Porcine reproductive and respiratory syndrome virus in Ontario, Canada 1999 to 2010: genetic diversity and restriction fragment length polymorphisms

Manreetpal Singh Brar,1 Mang Shi,1 Li Ge,2 Susy Carman,2 Michael P. Murtaugh3 and Frederick Chi-Ching Leung1

Classification of Ontario porcine reproductive and respiratory syndrome virus (PRRSV) field isolates (n=505) from 1999 to 2010, based on a global type 2 PRRSV ORF5 phylogenetic framework, revealed genetic diversity comparable to PRRSV in the USA, with sequences assigned to five of nine lineages (1, 2, 5, 8 and 9). Importantly, the tree topology indicated a Canadian ancestry for the highly virulent MN184-related strains that first emerged in 2001 in the USA. Mapping of the RFLP patterns onto the phylogenetic tree revealed numerous examples of different RFLP patterns located within the same phylogenetic cluster. Statistical analysis showed occurrences where similar RFLP patterns masked diverse genetic distances and instances of close genetic proximity with divergent RFLP patterns. Collectively, extensive genetic diversity prevails in type 2 PRRSV in one region of the North American swine industry, and it is not described adequately by RFLP typing, which might have value in differentiating strains at the local farm level.

The late 1980s to early 1990s witnessed the spread of porcine reproductive and respiratory syndrome (PRRS) disease amongst swine herds across several continents (Baron et al., 1992; Botner et al., 1994; Hirose et al., 1995; Paton et al., 1992; Wensvoort et al., 1991; Zimmerman et al., 1997). Porcine reproductive and respiratory syndrome virus (PRRSV), initially known as Lelystad virus, was identified as the causative agent in Europe (Terpstra et al., 1991). Not long after, the prototype North American isolate ATCC VR-2332 was shown to cause the same disease in the USA (Benfield et al., 1992; Collins et al., 1992). The positive-sense, ssRNA-enveloped virus is grouped with lactate dehydrogenase elevating virus, equine arteritis virus and simian hemorrhagic fever virus in the family Arteriviridae, order Nidovirales (Cavanagh, 1997; Conzelmann et al., 1993). Flanked by a 5′-cap and 3′-polyadenylation tail, the genome consists of nine ORFs spanning approximately 15 kb. ORF1a and ORF1b together comprise the majority of the genome and encode the ORF1a and ORF1ab polyproteins, which are cleaved into a series of non-structural proteins (NSPs) that function in replication and transcription (Ziebuhr et al., 2000). Substantial gains in elucidating more precise functional roles of these proteins have been made in recent times (Fang & Snijder, 2010). The structural viral proteins GP2, E, GP3, GP4, GP5, M and N are encoded by ORF2a, ORF2b and ORFs 3–7, respectively (Snijder & Meulenberg, 1998). The European (EU; type 1) and North American (NA; type 2) strains form the dichotomized genotypes of PRRSV. This division has been established through extensive investigation revealing divergent differences at both genomic (Murtaugh et al., 1995; Nelsen et al., 1999) and antigenic levels (Dea et al., 2000; Kim & Yoon, 2008; Magar et al., 1995). Significant heterogeneity is present within each genotype as well. Evolution of PRRSV is perceived to be driven by mutation and recombination (Murtaugh et al., 2010). Extensive genetic diversity and rapid evolution of PRRSV complicates efforts to develop broadly protective vaccines. Thus, it is important to evaluate descriptive methods that characterize genetic diversity.

Because GP5 is highly variable and important immunologically, it is widely used for investigation of genetic diversity and establishing phylogenetic relationships of PRRSV isolates (Kapur et al., 1996; Meng et al., 1995). Recently, a global PRRSV type 2 phylogenetic tree was constructed based on ORF5 (Shi et al., 2010) in an attempt...
to achieve a holistic perspective on North American PRRSV versus traditional phylogenetic inferences drawn from more limited datasets. Though Shi and colleagues analysed a voluminous ORF5 sequence dataset (n>8500), the sampling was not strictly random and regional concentrations of swine production may have been poorly represented (Shi et al., 2010). Here, we examined additional ORF5 sequences from 505 PRRSV diagnostic field cases collected by the Animal Health Laboratory at the University of Guelph in the Province of Ontario in Canada from 1999 to 2010, which were not included in the prototypical global type 2 PRRSV phylogenetic tree. Due to client confidentiality requirements, the sequence and other isolate information were not deposited in public databases, but individual sequences may be made available upon request. Importantly, inclusion of previously unavailable sequences will augment the robustness of the existing tree (Shi et al., 2010) by improving bootstrap values especially at moderately confident nodes. Additionally, improved representation of the genetic diversity of poorly sampled lineages and sublineages will be achieved with the inclusion of these additional sequences, thereby better aligning the perceived genetic richness of type 2 PRRSV with the actual state prevailing in the field. Periodic refinement of the global phylogeny with new data will help to maintain a comprehensive and representative type 2 PRRSV framework for categorizing novel isolates and estimating their origin. This approach is expected to be useful in understanding regional type 2 PRRSV relationships in a global context of geographical distributions and of animal and virus movement.

Since RFLP typing is widely used in North America to communicate genetic classification of type 2 PRRSV isolates to swine producers, we also investigated its use as an estimator of the genetic relatedness amongst PRRSV strains.

Isolates examined in this study were obtained from swine farms concentrated in south-western Ontario, Canada. The 505 Ontario ORF5 sequence dataset from 1999 to 2010 was fitted to the previously established global type 2 PRRSV backbone (n=612) using computational methods as described previously (Shi et al., 2010). The Ontario sequences were lodged in five of the nine extant lineages, namely 1, 2, 5, 8 and 9 (Shi et al., 2010). Eighty-five per cent of all sequences were in lineages 1 (n=337) and 2 (n=93), lineages 5 and 8 together contained 14% of the isolates and two sequences (0.4%) were in lineage 9 (Fig. 1). The majority of lineage 1 sequences were scattered throughout the existing nine sublineages, with prominence in sublineages 1.2-1.4 and 1.6-1.9. Also, a previously unknown sublineage 1.10 was identified. A number of sequences, at different locations in the tree, clustered to form ancestral groups in either single or multiple sublineages. Such isolates may provide valuable information that sheds light on the emergence of highly diverse or outbreak-associated clusters. In particular, they may serve to elucidate the ancestral source of sudden appearances of outbreak strains.

In 2001 the state of Minnesota, USA, witnessed an abrupt appearance of an exceptionally virulent type 2 PRRSV with a 1-8-4 RFLP pattern, and that was designated PRRSV strain MN184 (Han et al., 2006). At the time of its isolation, the origin of this MN184 cohort of viruses was shrouded in mystery. However, the analysis provided here reveals a group of Ontario sequences immediately ancestral to the Minnesota cluster (Fig. 1, rounded, black-bordered rectangle), and other Canadian clusters found in the phylogenetic vicinity. The absence of these sequences in public databases, at the time of the MN184 outbreak, helps to explain the apparently mysterious appearance of MN184. Moreover, the high diversity of the ancestral strains of MN184-type PRRSV also accounts for the genetic disparity observed between the whole genome sequences of the first two MN184 isolates (Han et al., 2006). Previously, it was conjectured that MN184 may have emerged from local populations of less-virulent viruses whose genotypes had not been sequenced. Since ancestral MN184 strains have been historically present in one or more Canadian provinces, it is more likely that virulent PRRSV of this type were circulating in Canada and reached Minnesota through transportation of infected pigs. Several million live pigs are exported annually from Canada to the USA for growth and slaughter. For example, in 2008, 9.3 million pigs arrived in the USA from Canada. Of these, 1.6 million were from Ontario (ThePigSite, 2009).

Our findings show that lineage 2 diversity is greater than was estimated previously (Shi et al., 2010). The new Ontario samples expanded the sample size of lineage 2 from 115 to 208 sequences. Collectively, with respect to lineages 1 and 2, it is apparent that Canadian type 2 PRRSV diversity is, at the very least, comparable to that of the USA, if not greater. Moreover, the large concentration of Ontario sequences in lineages 1 and 2, including clusters at basal positions, suggests that Ontario, Canada may be the ancestral home of these lineages. It is noteworthy that the earliest documented evidence of PRRSV in North America was in eastern Canada (Carman et al., 1995). All Ontario sequences in lineages 5 and 8 represented field reisolates of two vaccine strains, Ingelvac MLV in sublineage 5.1, and Ingelvac ATP in sublineage 8.9 (Shi et al., 2010).

RFLP analysis of PRRSV ORF5 was pioneered by Wesley and colleagues to differentiate Ingelvac MLV (then known as RespPRRS) vaccine from wild-type PRRSV (Wesley et al., 1998). The method was quickly adopted as a simple type 2 PRRSV typing system throughout North America, including in Ontario and other parts of Canada (Cai et al., 2002; Rosendal et al., 2010; Larochelle et al., 2003). Initially, RT-PCR amplified ORF5 fragments were subjected to digestion with the restriction enzymes MluI, HincII andSacII, separately. The resulting electrophoretic pattern of bands produced by each digestion, after running on agarose gels, was used to assign a three digit numerical code, with each digit corresponding to a specific pattern of bands (Wesley et al., 1998). PRRSV typing now is accomplished largely, but not exclusively, by prediction of restriction site
PRRSV in Ontario: ORF5 genetic diversity and RFLP patterns
locations in ORF5 from nucleic acid sequence data (Rosendal et al., 2010). The resulting RFLP patterns have been used in assessing genetic diversity and strain homology amongst different PRRSV isolates (Cheon & Chae, 2000; Itou et al., 2001; Wang et al., 2008). However, its use in assessing genetic relationships of PRRSV isolates has not been rigorously examined, even though altered RFLP patterns following simple PRRSV passage through a pig raised doubts about its predictive value (Wesley et al., 1999; Cha et al., 2004).

Here, we analysed the phylogenetic relevance of RFLP genotyping in a subset of 441/505 Ontario ORF5 sequences for which the RFLP pattern was available. The occurrence and frequency of various RFLP patterns at the lineage and sublineage levels are categorized and is presented in Table 1. Lineages 1 (intra-lineages 1.1-1.4 and 1.10), 5 and 9 did not show predominant RFLP patterns. Lineage 2 had patterns of moderate frequency (n>5 instances), in addition to a dominant pattern. All lineages possessed multiple minor RFLP patterns (n<5 instances) with intra-lineage 1.10

**Table 1. Classification and abundance of Ontario PRRSV ORF5 RFLP patterns in the global type 2 PRRSV phylogenetic tree**

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Intra-lineage designation</th>
<th>ORF5 RFLP patterns (frequency per 441 in total)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Major intra-lineage</td>
<td>Minor intra-lineage</td>
</tr>
<tr>
<td>1</td>
<td>1.1</td>
<td>1-8-2 (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-21-2 (2)</td>
</tr>
<tr>
<td></td>
<td>1.2-1.4</td>
<td>1-12-2 (20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-3-2 (19)</td>
</tr>
<tr>
<td></td>
<td>1.6</td>
<td>1-18-4 (99)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-13-4 (9)</td>
</tr>
<tr>
<td></td>
<td>1.7-1.9</td>
<td>1-8-4 (23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-12-4 (13)</td>
</tr>
<tr>
<td>1.10</td>
<td></td>
<td>1-16-4 (10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-16-2 (8)</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td>1-4-2 (1)</td>
</tr>
<tr>
<td>2</td>
<td>–</td>
<td>1-3-4 (44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-2-4 (12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-1-2 (10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-6-2 (8)</td>
</tr>
<tr>
<td>5</td>
<td>Ingelvac MLV associated</td>
<td>2-1-2 (19)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1-3-4 (19)</td>
</tr>
<tr>
<td>8</td>
<td>Ingelvac ATP associated</td>
<td>1-3-2 (1)</td>
</tr>
<tr>
<td>9</td>
<td>–</td>
<td>1-3-2 (1)</td>
</tr>
</tbody>
</table>

**Fig. 1.** Incorporation and classification of Ontario PRRSV ORF5 sequences into the global type 2 PRRSV phylogenetic tree. Sequence alignment was carried out using MUSCLE v3.6 (Edgar, 2004) followed by recombination screening (Martin et al., 2005). Phylogeny construction was performed by MrBayes v3.2 (Ronquist & Huelsenbeck, 2003) and midpoint rooted using parameters as described previously (Shi et al., 2010). (Left) The entire tree with the lineages containing Ontario strains shaded in grey. The number of strains and their percentage share from this study in the current dataset are indicated within each shaded box. Regions of the global tree containing Ontario samples are expanded to enable detailed viewing (right). The solid dots (●) represent Ontario taxonomic units and the solid triangles (▲) denote vaccine strains. For the sake of clarity, only representative Ontario strains are indicated. All new clusters or branches formed by the inclusion of Ontario strains are shown. For lineage 1, the sublineage designation is indicated by a vertical stroke to the right of the intra-lineage tree. The clustering of different RFLP patterns at various spots is shown in rectangular grey boxes. The MN184-related cluster and the immediately ancestral group of Ontario strains are depicted in sublineage 1.7-1.9 outlined by a broken line. Numbers at nodes are bootstrap values.
Fig. 2. Distribution and frequency of pair-wise patristic genetic distances of strains with identical ORF5 RFLP pattern scores. (a) Box plots generated in R depicting the distribution of patristic genetic distances of all possible pair-wise comparisons under set ORF5 RFLP scores using unweighted (left) and weighted (right) scoring methods. The lower and upper boundaries of the central box demarcate the first and third quartiles, respectively. The line in the middle of the box is the median value and circles are data outside the 95% confidence interval. (b) Frequency distribution of patristic genetic distances for each unweighted (left) or weighted (right) ORF5 RFLP score.
displaying ten different such patterns. Moreover, some of the minor RFLP patterns were present in multiple lineages or sublineages.

Two approaches were used to assess the confidence of RFLP in predicting genetic distance or similarity (actual pair-wise genetic distance) between field strains. First, clustering of RFLP patterns was examined on the phylogenetic tree constructed in Fig. 1. In numerous cases either an intra-lineage cluster comprising multiple RFLP patterns or multiple sublineages with the same RFLP pattern could be found (Fig. 1, shaded boxes). Thus, qualitatively, RFLP typing did not effectively predict phylogenetic relationships.

Second, pair-wise comparisons were made of RFLP patterns and patristic genetic distance (Fourment & Gibbs, 2006) for each strain in the Ontario sequence dataset to determine the ability of RFLP types to mirror genetic distance quantitatively. The distribution and frequency of the genetic distances obtained from these comparisons is shown in Fig. 2, using unweighted and weighted categorical RFLP scoring methods to enable graphical representation of the results. In the unweighted technique, strains with the same RFLP pattern in a pair-wise comparison were assigned a score of zero for the comparison. Pairs of strains with only a single difference in one of the three digits of the RFLP patterns were assigned a score of one. Similarly, scores of 2 and 3 were assigned for non-identity in two of three and three of three digits of the RFLP numerical code, respectively. Thus, all comparisons could be grouped into one of the four RFLP score categories: 0, 1, 2 and 3. In the weighted technique, a value of 4 was assigned to the first (MluI) or the third (SacI) digits of the RFLP code every time a difference occurred at these positions. A difference in the second (HincII) digit was given a value of 1. This variability in assigning values was due to the inherently lower specificity of the HincII restriction site, since the hexanucleotide recognition sequence contained two ambiguous residues (GTYRAC). Thus, HincII could cleave four different types of hexanucleotide sequences instead of one like MluI or SacI. As before, a value of zero was assigned to RFLP digits that were the same in a pair-wise comparison. The RFLP score of a particular comparison was then determined by adding the values at each of the three digits (cumulative score). Hence, in the weighted scoring method, all comparisons could be placed into one of the following RFLP score categories: 0, 1, 4, 5, 8 and 9. Both the box plots and the histograms in Fig. 2 highlight several salient points. Strains with the identical RFLP type (cumulative score=0) usually were genetically similar, but with notable exceptions. Scores greater than zero exhibited a diverse degree of genetic distances, which were independent of the score, using both weighted and unweighted methods. Numerous instances were found of strains with the same or similar RFLP patterns displaying great genetic diversity, and strains with highly contrasting RFLP scores showing close genetic similarity. Collectively, the analysis showed that RFLP typing is unreliable for estimating genetic relatedness of PRRSV isolates in a large swine producing region such as Ontario, Canada. In a region of limited genetic diversity, such as a swine farm with one type of PRRSV, RFLP typing may be useful for detection of a new virus.

In summary, phylogenetic analysis of 505 PRRSV ORF5 sequences gathered from swine in Ontario, Canada, over a 12-year-period revealed extensive genetic diversity equivalent to that previously described over an approximately 20-year-period in the USA. Given the genetic richness of PRRSV in eastern Canada and extensive cross-border live pig trade with the USA, it is not surprising that the ancestors of the highly virulent MN184-type viruses appear to have been in this region. This example highlights the value of analysing PRRSV isolates in the context of total population diversity. Comparison of RFLP patterns in ORF5 with phylogenetic analysis showed that RFLP typing is not adequate for large-scale genetic analysis, but may be useful for the detection of local introductions of new viruses.

Acknowledgements

This work was partially supported by the Strategic Research Theme of Infection and Immunology, The University of Hong Kong.

References


