

Module 2 Quiz

7/10 points (70%)

Quiz, 10 questions

✓ Congratulations! You passed![Next Item](#)1 / 1
points

1.

Which of the following strings cannot denote a DNA sequence:

☐

AGCTACTACGAGCT

**Un-selected is correct**☐

AAAAAAAAA

**Un-selected is correct**☐

CCCCCCCCC

**Un-selected is correct**☐

APCTSYFPEITHI

**Correct**1 / 1
points

2.

How many lines does it take to specify:

Module 2 Quiz

i) one fasta sequence? and ii) one fastq sequence?

7/10 points (70%)

Quiz, 10 questions

Select the best answer.

- ☐ Fasta – 1 line; fastq – 2 lines
- ☐ Fasta – 2 lines; fastq – 2 lines
- ☒ Fasta – a fasta header followed by any number of sequence lines; fastq – 4 lines

Correct



1 / 1
points

3.

Which of the following is incorrect:

- ☐ The CIGAR alignment field can contain insertions, deletions, matches or substitutions, and hard and soft clipping.
- ☐ Introns are represented with Ns in a CIGAR string.
- ☐ The symbols H and S in the CIGAR field of a SAM record represent 'hard' and 'soft' clipping, respectively.
- ☒ SAMtools can be used to convert data from SAM to BED format and vice-versa.

Correct



1 / 1
points

4.

Which of the following is NOT an alignment operation:

- ☐ Padding



Cut and paste

Module 2 Quiz

Correct

7/10 points (70%)

Quiz, 10 questions



Soft clipping



Hard clipping

1 / 1
points

5.

What is the minimum number of columns that are sufficient to specify a BED format?



8



3



Correct



12



9

1 / 1
points

6.

Which of the following represents the most accurate conversion into BED of the GTF record:

```
1   '''  
2   chr1  CLASS exon  516 811 100 + . gene_id "genA"; transcript_id "genA.1";  
3   chr1  CLASS exon  1001 1115 100 + . gene_id "genA"; transcript_id "genA.1";  
4   chr1  CLASS exon  3010 3312 100 + . gene_id "genA"; transcript_id "genA.1"  
5   '''
```



```
1   '''  
2   chr1  515 3312  genA.1  100 + 515 3312  0 3 296,115,303 0,485  
3   ''',2494
```



Correct



Module 2 Quiz

Quiz, 10 questions

7/10 points (70%)

```
1   ```
2   chr1 516 3312 genA + 516 3312 0 2 296,303 0,2494
3   ```
```

```
1   ```
2   chr1 515 811 genA 100 + . 800 811 0 1 296 0
3   ```
```

```
1   ```
2   chr1 516 3312 genA.1 100 + 800 900 0 3 296,115,303 0,485,2494
3   ```
```

0 / 1
points

7.

Determine the number of genes, transcripts, exons per transcript, gene orientation (strand), and the length of 5' most exon(s) from the GTF snippet below. Select the correct answer.

```
1   ```
2   chr1 HAVANA gene 3205901 3671498 . - . gene_id "MUSG51951.5";
3   chr1 HAVANA transcript 3205901 3216344 . - . gene_id "MUSG51951.5";
4   transcript_id "MUST162897.1";
5   chr1 HAVANA exon 3213609 3216344 . - . gene_id "MUSG51951.5"; transcript_id
6   "MUST162897.1";
7   chr1 HAVANA exon 3205901 3207317 . - . gene_id "MUSG51951.5"; transcript_id
8   "MUST162897.1";
9   chr1 HAVANA transcript 3206523 3215632 . - . gene_id "MUSG51951.5";
10  transcript_id "MUST159265.1";
11  chr1 HAVANA exon 3213439 3215632 . - . gene_id "MUSG51951.5"; transcript_id
12  "MUST159265.1";
13  chr1 HAVANA exon 3206523 3207317 . - . gene_id "MUSG51951.5"; transcript_id
14  "MUST159265.1";
15  ```
```



Genes: 1; Transcripts: 2; Exons: 2,2; Strand: -; Length of 5' exon(s): 2736, 2194.



Genes: 1; Transcripts: 2; Exons: 2,2; Strand: -; Length of 5' exon(s): 2735, 2193.



This should not be selected



Genes: 1; Transcripts: 1; Exons: 4; Strand: -; Length of 5' exon(s): 2736.



Genes: 1; Transcripts: 4; Exons: 1,1,1,1; Strand: -; Length of 5' exon(s): 2736, 1417,2194,795.

0 / 1
points

8.


Module 2 Quiz

7/10 points (70%)

```

1  ...
2  R1 83 chr12 9232390 255 50M = 9232180 0
3  ATGGCAGAGCCTAATATGTCTCCTAGAGAATGGGAGAGATGGGAAGTCAT
   HGHHHHHHHHHHHHHHHHHHHHHHIGIIIIHHHHHHHHHHHHHHF NM:i:0 NH:i:1 HI:i:0
4  R2 97 chr12 9232391 255 28M278N22M = 9242529
5  0 TGGCAGAGCCTAATATGTCTCCCAAACTGAGACAGAAGCTCGGCAGAT D
   >DDHHHHHHHHHHHHIHHHHHHHHHHHIGFFGGHHHHHHHHHHFB.F NM:i:4 NH:i:3 HI:i:0 XS:A
   :+ NS:i:2
6  R3 77 * 0 0 0 * * 0 0 CTGATATGAGGAAAGAGGATTGCTTAAGCCCAGGAGGTAGAGGCTGTACC
   @@FFDFHFHFHHJJJJJIEGFIGHHHHHIIGCDE?D?FGGBHDGGG
7  ...

```

- ☐ R1 has an exact match to the genome.
- ☐ R2 maps in 3 places within the genome.
- ☒ R1, R2 and R3 all have length 50.
- 
- This should not be selected**
- ☐ R2 has an exact match to the genome.

This should not be selected



0 / 1
points

9.

For the alignment below, which statements are FALSE? The binary encoding for 97 is 97₁₀ = 0000 0110 0001₂. Select all answers that apply.

```
1   ...  
2   R2_97_chr12_9232391_255_28M278N22M = 9242529  
3   0 TGGCAGAGCCTAATATGTCTCCCAAACTGAGACAGAAGCTCGGGCAGAT D  
    >DDDDHHHHHHHHHHIHHHHHHIHHHHIGFFGGGHHHHHHHHHHHFB.F NM:i:4 XS:A:+ NS:i:2  
4   ...
```

- The alignment represents a potential PCR or optical duplicate.

Correct

- ☒ Both the read and its mate are mapped.

This should not be selected

- ☐ The read matches to the genome with 4 differences.

Un-selected is correct



The two mates are identical in sequence.

Module 2 Quiz

Correct

7/10 points (70%)

Quiz, 10 questions



The sequence of the read's mate is reverse complemented in its alignment.

Un-selected is correct



The alignment passes quality checks.

Un-selected is correct



1 / 1
points

10.

Files 'A.bed' and 'B.bed' contain the following sets of intervals:

```

1  ``
2  File A           File B
3  chr1  100 400      chr1  300 500
4  chr1 1000 1400      chr1  900 1600
5  chr1 2000 2400      chr12 2000 2200
6  ``
7  What would be the answers for the following sequence of commands:
8  ``
9  bedtools intersect -wao -a A.bed -b B.bed | sort -u | wc -l
10 bedtools intersect -wo -a A.bed -b B.bed | cut -f1-3 | sort -u | wc -l
11 bedtools intersect -wo -a A.bed -b B.bed | cut -f4-6 | sort -u | wc -l
12 ``

```



9, 3, 4



3, 6, 6



9, 3, 6



4, 2, 6



9, 6, 3



3, 2, 2

Correct



9, 3, 3

☐ 5, 5, 5

Module 2 Quiz

☐ 9, 2, 2

7/10 points (70%)

Quiz, 10 questions

☐ 3, 6, 3

☐ 3, 4, 2

☐ 3, 2, 6

☐ 3, 2, 3

☐ 4, 2, 3

