seqkit Cheat Sheet

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

Cheatography	y
--------------	---

grep		
-f	pattern file	
-p	-p search pattern	
-V	invert (non-matching)	
-r	-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		

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

split	
-i	by id
-р	into # parts
-S	by size
-0	output directory (def. is \$infile.split)

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

-n print names -i print id (instead of full header)

- print gc content -g -G print gc-skew
- -| print length
- -B print base content (e.g. -B AT -B N)

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)	

-1	print length	
-B	print base content (e.gB AT -B N)	-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)
sort		
-1	by length	other
-n	by full name (not id)	faidx Create fasta index file
-S	by sequence	fq2a fastq to fasta
-r	reverse	rmdup remove duplicated sequences

By ac1513

cheatography.com/ac1513/

Published 2nd September, 2020. Last updated 22nd March, 2022. Page 1 of 1.

Sponsored by Readable.com Measure your website readability! https://readable.com