**Supplementary File 1**

**Bioinformatics Tool Parameters**

#----Bbduk----

bbduk.sh in1=<input\_R1\_fastq> out1=<output\_R1\_fq> in2=<input\_R2\_fastq> out2=<output\_R2\_fq> ref=./bbmap/resources/adapters.fa ktrim=r k=23 mink=11 hdist=1 tpe tbo qtrim=rl trimq=30 minlen=50 maq=30

# in1 and in2 are used to indicate the full path to trim both files. BOTH files must be processed in the same command-line to correctly process each read with its pair

# ktrim=r. Use to trim the adapters; kmers will only come from the right end of the read (3' adapters).

# k=23. “k” specifies the maximum kmer size to use.

# mink=8. Bbduk will additionally look for shorter kmers with lengths 22 to 8 (in this case).

# hdist=1. Controls the hamming distance for all kmers (this allows one mismatch).

# tpe. Use this flag to trim both reads to the same length (in the event that an adapter kmer was only detected in one of them).

# tbo. Use this flag to also trim adapters based on pair overlap detection using BBMerge.

# qtrim=rl. This command will trim both sides of the reads (left and rigth). It happens AFTER all kmer-based operations. REVISAR!!!

# trimq=30. This will quality-trim to Q30 using the Phred algorithm.

# minlen=50. This will discard reads shorter than 50bp after trimming.

# maq=30. This will discard reads with average quality below 30.

#----Bowtie----

bowtie2-build FMDV.fasta FMDVref

# this step builds the index from the reference file

bowtie2 -p 4 --no-mixed --no-discordant -x <Ref\_name> -1 <input\_R1\_fq> -2 <input\_R1\_fq> -S <output\_sam>

# This step will produce the actual alignment

# -p 4. The -p option causes Bowtie 2 to launch a specified number of parallel search threads (for computer with multiple processors/cores).

# --no-mixed. By default, when bowtie2 cannot find a concordant or discordant alignment for a pair, it then tries to find alignments for the individual mates.

# This option disables that behavior.

# --no-discordant. This option disables the search for discordant alignments if it cannot find any concordant alignments.

# A discordant alignment is an alignment where both mates align uniquely, but that does not satisfy the paired-end constraints.

# -x. Indicates the base name of the index for the reference genome.

# -1 and -2 are used to indicate the fastq files with the reads to align.

# -S outputs file name.

#----SamTools----

samtools view -bST <Ref\_fasta> <input\_sam> > <output\_bam>

# This step exports the alignment from SAM format to the BAM format.

# -b. Forces output in the BAM format.

# -S. Ignored for compatibility with previous samtools versions.

# -T. Indicates the FASTA format for the reference file

samtools sort <input\_bam> > <output\_sorted\_bam>

# Sorts the alignment by leftmost coordinates

samtools view -h -F 4 -b <input\_sorted\_bam> > <output\_map\_bam>

# This step discards all reads that do not map to the reference

# -h. Includes the header in the output.

# -F. Filters aligment by a specific flag (include only mapped reads)

# -b. Forces output in the BAM format.

samtools index <input\_map\_bam> <output\_map\_bai>

# This step generates the index file (.bai) of the alignment.

samtools depth -d10000000 <input\_map\_bam> > <output\_map\_coverage\_txt>

# This step (optional) generates a file with the raw coverage at each reference position

# -d. Sets the maximum cutoff for coverage (default is 8000X).

#----LoFreq----

lofreq call-parallel --pp-threads 4 -f <Ref\_fasta> -o <output\_vcf> <input\_map\_bam>

# This step generates a VCF file with the SNV detected from the BAM file.

# --pp-threads 4.The --pp-threads option causes LoFreq to launch a specified number of parallel search threads (for computer with multiple processors/cores).

# -f. Sets the reference file to use.

# -o. Sets the output file name

bgzip <input\_vcf>

# Compress the VCF file.

tabix <input\_vcf\_gz>

# Creates an index file for the VCF file

bcftools stats <input\_vcf\_gz> > <output\_stats\_vchk>

# Generates a text file with different stats from the VCF file

bcftools filter -i "DP>1000" <input\_vcf\_gz> -o <output\_filtered\_vcf>

# Filters the SNV for a specific parameter.

# -i "DP>1000". Flag to filter by (i.e. only keeps calls with raw coverage > 1000X).

bgzip <input\_filtered\_vcf>

tabix <input\_filtered\_vcf.gz>

bcftools filter -i "AF>0.01" <input\_filtered\_vcf.gz> > <output\_filtered\_freq\_vcf>

# Filters the SNV for a specific parameter.

# -i "AF>0.01". Flag to filter by (i.e. only keeps calls with allele frequency > 1%).

----QuRe----

samtools sort -n <input\_map\_bam> > < input\_map\_sorted\_bam>

# Sorts the alignment by read name.

bam2fastx -q -P -N -M -o <Reads\_fq> <input\_map\_sorted\_bam>

#-q. Force fastq output.

#-P. Paired-end data.

#-N. Append /1 and /2 suffixes.

#-M. Output only mapped reads.

cat <Reads.1\_fq> <Reads.2\_fq> > <Reads\_all\_fq>

#Concatenate read files

seqtk seq -A <Reads\_all\_fq> > <Reads\_all\_fasta>

# This step forces fasta format (discards Quality info)

java -cp <path\_to/QuRe> -Xmx32G QuRe <Reads\_all\_fasta> <Ref\_fasta> homopolymericErrorRate nonHomopolymericErrorRate iterations

# Runs Qure software to reconstruct haplotypes

# -cp. Indicates full path to the qure class.

# -Xmx. Controls the maximum amount of memory (RAM) my Java program uses. Default is too low for our dataset.

# homopolymericErrorRate nonHomopolymericErrorRate iterations. If the last three parameters are not inserted, default values are used (0.01, 0.005, 3000).

----CliqueSNV----

java -Xmx13G -jar <path\_to/clique-snv.jar> -m snv-illumina -tf 0.1 -t 1000 -in <input\_map\_bam> -log

# Runs CliqueSNV software to reconstruct haplotypes

# -Xmx. Controls the maximum amount of memory (RAM) my Java program uses. Default is too low for our dataset.

# -jar. Indicates full path to the CliqueSNV jar file.

# -m snv-illumina. Sets specific mode for illumina data.

# -t. Minimum threshold for O22 value. Default is 10 (only for Illumina reads)

# -tf. Minimum haplotype expected frequency.

# -in. Indicates alignment file to use.

----ViQuaS----

Rscript ViQuaS.R <Ref\_fasta> <input\_map\_bam> r o

# Runs ViQuaS software to reconstruct haplotypes

# r is the minimum number of reads needed to call a base during an extension. Default is 3.

# o is the minimum base ratio used to accept an overhang consensus base. Default is 0.7.