

gatk4



Opis

Genome Analysis Toolkit (GATK) softverski je paket razvijen na Institutu Broad za analizu podataka sekvenciranja visoke propusnosti.

Skup alata uključuje širok izbor alata, s fokusom na otkrivanje varijanti i genotipizaciju, kao i naglasak na osiguranje kvalitete podataka.

Verzije

Verzija	Modul	Supek	Padobran
4.4.0.0	scientific/gatk4/4.4.0.0	✗	✓

Službena dokumentacija

- Dostupni alati: [gatk-alati](#)
- Git: <https://github.com/broadinstitute/gatk>
- Službena stranica: <https://gatk.broadinstitute.org/hc/en-us>

Primjer korištenja

Padobran

```
#!/bin/bash

#PBS -N gatk4-HTcaller
#PBS -q cpu
#PBS -l select=1:ncpus=1:mem=10GB

cd $PBS_O_WORKDIR

module load scientific/gatk4/4.4.0.0

gatk4.sh gatk --java-options "-Xmx10G" HaplotypeCaller -R reference.fasta -I input.bam -O output.vcf
```



S obzirom da veliki broj alata nema mogućnost kontrole broja jezgara koji se koristi, poslovi se šalju kao serijski. Kod korištenja alata koji imaju mogućnost korištenja više dretvi ili više jezgara bitno je navoditi da se koristi jednak broj jezgara koji je i tražen u skripti za podnošenje poslova (\$NCPUS).