Interpreting PRRS ORF5 sequencing, can we do better?

D. Marthaler, M. Murtaugh

Key points:

- **RFLP (restriction fragment length polymorphism) cut patterns may misconstrue genetic relationships among PRRS viruses.**
- **Phylogenetic relationship and nucleotide percent identity together help to interpret PRRS ORF5 sequencing.**
- **Through a collaborative effort, we hope to propose a diagnostic report process, using phylogenetic relationship, percent similarity, and RFLP cut pattern, to better assist the swine producers and veterinarians in understanding results. The methodology will be compatible with whole genome sequencing as well.**

Currently, interpreting PRRS ORF5 sequencing is not an easy task, and things will only get more complicated when complete genome sequencing becomes a routine service. Interpreting by RFLP is a problem since large PRRS outbreaks from a single introduction can give multiple RFLP patterns, and because a single RFLP pattern can be found in multiple, genetically unrelated, clusters. RFLP was useful when there was less genetic variation in PRRSV and fewer types. Now, there are hundreds of RFLP types and routine sequencing within systems shows various RFLP cut patterns in closely related sequences.

For these reasons, RFLP by itself can miss or misrepresent important elements of the diagnostic picture. Adding phylogenetic information into a naming scheme will clarify when the same RFLP is in the same or different genetic clusters (also called lineages, see Figure 1). Routine use of diagnostic dendrograms and percent similarity will provide more accurate and better information. Interestingly, phylogenetics is routinely used for other porcine pathogens, including:

- Swine influenza A viruses are classified by subtype (H1 or H3), and these subtypes have been further classified based on their phylogenetic relationships (eg. alpha, beta, cluster I, II, etc.). This information is useful for vaccine selection.
- Phylogenetic analysis is used to discriminate PCV2 strains (2a-2e). Because of phylogenetics, it was possible to recognize the recent emergence of PCV2d, and the existence of PCV2e, which sat unrecognized in VDL databases since 2006.
- Genotype classification of rotavirus group B and group C VP7 gene segments was recently proposed so that diagnostic sequence can be used to identify emergence of new genotypes in a herd.

Genetics-based terminology is used every day for IAV, PCV2, and rotavirus diagnostic reports. We are missing this key component when discussing and understanding the diversity and emergence of new PRRS strains.

Currently, type 2 PRRS viruses are diagnostically classified by ORF5 cut patterns. Patterns are based on the presence and distribution of three specific 4 or 6-base sequences in the ~600 base gene that are recognized by three DNA enzymes, known as MluI, HincII, and SacII. Today, there are 5 known patterns for MluI, 156 for HincII, and 15 for SacII, totaling 11,700 possible combinations. And, they have little connection to genetic relatedness. The problem with using RFLP as the main tool for describing PRRSV isolates is illustrated by two examples:

1) Severe 1-7-4 outbreaks swept through large and small operations starting in 2014. By 2015, individual systems had dozens of 1-7-4 viruses, but also various others, including 1-6-4, 1-21-4, 1-4-4, and 1-7-3. What were these? A dendrogram showed they all were from same 1-7-4 outbreak. Minor mutations were changing cut patterns, but the viruses were still derived from the outbreak virus. They were not new introductions. RFLP did not could show this result.
2) Figure 1 below shows a set of PRRSV ORF5 sequences spanning the full range of type 2 diversity. Major groups are identified by color and lineage (L1-L9). Asterisks show strikingly different viruses, all of which are a 1-4-4 RFLP. In this situation, RFLP by itself is an unhelpful, or a misleading, indicator of virus genotype.

Should we disregard the RFLP pattern completely and develop a new classification? Let's not throw the baby out with the bath water. Since RFLP patterns are so familiar, we want to improve it by developing a classification system that is scientifically sound and that integrates our knowledge on phylogenetic relationship, nucleotide identity, in conjunction with RFLP pattern. In addition, we want to provide useful reference strains to easily interpret the diagnostic reports. Most importantly, the system must be user friendly, practical, and helpful to producers and veterinarians. Going forward, it also can lay the foundation for whole genome comparisons. We will appreciate your comments and feedback on these goals.

Next week, we will discuss using phylogenetic relationships and percent sequence identity in combination with RFLP to interpret virus relationships.

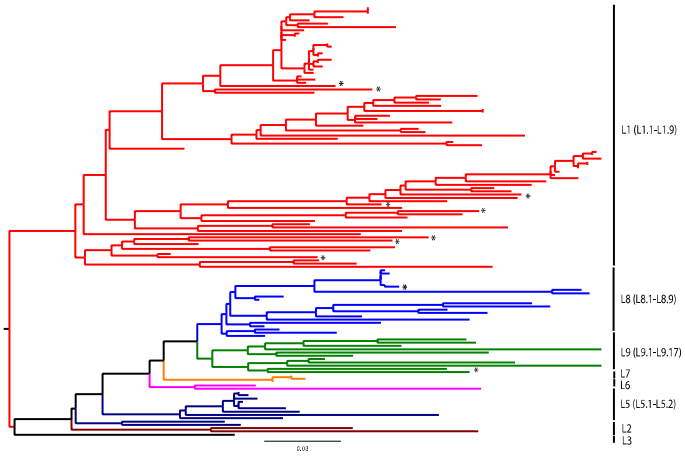

Figure 1. Phylogenetic tree of PRRS ORF5 strains generated at the MVDL in July 2016 plus MVDL's reference strains. The different lineages represented by color. Previous proposed sub-lineages represented in the parenthesis. Asterisk represents 144 strains separate throughout the different lineages.