

- **Virus Identification**  
(Genomad, VS2, CT3)
- **pre-processing analysis output**
- **Recombination analysis output**

- fasta files

## Motif

### Regex

- input: fasta
- motif pattern
- + seqkit

### *de novo* discovery

- input: fasta
- + MEME
- + ScanProsite

```
output/motif/.
├─ logo.pdf
├─ logo_split.pdf
│  # pdf figures
├─ motif.log
├─ pattern_positions.txt
│  # filte containg the pattern positions (seqID patternName pattern strand start end matched)
├─ seq_logo.log
│  # log file
│  # sequences if option (--split or/and --remove-gaps) are selected
├─ ungapped_sequences.fasta
├─ split_sequences_1.fasta
└─ split_sequences_2.fasta
```

```
.
├─ consensus_table.csv
│  # table with the consensus motif sequences (id,consensus,length,occurrences)
├─ logo1.eps
│  # sequence logo in eps format (meme output) for each discovered motif.
├─ MEME-1_pwm_matrix.csv
│  # Position Weight Matrix (PWM) matrix for each discovered motif.
├─ meme.html # meme output
├─ meme.txt # meme output
├─ meme.xml # meme output
├─ motif_discovery.log
│  # log file
├─ motif_table.csv
│  # table with all the motif sequences and attributes (length,motif_name,pvalue
└─ scanprosite_results.csv (if --scanprofile selected)
   # ScanProsite table (sequence_ac,start,stop,signature_ac,score,level,sequenc
```