VirSorter2··········contig·····

·····VirSorter2·····FASTA·····

VirSorter2·····
·····contig·····

·· | ·· | ··

··· | setup····· | ·····

···· | ····· | ·····dsDNAphage·ssDNA

···· | 1500 nt··· | ·····

·· | ·····work·· | ·····

·····

·····VirSorter2 setup·····FASTA·····HMM·····setup·····
·····

·····

SAMPLES·····TAB·····ID·contig FASTA···VIRAL_GROUPS·····dsDNApha
ge,ssDNA·····VirSorter2 2.2·····

```
#!/usr/bin/env bash
set -euo pipefail
: "${DB:?set the prepared VirSorter2 database directory}"
: "${SAMPLES:?set a two-column sample_id<TAB>contigs_fasta manifest}"
: "${VIRAL_GROUPS:?set groups, e.g. dsDNAphage,ssDNA}"
: "${THREADS:?set allocated threads}"
: "${OUT:?set a new output directory}"
[[ "$THREADS" =~ ^[1-9][0-9]*$ ]] || exit 2
test -d "$DB" && test -s "$SAMPLES"
test ! -e "$OUT" && test ! -L "$OUT" || exit 2
mkdir -p "$OUT"
virsorter run -h > "$OUT/help.txt" 2>&1
completed=0
while IFS=$'\t' read -r sample contigs || [[ -n "${sample}${contigs}" ]]; do
  [[ "$sample" =~ ^[A-Za-z0-9_-]+$ ]] || exit 2
  test -s "$contigs" || exit 2
  root="$OUT/$sample"
  mkdir "$root"
  sha256sum "$contigs" > "$root/input.sha256"
  /usr/bin/time -v -o "$root/resources.txt" \
    virsorter run -w "$root/work" -i "$contigs" --db-dir "$DB" \
    --include-groups "$VIRAL_GROUPS" --min-length 1500 -j "$THREADS" all \
    > "$root/run.log" 2>&1
  test -f "$root/work/final-viral-combined.fa"
  test -s "$root/work/final-viral-score.tsv"
  python3 check_candidate_ids.py "$root/work/final-viral-combined.fa" \
    > "$root/candidate-count.txt"
  completed=$((completed+1))
done < "$SAMPLES"
test "$completed" -gt 0
```

·····check_candidate_ids.py·····FASTA···ID·····
·····score·····

```
import sys
with open(sys.argv[1]) as handle:
    ids=[line[1:].split()[0] for line in handle if line.startswith(">")]
if len(ids)!=len(set(ids)):
    raise SystemExit("duplicate candidate FASTA IDs")
print("candidate records",len(ids))
```

·····

final-viral-combined.fa·final-viral-score.tsv·final-viral-boundary.tsv·····
·····seqname·····contig·····full·partial·lt2gene·····
·····

·····CheckV·····
·····
.

·····

VirSorter2·····contig·····
·····RAM·····I/O·····

……Snakemake……
……
……

….

VirSorter2··README····setup····· https://github.com/jiarong/VirSorter2/blob/master/README.md

VirSorter2····FASTA https://github.com/jiarong/VirSorter2/blob/master/test/8seq.fa