**Viral Genome Mapping Pipeline Notes**

CPU Requirements: To run on local computer or server

Languages: Nextflow, Python

Dependencies: bbtools, samtools, bcftools, SNAP, local viral whole genome database

1. Input: fastq.gz files (1-15 million reads)
2. Read quality filtering and adapter filtering (see CLOMP)
3. Host read subtraction (see CLOMP)
   1. Align to Human Genome
4. Align to local database (build using all respiratory virus whole genome sequences: biofire, ncbi). Alignment tool: SNAP aligned
5. Map to whole genome that matches
6. Match to the following criteria: highest read match, highest length of reference match
7. Create consensus sequence FASTA (see LAVA)
8. Output: consensus sequence

Ideas:

Add option to perform VAPID workflow to prepare for Genbank submission.