

Drugi domaći zadatak iz predmeta Genomska Informatika

Student: Dunja Đorđević 3346/20

```
In [1]: import pysam

samfile = pysam.AlignmentFile("merged-tumor.bam", "rb")

total_count = 0
count_of_unmapped = 0
count_of_maq_zero = 0
sum_of_maqs = 0

iter = samfile.fetch(until_eof=True)

for x in iter:
    total_count += 1
    sum_of_maqs += x.mapping_quality
    if x.is_unmapped:
        count_of_unmapped += 1
    if (x.mapping_quality == 0):
        count_of_maq_zero += 1

print("Broj unmapped readova u BAM fajlu je: ")
print(count_of_unmapped)

print("Broj read-ova u BAM fajlu je: ")
print(total_count)

print("Broj readova sa kvalitetom mapiranja jednakim nuli u BAM fajlu je: ")
print(count_of_maq_zero)

print("Prosecni kvalitet mapiranja (za sve read-ove): ")
print(sum_of_maqs/total_count)

print("Prosecni kvalitet mapiranja (samo za read-ove koji imaju ne nula MAQ): ")
print(sum_of_maqs/(total_count - count_of_maq_zero))
```

```
Broj unmapped readova u BAM fajlu je:
17765
Broj read-ova u BAM fajlu je:
2921629
Broj readova sa kvalitetom mapiranja jednakim nuli u BAM fajlu je:
126628
Prosecni kvalitet mapiranja (za sve read-ove):
55.91379158681681
Prosecni kvalitet mapiranja (samo za read-ove koji imaju ne nula MAQ):
58.446975510921106
```

```
In [2]: samfile = pysam.AlignmentFile("merged-tumor.bam", "rb")
iter = samfile.fetch(until_eof=True)

for x in iter:
    print(x)
    print("\nFlag polje za prvi read je: ")
```

```
print (x.flag)
break
```

[illegible]

```
Flag polje za prvi read je:
1187
```

Na osnovu dobijenog flag-a za prvi read u BAM fajlu može se zaključiti sledeće za posmatrani read:

Summary:

```
read paired (0x1)
read mapped in proper pair (0x2)
mate reverse strand (0x20)
second in pair (0x80)
read is PCR or optical duplicate (0x400)
```