

jellyfish

Opis

Jellyfish je alat za brzo, memorijski učinkovito brojanje k-mera u DNK. K-mer je podniz duljine k, a brojanje pojavljivanja svih takvih podnizova središnji je korak u mnogim analizama sekvence DNK. Jellyfish može brojati k-mere koristeći red veličine manje memorije i red veličine brže od ostalih paketa za brojanje k-mer koristeći učinkovito kodiranje hash tablice i iskorištavanjem CPU instrukcije "usporedi i zamijeni" za povećanje paralelizam.

JELLYFISH je program naredbenog retka koji čita FASTA i multi-FASTA datoteke koje sadrže sekvence DNK. Izbacuje svoje k-mer brojeve u binarnom formatu, koji se može prevesti u tekstualni format čitljiv ljudima pomoću naredbe "jellyfish dump" ili postaviti upit za određene k-mere pomoću "jellyfish query".

Verzije

Verzija	Modul	Supek	Padobran
2.3.1	scientific/jellyfish/2.3.1		

Službena dokumentacija

fasta datoteka korištena za primjer: http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/human_g1k_v37.fasta.gz

Dokumentacija : <https://github.com/gmarcais/Jellyfish/tree/master/doc#query>

Primjer

Brojanje k-mera(Padobran)

```
#PBS -N k-mer_counting
#PBS -q cpu
#PBS -l select=1:ncpus=60

cd ${PBS_O_WORKDIR}

module load scientific/jellyfish/2.3.1

jellyfish.sh jellyfish count -m 21 -s 200M -t $NCPUS -C human_g1k_v37.fasta
```