

BCoV_nanopore

basecalling from fast5

tool: guppy (version)

```
guppy_basecalled -qscore 7
```

Alignment

tool: #minimap2

ref: U00735.2, NC, GRCh38

```
minimap2 -Y -k 20 -w 1 --splice -g 30000 -G 30000 -F  
40000 -N 32 --splice-flank=no --max-chain-skip=40 -u n --  
MD -a -t 10 --secondary=no [ref] [query]
```

Check the error rate

tools: #Alfred

```
alfred qc -r [ref] -j [out.json] -o [out.tsv] [in.bam]
```

Viral depth

Tool: #samtools

```
samtools depth [in.bam] -a -o [out.txt]
```

Subset read

Tool: #samtools

```
samtools view -N [read_name] [in.bam] -o [out.bam]
```

polishment

Tool: #Transcriptclean

```
python Transcriptclean.py
```

file transformation [4]

tools: #samtools, #bedtools

```
samtools view -b -f 0 | samtools -@ 10 | bedtools  
bamtobed -split > out.bed
```

Read to fasta by bed file [5]

tools: biotools-master #bedTobed12, #bedtools

ref: U00735.2

```
python3 ~/Analysis/tools/biotools-master/bed6Tobed12.py  
[in.bed] > [out.bed12]  
bedtools getfasta -fi [ref] -bed [in.bed12] -fo  
[out.fasta] -name -split
```

Predict the open reading frame and it's translated peptide in R

required files: 4, 5

generated file: [out.RData]

Classified the reads into sgmRNA and DVG by known peptide in R

required files: 4, 5, virus ORF peptide sequence([out.RData])

Grep the soft clip region by cigar string

Alignment and cut the specific columns (read, flag, cigar, seq)

The condition is the same as viral alignment.

```
minimap2 -Y -k 20 -w 1 --splice -g 30000 -G 30000 -f 40000 -N 32 --splice-flank=no --max-chain-skip=40 -u n --MD -a -t 10 --secondary=no [in.fq] [ref.viral] | cut -f1,2,6,10 > [out.txt]
```

Grep the clip region sequence by R(read, flag, cigar, seq)

Filter: Those reads with soft clip being 0.

language-r

Align the clip sequence and cut the aligned ref (read, flag, ref)

```
minimap2 -Y -k 20 -w 1 --MD -a -t 10 --secondary=no [clip.fa] [ref.viral_RCS_human] | cut -f1,2,3 > [out.txt]
```