 2010. Phylogeny-based evolutionary, demographical, and geographical dissection of north American type 2 porcine reproductive and respiratory syndrome viruses. J. Virol. 84, 8700–8711. https://doi.org/10.1128/Jvi.02551-09.

Shi, M., Lemey, P., Brar, M.S., Suchard, M.A., Murtaugh, M.P., Carman, S., D'Allaire, S., Delisle, B., Lambert, M.E., Gagnon, C.A., Ge, L., Qu, Y.H., Yoo, D., Holmes, E.C., Leung, F.C.C., 2013. The spread of type 2 porcine reproductive and respiratory syndrome virus (PRRSV) in North America: a phylogeographic approach. Virology 447, 146–154. https://doi.org/10.1016/j.virol.2013.08.028.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal omega. Mol. Syst. Biol. 7, 539. https://doi.org/10.1038/msb.2011.75.

Wesley, R.D., Mengeling, W.L., Lager, K.M., Clouser, D.F., Landgraf, J.G., Frey, M.L., 1998. Differentiation of a porcine reproductive and respiratory syndrome virus vaccine strain from north American field strains by restriction fragment length polymorphism analysis of ORF 5. J. Vet. Diagn. Investig. 10, 140–144. https://doi.org/10.1177/104063879801000204.

Zhou, L., Kang, R., Zhang, Y., Ding, M., Xie, B., Tian, Y., Wu, X., Zuo, L., Yang, X., Wang, H., 2018. Whole genome analysis of two novel type 2 porcine reproductive and respiratory syndrome viruses with complex genome recombination between lineage 8, 3, and 1 strains identified in southwestern China. Viruses 10. https://doi.org/10.3390/v10060328.