

# salmon



## Opis

Salmon je alat za brzu kvantifikaciju prijepisa iz RNA-seq podataka. Za kvantificiranje je potreban skup ciljnih prijepisa (bilo iz referentnog ili de-novo sklopa). Sve što trebate za pokretanje Salmona je FASTA datoteka koja sadrži vaše referentne transkripte i (skup) FASTA/FASTQ datoteka(a) koje sadrže vaša čitanja. Po izboru, Salmon može koristiti unaprijed izračunata poravnanja (u obliku SAM /BAM datoteke) transkripata umjesto sirovih čitanja.

Način Salmona koji se temelji na kartiranju odvija se u dvije faze; indeksiranje i kvantificiranje. Korak indeksiranja neovisan je o čitanjima i potrebno ga je pokrenuti samo jednom za određeni skup referentnih prijepisa. Korak kvantifikacije je očito specifičan za skup očitavanja RNA-seq i stoga se izvodi češće.

## Verzije

Verzija	Modul	Supek	Padobran
1.10.2	scientific/salmon/1.10.2	✓	✓

## Službena dokumentacija

<https://salmon.readthedocs.io/en/latest/index.html>

Primjer:

- <https://combine-lab.github.io/alevin-tutorial/2019/selective-alignment/>
- [ftp://ftp.ensemblgenomes.org/pub/plants/release-28/fasta/arabidopsis\\_thaliana/cdna/Arabidopsis\\_thaliana.TAIR10.28.cdna.all.fa.gz](ftp://ftp.ensemblgenomes.org/pub/plants/release-28/fasta/arabidopsis_thaliana/cdna/Arabidopsis_thaliana.TAIR10.28.cdna.all.fa.gz) -o athal.fa.gz

## Primjer

- Alociranje 8-12 threadova za Salmon postiže najviše brzine za izračune. Odabir threadova više od 12 će rezultirati da ti threadovi budu zauzeti i neiskorišteni.



Zbog tehničkih poteškoća Salmon-a sa radom na Padobranu, privremeno je dopušteno korištenje samo na Supeku.

## Supek

```
#!/bin/bash

#PBS -N salmon-test
#PBS -l select=1:ncpus=12:mem=20GB
#PBS -q cpu

cd $PBS_O_WORKDIR

module load scientific/salmon/1.10.2

salmon.sh salmon index -t athal.fa.gz -i athal_index --threads $NCPUS
```