

BWA + SAMTOOLS_(2)

2022/02/10

20193852 문유빈

2-3> Index, Idxstats, Flagstat

```
echo "indexing"
$ samtools index ${file}/megahit/readsMapping/bam/${stub}_paired.bam
echo "contig stats"
$ samtools idxstats ${file}/megahit/readsMapping/bam/${stub}_paired.bam > ${file}/megahit/readsMapping/stat/${stub}_stat.contig.txt
echo "mapping stat and genome mapped coverage calculation"
$ samtools flagstat ${file}/megahit/readsMapping/bam/${stub}_paired.bam > ${file}/megahit/readsMapping/stat/${stub}_stat.txt
done
```

① index

```
/home/bioware/samtools-1.8/samtools
```

index

```
/home/guest01/2021/yb/yb01/ybMegahit/readsMapping/bam/PG-16-sps-01_paired.bam
```

```
[guest01@smell:bam]$ ll
total 2737248
-rw-rw-r-- 1 guest01 guest01 2746200573 Feb  9 19:10 PG-16-sps-01_paired.bam
-rw-rw-r-- 1 guest01 guest01  56727568 Feb 10 14:11 PG-16-sps-01_paired.bam.bai
```

=> 완료시 bam 폴더 안에 bam.bai 파일 생성됨

[illegible]

=> 바이너리 파일이여서 우리가 해석 x 컴퓨터는 알아 들음

② idxstats

```
/home/bioware/samtools-1.8/samtools idxstats
/home/guest01/2021/yb/yb01/ybMegahit/readsMapping/bam/PG-16-sps-01_paired.b
am >
/home/guest01/2021/yb/yb01/ybMegahit/readsMapping/stat/PG-16-sps-01_stat.conti
g.txt
```

```
[guest01@smell:stat]$ ll
total 14720
-rw-rw-r-- 1 guest01 guest01 15065533 Feb 10 14:14 PG-16-sps-01_stat.contig.txt
-rw-rw-r-- 1 guest01 guest01 452 Feb 10 14:20 PG-16-sps-01_stat.txt
```

=> 완료시 stat 폴더 안에 stst.contig.txt 파일 생성됨

```
k141_2 515 16 0
k141_3 648 14 0
k141_4 512 13 0
k141_7 567 10 0
k141_9 761 19 0
k141_10 825 12 0
k141_11 814 25 0
k141_12 814 21 0
k141_13 974 29 0
k141_15 1098 23 0
k141_17 540 17 0
k141_18 1177 27 0
k141_19 607 14 0
k141_20 827 28 0
k141_22 528 9 0
k141_23 867 22 0
k141_24 512 12 0
k141_26 988 28 0
k141_27 519 28 0
k141_28 1189 46 0
k141_29 1179 37 0
```

=> stat.contig.txt 파일

** The output is TAB-delimited with each line consisting of reference sequence name, sequence length, # mapped read-segments and # unmapped read-segments.

** 즉 참조 시퀀스 이름, 시퀀스 길이, # 맵핑된 리드-조각, # 맵핑되지 않은 리드-조각

③ flagstat

```
/home/bioware/samtools-1.8/samtools flagstat
/home/guest01/2021/yb/yb01/ybMegahit/readsMapping/bam/PG-16-sps-01_paired.b
am >
/home/guest01/2021/yb/yb01/ybMegahit/readsMapping/stat/PG-16-sps-01_stat.txt
```

```
[guest01@smell:stat]$ ll
total 14716
-rw-rw-r-- 1 guest01 guest01 15065533 Feb 10 14:14 PG-16-sps-01_stat.contig.txt
```

=> 완료시 stat 폴더 안에 stst.txt 파일 생성됨

```
34856726 + 0 in total (QC-passed reads + QC-failed reads)
0 + 0 secondary
948696 + 0 supplementary
0 + 0 duplicates
34856726 + 0 mapped (100.00% : N/A)
33908030 + 0 paired in sequencing
16988376 + 0 read1
16919654 + 0 read2
28375960 + 0 properly paired (83.69% : N/A)
32766378 + 0 with itself and mate mapped
1141652 + 0 singletons (3.37% : N/A)
4351922 + 0 with mate mapped to a different chr
2969674 + 0 with mate mapped to a different chr (mapQ<=5)
```

** stat.txt 파일

** 여기 flagstat에 대해 잘 설명 돼 있음

<https://hhj6212.github.io/biology/tech/2021/06/13/samtools-flagstat.html>