

STAR

Opis

STAR, što znači "Spliced Transcripts Alignment to a Reference" (Usporedba Izrezanih Transkripata s Referencom), popularan je softverski alat u području bioinformatike. STAR se koristi prije svega za usklađivanje RNA-Seq podataka, tehnike sekvenciranja visokog protoka koja se koristi za proučavanje izražavanja gena i profila transkriptoma. Razvijen je kako bi učinkovito i točno mapirao RNA-Seq očitavanja na referentni genom, identificirao mjesto spajanja između izreza i kvantificirao razine izražavanja gena. Evo nekih ključnih značajki i funkcija softvera STAR:

- **Brzo i Točno Uspoređivanje:** STAR je poznat po svojoj visokoj brzini usklađivanja i preciznosti. Koristi dvostupanjski proces, uključujući usklađivanje sjemena i usklađivanje mjesta spajanja, kako bi učinkovito obrađivao poznate i nove spojne točke.
- **Podrška za Različite Vrste Podataka:** STAR može uskladiti različite vrste RNA-Seq podataka, uključujući očitavanja jednostrukog kraja i očitavanja parnog kraja, kao i podatke s različitih platformi za sekvenciranje (Illumina, Ion Torrent itd.).
- **Otkrivanje Mjesta Spajanja:** STAR je posebno dobar u otkrivanju mjesta spajanja i preciznom usklađivanju očitavanja preko egzonskih granica, što je važno za identificiranje alternativnih splicing događaja i novih transkripata.
- **Kvantifikacija Izražavanja Gena:** Softver kvantificira razine izražavanja gena brojanjem broja očitavanja koja se usklađuju s svakim genom ili transkriptom, omogućavajući istraživačima točno mjerenje izražavanja gena.
- **Spajanje Kimeričkih Izraza:** STAR također može identificirati kimeričke fuzijske događaje, što je važno u istraživanju raka za otkrivanje fuzija gena i rearanžmana.
- **Višenitno i Paralelno Procesiranje:** STAR je dizajniran kako bi iskoristio više CPU jezgara, što ga čini prikladnim za okoline visokih performansi (HPC).
- **Prilagodljiv:** Korisnici mogu konfigurirati različite parametre i opcije kako bi prilagodili STAR specifičnim eksperimentalnim postavkama i referentnim genomima.
- **Otvorenog Koda:** STAR je otvoreni softver, a izvorni kôd je besplatno dostupan za prilagodbu i izmjene.

STAR je postao široko prihvaćeni alat u analizi RNA-Seq zbog svoje brzine i preciznosti, što ga čini važnim dijelom bioinformatičkog procesa za proučavanje izražavanja gena, alternativnog spajanja i drugih aspekata transkriptomike. Istraživači ga koriste za obradu sirovih RNA-Seq podataka, usklađivanje očitavanja s referentnim genomom i kvantificiranje razina izražavanja gena, što je ključno za mnoga genomika i transkriptomika istraživanja.

Dostupne verzije

Verzije	Modul	Supek	Padobran
2.7.11a	scientific/STAR/2.7.11a		

Primjer korištenja

- Službena dokumentacija: <https://github.com/alexdobin/STAR/blob/master/doc/STARmanual.pdf>
- Podaci za primjer: https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001735.4/

```
STAR.pbs

#!/bin/bash

#PBS -N STAR-ex
#PBS -l select=1:ncpus=32:mem=10GB
#PBS -j oe

cd $PBS_O_WORKDIR

module load scientific/STAR/2.7.11a

STAR --runThreadN $NCPUS --runMode genomeGenerate --genomeDir index --genomeFastaFiles GCF_000001735.4_TAIR10.1_genomic.fna --sjdbGTFfile genomic.gff --sjdbOverhang 149
```