For developing own method:

~~Idea: predict the emergence of new alphavirus, or predict which pre-existing strain will be the next epidemic?~~

~~-Character: death? Disease in humans? Nucleotide mutation based on mutation rate?~~ ~~Flu can use antigenic distance on maximum likelihood tree compared to the phylogeny~~

Look at flu prediction papers

Method:

-Get full-length genomes of a wide variety of alphaviruses. Make a phylogenetic tree out of them

-Examine: where are the areas on the genome where mutations are most likely to occur and drive changes in \_\_\_\_\_?

Yang 1995 gamma

Substitution rates

Estimate rate of evolution using sliding window/entropy. Do not use Shannon entropy, but figure out a method that runs similarly, but still checks diversity in a way useful for phylogeny. Infer evolution from diversity?

-how to set up method that can do this in R?

Sliding window has some good direction – has been used on some RNA viruses

-most papers detail specific programs for this. Possible to implement into R?

-also run tests to check for sites with anomalous nonsynonymous/synonymous rates

Do on E2 protein of VEEV