Cheatography

seqkit Cheat Sheet

by ac1513 via cheatography.com/78065/cs/19074/

grep

-f	pattern file
-p	search pattern
-v	invert (non-matching)
-r	patterns use regular expression
-n	by name
-s	by seq
-i	ignore case

search sequences by pattern(s) of name or sequence motifs

fx2tab

-n	print names
-i	print id (instead of full header)
-g	print gc content
-G	print gc-skew
-l	print length
-B	print base content (e.g. -B AT -B N)

sort

-l	by length
-n	by full name (not id)
-s	by sequence
-r	reverse

stats

-a	all
-T	tabular (machine readable
-j	number of threads

split

-i	by id
-p	into # parts
-s	by size
-O	output directory (def. is $infile.split)

use split2 for fastq/paired end (-1 + -2 for paired end)

seq

-m	min. length of reads to output e.g. 500 for reads over 500bp
-M	max. length of reads to output e.g. 500 for reads under 500bp
-n	only print read names
-w	defines line width, 0 for no wrap (i.e. to turn into one-line fastx)
-i	print ID instead of full head (shorten ID)

other

faidx	Create fasta index file
fq2a	fastq to fasta
rmdup	remove duplicated sequences

By **ac1513**
cheatography.com/ac1513/

Published 2nd September, 2020.
Last updated 22nd March, 2022.