**POST-DADA2**

**VGSC TUTORIAL**

**By Raquel Lima-Cordon**

This tutorial has been prepared to analyze the insecticide resistance gene VGSC. The files required from DADA2 are: ASVBimeras.txt and VGSC2\_seqtab.tsv. **Keep in mind that these files need filtration**, starting by removing bimeras and trimming the ASVs.

1. Run the R script named “Script1.R”. This script removes bimeras and trim the ASV to 132bp as well as prepares the clean data for further steps. This script can be found in our [Drive](https://drive.google.com/drive/folders/1-FLbgyRh4ui_cUWkenJWzHbLCK5U1n68?usp=drive_link). The only change you might need to make is to adjust the working directory path to your own working directory.
2. Once you finish running “Script1.R”, Run the first half of the R script named “Script2\_PRE\_&\_POST\_KMER.R” just before the section labeled as the “POST-KMER” SECTION. In this first half, what we are doing is to filter the data, we are removing any ASV below 50 read count and any secondary most abundant ASV with less than 10% of the reads relative to the top ASV (We still need to decide this threshold as it would impact on whether a sample is homozygous or a possible heterozygous).
3. After running the first half of the “Script2\_PRE\_&\_POST\_KMER.R”, upload the output file named “Top2ASV\_sequences.fasta” to the BROAD SERVER.
4. On the BROAD SERVER, create a new directory (name it as you prefer). The required files for running the kmer script, which should be uploaded into the newly created directory, can be found at <https://drive.google.com/drive/folders/1-FLbgyRh4ui_cUWkenJWzHbLCK5U1n68?usp=drive_link> . The necessary files include:

* FindKmersASVs.pl
* VGSC\_kmer\_table.txt (This table should be updated as new variants arise)
* Top2ASV\_sequences.fasta (this is the output from the first half of “Script2\_PRE\_&\_POST\_KMER.R”)

Once you have these files, run the following command, adjusting the -o to your preferred name:

FindKmersASVs.pl -k VGSC\_kmer\_table.txt -f Top2ASV\_sequences.fasta -o vgsc\_R7\_R9.txt

5. Save the output file “vgsc\_R7\_R9.txt” from the KMER script and return to R to finish running the second part of the “Script2\_PRE\_&\_POST\_KMER.R”. You should get the same output file as the one I just shared.