

Feedback — Module 2 Exam

[Help Center](#)

Thank you. Your submission for this exam was received.

You submitted this exam on **Sat 19 Sep 2015 3:31 PM PDT**. You got a score of **20.00** out of **20.00**.

For this project, it is recommended that you use the VMBox virtual environment provided with the Course package and the tools therein. You may also use your own system and software, however make sure that appropriate versions are installed. The answers are compatible with the following versions of the software: samtools v.1.2, bedtools v.2.24.0.

As part of a larger project cataloging genetic variation in the plant *Arabidopsis thaliana*, you sequenced and assembled the genome of one strain ('wu_0_A'), then mapped back the reads to the assembled genome. The resulting BAM file is included in the package 'gencommand_proj2_data.tar.gz'. Using SAMtools and BEDtools as well as other Unix commands introduced in this course, examine the files and answer the following questions. NOTE: Input data have been obtained and modified from those generated by the 1001 Genomes Project, accession 'Wu_0_A'.

[Click here to download the Project 2 Data Files](#)

Apply these rules and steps to the questions marked above each rule.

Questions 1 - 5:

For the original set of alignments (file 'athal_wu_0_A.bam'):

Questions 6 - 10:

Extract only the alignments in the range "Chr3:11,777,000-11,794,000", corresponding to a locus of interest. For this alignment set:

Questions 11 - 15:

Determine general information about the alignment process from the original BAM file.

Questions 16 - 20:

Using BEDtools, examine how many of the alignments at point 2 overlap exons at the locus of interest. Use the BEDtools '-wo' option to only report non-zero overlaps. The list of exons is given in the included 'athal_wu_0_A_annot.gtf' GTF file.

Question 1

How many alignments does the set contain?

You entered:

Your Answer		Score	Explanation
221372	✓	1.00	
Total		1.00 / 1.00	

Question 2

How many alignments show the read's mate unmapped?

You entered:

Your Answer		Score	Explanation
65521	✓	1.00	
Total		1.00 / 1.00	

Question 3

How many alignments contain a deletion (D)?

You entered:

Your Answer		Score	Explanation
2451	✓	1.00	
Total		1.00 / 1.00	

Question 4

How many alignments show the read's mate mapped to the same chromosome?

You entered:

Your Answer		Score	Explanation
150913	✓	1.00	
Total		1.00 / 1.00	

Question 5

How many alignments are spliced?

You entered:

Your Answer		Score	Explanation
0	✓	1.00	
Total		1.00 / 1.00	

Question 6

How many alignments does the set contain?

You entered:

7081

Your Answer	Score	Explanation
7081	✓ 1.00	
Total	1.00 / 1.00	

Question 7

How many alignments show the read's mate unmapped?

You entered:

1983

Your Answer	Score	Explanation
1983	✓ 1.00	
Total	1.00 / 1.00	

Question 8

How many alignments contain a deletion (D)?

You entered:

31

Your Answer		Score	Explanation
31	✓	1.00	
Total		1.00 / 1.00	

Question 9

How many alignments show the read's mate mapped to the same chromosome?

You entered:

Your Answer		Score	Explanation
4670	✓	1.00	
Total		1.00 / 1.00	

Question 10

How many alignments are spliced?

You entered:

Your Answer		Score	Explanation
0	✓	1.00	
Total		1.00 / 1.00	

Question 11

How many sequences are in the genome file?

You entered:

Your Answer		Score	Explanation
7	✓	1.00	
Total		1.00 / 1.00	

Question 12

What is the length of the first sequence in the genome file?

You entered:

Your Answer		Score	Explanation
29923332	✓	1.00	
Total		1.00 / 1.00	

Question 13

What alignment tool was used?

You entered:

Your Answer		Score	Explanation
stampy	✓	1.00	
Total		1.00 / 1.00	

Question 14

What is the read identifier (name) for the first alignment?

You entered:

GAI05_0002:1:113:7822:3886#0

Your Answer		Score	Explanation
GAI05_0002:1:113:7822:3886#0	✓	1.00	
Total		1.00 / 1.00	

Question 15

What is the start position of this read's mate on the genome? Give this as 'chrom:pos' if the read was mapped, or "*" if unmapped.

You entered:

Chr3:11700332

Your Answer		Score	Explanation
Chr3:11700332	✓	1.00	
Total		1.00 / 1.00	

Question 16

How many overlaps (each overlap is reported on one line) are reported?

You entered:

Your Answer		Score	Explanation
3101	✓	1.00	
Total		1.00 / 1.00	

Question 17

How many of these are 10 bases or longer?

You entered:

Your Answer		Score	Explanation
2899	✓	1.00	
Total		1.00 / 1.00	

Question 18

How many alignments overlap the annotations?

You entered:

Your Answer		Score	Explanation
3101	✓	1.00	
Total		1.00 / 1.00	

Question 19

Conversely, how many exons have reads mapped to them?

You entered:

Your Answer		Score	Explanation
21	✓	1.00	
Total		1.00 / 1.00	

Question 20

If you were to convert the transcript annotations in the file “athal_wu_0_A_annot.gtf” into BED format, how many BED records would be generated?

You entered:

Your Answer		Score	Explanation
4	✓	1.00	
Total		1.00 / 1.00	

