The bioinformatic PRRSV analysis program will be developed in Python. To ensure user friendliness all processes will be automated. All the user will have to do is place their PRRSV sequences into a directory and click a run button on a GUI then the generated data will be placed in an output directory for viewing.

The open reading frames will be annotated using a weighted scoring algorithm trained from published PRRSV genomes accessed on GenBank. ORF 1 is identified first and the remaining ORFs are identified dynamically based on the site of ORF1a/b. If the input sequence is only a partial genome it will be aligned to a reference genome to replace missing key data points for ORF identification.

Once all open reading frames are identified and annotated alignments will be created using modules from BioPython. From these annotated alignments…this is where you describe all of the output we’ll be generating which will all be developed by common coding practices