



VIRAL DISEASES

VVD-035

AN AUTOMATED CLASSIFICATION SYSTEM FOR PRRS ORF5 SEQUENCES

M.É. Lambert, P. Audet, J. Arsenault, B. Delisle, S. D'allaire.

University of Montreal, Faculty of Veterinary Medicine, Canada.

A classification system based on PRRS ORF5 sequences would be useful to assess progress in control program and for surveillance purposes. The objective of this study was to evaluate the feasibility of an automated system that could be used for very large datasets and would consistently give reliable results over time.

A database of 3661 sequences (January 1998 to December 2013) was used as a baseline dataset and classified into distinct groups. Briefly, a maximum likelihood phylogeny was inferred using RAxML and computed on 1000 randomized stepwise Maximum parsimony starting trees to select the best-scoring ML tree. Criteria used for grouping in the classification were a minimum number of sequences of 15 per group with a rapid bootstrap >70%. Following this, approximately 1300 sequences were progressively added to the baseline dataset according to sampling date, in order to mimic sequences that our laboratory would receive in a 3-month period and, analyses were repeated to obtain a total of eleven additional runs.

The baseline dataset led to the formation of 27 distinct major groups, whereas the final dataset (n=4958) led to 33 major groups classifying more than 75% of the sequences. Unclassified sequences were either insufficient in number to form a group or branches were not well supported. Analyses revealed that most clusters were stable through time under these criteria, such that sequences attributed to one group in initial classification stayed in the same group for the 12 different analyses. However, with time and an increasing number of sequences, some initial groups were subdivided into subgroups. This is an important aspect for monitoring since one specific wild-type group increased from 0% in 2007 to 45% of all sequences in 2016.

This automated classification system will be useful for assessing spatial and temporal evolution and for surveillance purposes of PRRS virus strains.