

## Module 2 Quiz

**9/10 points (90%)**

Quiz, 10 questions

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

1.

Which of the following strings cannot denote a DNA sequence:

☐

AAAAAAAAA

**Un-selected is correct**☐

APCTSYFPEITHI

**Correct**☐

CCCCCCCCC

**Un-selected is correct**☐

AGCTACTACGAGCT

**Un-selected is 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?

9/10 points (90%)

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

- ☐ Fasta – 100 lines; fastq – 4 lines



1 / 1  
points

3.

Which of the following is incorrect:

- ☒ BEDtools can be used to align sequences to the genome.

Correct

- ☐ The length of a read with CIGAR 51M1D24M is 75 bp.
- ☐ An unmapped read can be represented as a SAM record with a “\*” in column 7.
- ☐ “SRA” stands for “Short Read Archive”, an NCBI database that stores short read sequences.



1 / 1  
points

4.

Which of the following is NOT an alignment operation:

- ☐ Padding
- ☒ Cut and paste

## Module 2 Quiz

**9/10 points (90%)**

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?

- ☐ 7
- ☒ 3

**Correct**

- ☐ 5
- ☐ 6

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

**Correct**

## Module 2 Quiz

**9/10 points (90%)**

Quiz, 10 questions



```

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

```

1 / 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   chr1  HAVANA  exon  3213609 3216344 . - . gene_id "MUSG51951.5"; transcript_id
5   chr1  HAVANA  exon  3205901 3207317 . - . gene_id "MUSG51951.5"; transcript_id
6   chr1  HAVANA  transcript  3206523 3215632 . - . gene_id "MUSG51951.5";
7   chr1  HAVANA  exon  3213439 3215632 . - . gene_id "MUSG51951.5"; transcript_id
8   chr1  HAVANA  exon  3206523 3207317 . - . gene_id "MUSG51951.5"; transcript_id
9   ```

```



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

**Correct**

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



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.

1 / 1  
points

8.


## Module 2 Quiz

**9/10 points (90%)**

```

1 ...
2 R1 83 chr12 9232390 255 50M = 9232180 0
3 ATGGCAGAGCCTAATATGTCTCCTAGAGAATGGGAGAGATGGGAAGTCAT
   HGHHHHHHHHHHHHHHHHHHHHHHHHHHHHIGIIIIHHHHHHHHHHHGHFH NM:i:0 NH:i:1 HI:i:0
4 R2 97 chr12 9232391 255 28M278N22M = 9242529
5 0 TGCAGAGCCTAATATGTCTCCCAAACTGACACAGAAGCTCGGGCAGAT D
   >DDHHHHHHHHHHHHIHHHHHHHHHHHIGFFGGHHHHHHHHHHF.F NM:i:4 NH:i:3 HI:i:0 XS:A
   :+ NS:i:2
6 R3 77 * 0 0 0 * * 0 0 CTGATATGAGGAAGAGGATTGCTTAAGCCCGAGGAGTAGAGGCTGTACC
   @@@FDFHFHFFHJJJJIJEGFGIGHHHIHHIIIIGCDE?D?FGGCBHDGGG
7 ...

```

- ☐ R1 maps uniquely to the genome.
- ☐ R3 is unmapped.
- ☒ R2's mate is unmapped.
- 
- Correct**
- ☐ The R1 alignment is the primary mapping (hit index 0) for that read.

**Correct**


✖ 0 / 1 points

For the alignment below, which statements are FALSE? The binary encoding for 97 is 97<sub>2</sub> = 0000 0110 0001<sub>2</sub>. Select all answers that apply.

```
1  ...
2  R2 97 chr12 9232391 255 28M278N22M = 9242529
3  0 TGGCAGAGCCTAATATGTCTCCCAAACTGAGACAGAAGCTCGGGCAGAT D
  >DDDDHHHHHHHHHHIHHHHHHHHHHIIGFFGGGHHHHHHHHHHHFB.F NM:i:4 XS:A:+ NS:i:2
4  ...
```

-  This is the first read in the pair.

**This should not be selected**

-  The length of the read is 50 bp.

**Un-selected is correct**

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

**Un-selected is correct**



The read sequence is reverse complemented in the alignment.

## Module 2 Quiz

9/10 points (90%)

Quiz, 10 questions

Correct



The two mates are identical in sequence.

Correct



Both the read and its mate are mapped.

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, 2, 2



5, 5, 5



3, 4, 2



3, 2, 2

Correct



## Module 2 Quiz

**9/10 points (90%)**

Quiz, 10 questions