# Feedback — Module 2 Quiz

Thank you. Your submission for this quiz was received.

You submitted this quiz on **Thu 17 Sep 2015 10:35 PM PDT**. You got a score of **10.00** out of **10.00**.

## Question 1

Which of the following strings cannot denote a DNA sequence:

| Your Answer | | Score | Explanation |
|---|---|---|---|
| ○ GGGGGGGGG | | | |
| ○ GACTACGAGCGATTTACAGCGAGCATT | | | |
| ⊙ MASLLRG | ✔ | 1.00 | |
| ○ AGCTACTACGAGCT | | | |
| Total | | 1.00 / 1.00 | |

## Question 2

How many lines does it take to specify:

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

Select the best answer.

| Your Answer | | Score | Explanation |
|---|---|---|---|
| ⊙ Fasta – a fasta header followed by any number of sequence | ✔ | 1.00 | |

lines; fastq – 4 lines

○ Fasta – 3 lines, if one of the lines is a blank line; fastq – 4
lines

○ Fasta – 3 lines; fastq – 3 lines

○ Fasta – 1 line; fastq – 4 lines

| Total | 1.00 / 1.00 |
| --- | --- |

# Question 3

Which of the following is incorrect:

| Your Answer | Score | Explanation |
| --- | --- | --- |
| ○ BEDtools can be used to extract alignments within a certain range. | | |
| ○ 'SRA" stands for "Short Read Archive", an NCBI database that stores short read sequences. | | |
| ○ The symbols H and S in the CIGAR field of a SAM record represent 'hard' and 'soft' clipping, respectively. | | |
| ● SAMtools flagstats reports the total number of mapped reads. | ✔ 1.00 | |
| Total | 1.00 / 1.00 | |

# Question 4

Which of the following is NOT an alignment operation:

| Your Answer | Score | Explanation |
| --- | --- | --- |

○ Padding

○ Insertion

◉ Cut and paste         ✔        1.00

○ Substitution

Total                  1.00 / 1.00

# Question 5

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

| Your Answer | Score | Explanation |
|---|---|---|
| ○ 7 | | |
| ◉ 3 | ✔   1.00 | |
| ○ 1 | | |
| ○ 9 | | |
| Total | 1.00 / 1.00 | |

# Question 6

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

```
chr1    CLASS    exon    516    811    100    +    .    gene_id "ge
nA"; transcript_id "genA.1";
chr1    CLASS    exon    1001    1115    100    +    .    gene_id "ge
nA"; transcript_id "genA.1";
chr1    CLASS    exon    3010    3312    100    +    .    gene_id "ge
nA"; transcript_id "genA.1"
```

| Your Answer | Score | Explanation |
|---|---|---|

○

```
chr1 515  3312     genA.1    +    515 3312     0    3    29
6,115,303       516,1001,3010
```

○

```
chr1      515  811  genA.Exon1    100 +    515 811 255,0,0
chr7      3009    3312     genA.Exon3    100 +    3009
3312      255,0,0
```

◉  ✔  1.00

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

○

```
chr1      515  811  genA.Exon1    100 +    515 811 255,0,0
chr7      1000    1115     genA.Exon2    100 +    1000
1115      255,0,0
chr7      3009    3312     genA.Exon3    100 +    3009
3312      255,0,0
```

| Total | 1.00 / 1.00 | |
|---|---|---|

# Question 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.

```
chr1    HAVANA   gene    3205901 3671498 .      -      .      gene_id "MU
```

```
SG51951.5";
chr1    HAVANA  transcript      3205901 3216344 .       -       .       gen
e_id "MUSG51951.5"; transcript_id "MUST162897.1";
chr1    HAVANA  exon    3213609 3216344 .       -       .       gene_id "MU
SG51951.5"; transcript_id "MUST162897.1";
chr1    HAVANA  exon    3205901 3207317 .       -       .       gene_id "MU
SG51951.5"; transcript_id "MUST162897.1
chr1    HAVANA  transcript      3206523 3215632 .       -       .       gen
e_id "MUSG51951.5"; transcript_id "MUST159265.1";
chr1    HAVANA  exon    3213439 3215632 .       -       .       gene_id "MU
SG51951.5"; transcript_id "MUST159265.1";
chr1    HAVANA  exon    3206523 3207317 .       -       .       gene_id "MU
SG51951.5"; transcript_id "MUST159265.1";
```
```

| Your Answer | Score | Explanation |
|---|---|---|
| ● Genes: 1; Transcripts: 2; Exons: 2,2; Strand: -; Length of 5' exon(s): 2736, 2194. | ✔ 1.00 | |
| ○ Genes: 2; Transcripts: 4; Exons: 1,1,1,1; Strand: -; Length of 5' exon(s): 2736,1417, 2194,795. | | |
| ○ Genes: 2; Transcripts: 2; Exons: 3,2; Strand: -; Length of 5' exon(s): 2736, 2194. | | |
| ○ Genes: 1; Transcripts: 2; Exons: 2,2; Strand: +; Length of 5' exon(s): 2736, 2194. | | |
| Total | 1.00 / 1.00 | |

# Question 8

Which of the following is FALSE for the following read alignments:

```
R1      83      chr12   9232390 255     50M     =       9232180 0
ATGGCAGAGCCTAATATGTCTCCTAGAGAATGGGAGAGATGGGAAGTCAT      HGHHHHHHHHHHHHHHHHHH
HHHHHHHHHIGIIIIHHHHHHHHHHHGHHFH NM:i:0  NH:i:1  HI:i:0
```

```
R2      97      chr12   9232391 255     28M278N22M      =       9242529
0       TGGCAGAGCCTAATATGTCTCCCAAAACTGAGACAGAAGCTCGGGCAGAT      D>DDDHHHHHH
HHHHIHIHHHHHIHHHHIGFFGGGHHHHHHHHHHHFB.F  NM:i:4  NH:i:3  HI:i:0  XS:A:+  N
S:i:2
R3      77      *       0       0       0       *       *       0       0
CTGATATGAGGAAAGAGGATTGCTTAAGCCCAGGAGGTAGAGGCTGTACC      @@@FFDFFHFFHHJJJJIJ
EGFGIGHHIHIIIIGCDE?D?FGGCBHDGGG
```

| Your Answer | Score | Explanation |
|---|---|---|
| ○ R3 is unmapped. | | |
| ○ The R1 alignment is the primary mapping (hit index 0) for that read. | | |
| ⦿ R1, R2 and R3 all map to chromosome 12. | ✔ 1.00 | |
| ○ R2 maps in 3 places within the genome. | | |
| Total | 1.00 / 1.00 | |

# Question 9

For the alignment below, which statements are FALSE? The binary encoding for 97 is 972 = 0000 0110 00012. Select all answers that apply.

```
R2      97      chr12   9232391 255     28M278N22M      =       9242529
0       TGGCAGAGCCTAATATGTCTCCCAAAACTGAGACAGAAGCTCGGGCAGAT      D>DDDHHHHHH
HHHHIHIHHHHHIHHHHIGFFGGGHHHHHHHHHHHFB.F  NM:i:4  XS:A:+  NS:i:2
```

| Your Answer | Score | Explanation |
|---|---|---|
| ☑ Both the read and its mate are mapped. | ✘ 0.00 | |
| ☑ The read sequence is reverse complemented in the alignment. | ✔ 0.25 | |

| | | |
|---|---|---|
| ☐ The read is part of a pair. | ✖ | 0.00 |
| ☑ The alignment represents a potential PCR or optical duplicate. | ✔ | 0.25 |
| ☑ This is the second read in the pair. | ✔ | 0.25 |
| ☐ The alignment passes quality checks. | ✖ | 0.00 |
| ☑ The two mates are identical in sequence. | ✔ | 0.25 |
| Total | | 1.00 / 1.00 |

# Question 10

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

```
```
File A                                          File B
chr1    100     400                             chr1     300      500
chr1    1000    1400                            chr1     900      1600
chr1    2000    2400                            chr12    2000     2200
```

What would be the answers for the following sequence of commands:
```

bedtools intersect –wao –a A.bed –b B.bed | sort –u | wc -l
bedtools intersect –wo –a A.bed –b B.bed  | cut –f1-3 | sort –u | wc -l
bedtools intersect –wo –a A.bed -b B.bed  | cut -f4-6 | sort –u | wc -l
```
```

| Your Answer | Score | Explanation |
|---|---|---|
| ◉ 3, 2, 2 | ✔ 1.00 | |
| ◯ 9, 3, 3 | | |
| ◯ 3, 6, 3 | | |
| ◯ 3, 2, 6 | ✖ | |

| Total | 1.00 / 1.00 |