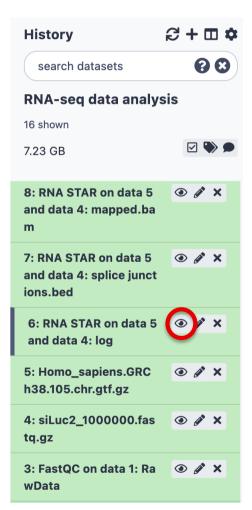
Analysis of RNA-seq data: answers to questions

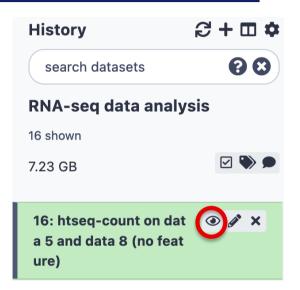
Number of uniquely mapped reads

```
Started job on
                                                       Feb 04 09:10:08
                           Started mapping on
                                                       Feb 04 09:14:43
                                                       Feb 04 09:14:54
                                  Finished on
                                                       327.27
     Mapping speed, Million of reads per hour
                        Number of input reads
                                                       1000000
                                                       50
                    Average input read length
                                  UNIQUE READS:
                                                      852434
                 Uniquely mapped reads number
                      Uniquely mapped reads %
                        Average mapped length
                                                       49.84
                     Number of splices: Total
                                                       137459
          Number of splices: Annotated (sidb)
                                                       136335
                     Number of splices: GT/AG
                                                       136060
                     Number of splices: GC/AG
                                                       1157
                     Number of splices: AT/AC
                                                       108
             Number of splices: Non-canonical
                                                       134
                    Mismatch rate per base, %
                                                       0.15%
                       Deletion rate per base
                                                       0.01%
                      Deletion average length
                                                       1.60
                      Insertion rate per base
                                                       0.00%
                     Insertion average length
                                                       1.29
                           MULTI-MAPPING READS:
      Number of reads mapped to multiple loci
                                                       133958
           % of reads mapped to multiple loci
                                                       13.40%
      Number of reads mapped to too many loci
                                                       4067
           % of reads mapped to too many loci
                                                       0.41%
                                UNMAPPED READS:
Number of reads unmapped: too many mismatches
     % of reads unmapped: too many mismatches
                                                       0.00%
          Number of reads unmapped: too short
                                                       7302
               % of reads unmapped: too short
                                                       0.73%
              Number of reads unmapped: other
                                                       2239
                   % of reads unmapped: other
                                                       0.22%
                                CHIMERIC READS:
                     Number of chimeric reads
                          % of chimeric reads
                                                       0.00%
```



Category	RNA STAR on data 5 and data 4: mapped.bam
no_feature	67657
ambiguous	32425
too_low_aQual	0
not_aligned	13608
alignment_not_unique	450475

- No feature reads
 - Number
 - **67657**
 - Proportion :
 - 67657*100/852434 = 7.94
- Ambiguous reads
 - Number
 - **32425**
 - Proportion
 - **32425*100/852434 = 3.80**



■ Proportion of reads among uniquely aligned reads

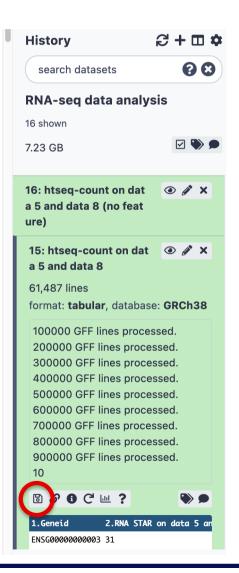
Assigned : 100-7.94-3.80 = 88.26 %

No feature : 7.94 %

■ Ambiguous : 3.80 %

Number of assigned reads

Geneid	RNA STAR on data 5 and data 4: mapped.bam
ENSG00000000003	31
ENSG0000000005	0
ENSG00000000419	95
ENSG00000000457	18
ENSG00000000460	55
ENSG00000000938	0
ENSG00000000971	3
ENSG0000001036	66
ENSG0000001084	50
ENSG0000001167	43
ENSG0000001460	6
ENSG00000001461	18
ENSG0000001497	72
ENSG00000001561	2
ENSG00000001617	2
ENSG00000001626	0
ENSG00000001629	53
ENSG0000001630	5
ENSG00000001631	5
ENSG00000002016	7
ENSG00000002079	0
ENSG00000002330	27
ENSG00000002549	68
ENSG00000002586	123
ENSG00000002587	1
ENSG00000002726	0
ENSG00000002745	0



- Number of assigned reads
 - Open the downloaded file with excel
 - Calculate the total number of reads in the second column

B61488 $\stackrel{\blacktriangle}{\checkmark}$ \times \checkmark f_X =SOMME(B1:B61487)						
	A		В	С	D	
61477	ENSG0000	0289634	0			
61478	ENSG0000	00289635	0			
61479	ENSG0000	00289636	0			
61480	ENSG0000	0289637	0			
61481	ENSG0000	0289638	0			
61482	ENSG0000	0289639	0			
61483	ENSG0000	00289640	0			
61484	ENSG0000	00289641	0			
61485	ENSG0000	0289642	0			
61486	ENSG0000	0289643	0			
61487	ENSG0000	0289644	0			
61488			752352			

- → Number of assigned reads = 752352
- → Proportion of assigned reads = 752352 *100/852434 = 88.26

Number of assigned reads

- = number of uniquely aligned reads number of no feature reads number of ambiguous reads
- = 852434 67657 32425 = 752352