Module 4 Quiz

6/10 points (60%)

Quiz, 10 questions

✖ Required to pass: 70% or higher

You can retake this quiz up to 3 times every 8 hours.

Back to Week 4

Retake

✔ 1 / 1 points

1.
Which of the following is FALSE:

- A human gene can express at most 12 splice variants.

Correct

- Genes form discrete units along the genome.
- Eukaryotic genes are interrupted, generally consisting of multiple exons.
- More than 90% of human genes are alternatively spliced.

✔ 1 / 1 points

2.
Which of the following is FALSE about the organization of a eukaryotic gene:

- The length of an intron cannot be a multiple of 3.

Correct

- The number of introns in a transcript is one less than the number of exons.
-

Some eukaryotic gene transcripts can consist of a single exon.

Module 4 Quiz

Genes that have only one exon are not alternatively spliced.

6/10 points (60%)

Quiz, 10 questions

0 / 1 points

3.

What programs could you use to align RNA-seq reads to: i) a reference genome, and ii) a transcript database?

- tophat, bowtie
- bowtie, tophat

This should not be selected

- bowtie, bowtie
- bowtie, samtools

1 / 1 points

4.

Which of the following is FALSE:

- RNA-seq analyses can reveal known genes and their splice variants, as well as novel genes.
- RNA-seq can be used to quantify the expression levels of proteins.

Correct

- Spliced reads can be used to determine the introns in a gene.
- 'Transfrag' stands for 'transcript fragments', a reference to the fact than transcript assemblers cannot always reconstruct full-length splice variants.

Module 4 Quiz

Quiz, 10 questions

6/10 points (60%)

0 / 1 points

5.
What programs could be used to: i) assemble transcripts from RNA-seq reads, and ii) identify potentially novel transcripts and genes?

- cufflinks, cuffcompare
- cufflinks, cufflinks
- cufflinks, cuffmerge

This should not be selected

- tophat, cufflinks

1 / 1 points

6.
Which of the following is FALSE about the gene annotations in the following GTF snippet:

```
1  ```
2  chr1  MGF gene  3413609 3671498 . - . gene_id "MG051951";
3  chr1  MGF transcript  3413609 3416344 . - .gene_id "MG051951"; transcript_id
     "MT162897";
4  chr1  MGF exon  3413609 3416344 . - . gene_id "MG051951"; transcript_id
     "MT162897";
5  chr1  MGF transcript  3421702 3671498 . - . gene_id "MG051951"; transcript_id
     "MT070533";
6  chr1  MGF exon  3670552 3671498 . - . gene_id "MG051951"; transcript_id
     "MT070533";
7  chr1  MGF CDS 3670552 3671348 . - 0 gene_id "MG051951"; transcript_id "MT070533"
     ;
8  chr1  MGF exon  3421702 3421901 . - . gene_id "MG051951"; transcript_id
     "MT070533";
9  chr1  MGF CDS 3421792 3421901 . - 1 gene_id "MG051951"; transcript_id "MT070533"
     ;
10 ```
```

- The two transcripts for gene MG051951 overlap on the genome.

Correct

- It contains only one gene, MG051951.
- Gene MG051951 has two transcripts, MT162897 and MT070533.
- Transcript MT162897 has a single exon.

Module 4 Quiz

6/10 points (60%)

Quiz, 10 questions

0 / 1 points

7.

What does the following code NOT do:

```
```
BWT2IDX=/home/me/genomes/hg20/hg20
ANNOT=/home/me/genomes/hg20/myannot.gtf
ANNOTIDX=/home/me/genomes/hg20/myannot/myannot

mkdir -p /home/me/SRR100000
tophat2 -o /home/me/SRR100000 -p 10 --max-multihits 10 \
        -r 26 --mate-std-dev 25 \
        -a 6 \
        -G $ANNOT --transcriptome-index $ANNOTIDX \
        $BWT2IDX \
        /home/me/SRR100000_1.fastq.gz /home/me/SR100000_2.fastq.gz
```
```

- Report spliced reads with at most 6 mismatches in the anchor site
- Use information about the insert size distribution, e.g. its average (26) and standard deviation (25)
- Use known gene annotations in the first mapping stage

This should not be selected

- Load the gene annotations index from /home/me/genomes/hg20/myannot/myannot

1 / 1 points

8.

What does the following code NOT do:

```
```
TOPHATDIR=/home/florea/Tophat/

mkdir -p Test1
cd Test1
ln -s $TOPHATDIR/accepted_hits.bam .
cufflinks -L Test1 -p 8 -j 0.10 -F 0.05 accepted_hits.bam
```
```

- Label cufflinks transcripts with the prefix ‘Test1’
- Run cufflinks to assemble transcripts

Use the default reference transcript annotation to guide assembly

Module 4 Quiz

6/10 points (60%)

Quiz, 10 questions

Correct

Create a soft link to the BAM read alignment file in the Test1 directory

1 / 1 points

9.

Which of the following is NOT described in the following summary file produced by tophat:

```
```
Left reads:
          Input   : 60586968
           Mapped   : 58163843 (96.0% of input)
            of these: 6832240 (11.7%) have multiple alignments (359075 have
              >10)
Right reads:
          Input   : 60586968
           Mapped   : 56969290 (94.0% of input)
            of these: 6668479 (11.7%) have multiple alignments (358573 have
              >10)
95.0% overall read mapping rate.

Aligned pairs: 55880048
     of these: 6491876 (11.6%) have multiple alignments
                 2795712 (5.0%) are discordant alignments
87.6% concordant pair alignment rate.
```
```

The number of read pairs in which both mates were aligned is 55,880,048

The reads were 100 bp long

Correct

Of the pairs in which both reads were aligned, 5.0% were at a distance or in an orientation inconsistent with the insert size and/or library-type

Of the aligned pairs, 6,491,876 had multiple matches on the genome

0 / 1 points

Module 4 Quiz

Quiz, 10 questions

10.

Which of the following is NOT TRUE about the output below, obtained from a cuffdiff differential expression analysis:

6/10 points (60%)

```
1  ```
2  XLOC_000002 XLOC_000002 AT1G01020 1:5927-8737 q1  q2  OK  1.13032 3.48406 1
     .62404  0.694576  0.5277  0.998846  no
3  XLOC_000004 XLOC_000004 AT1G01073 1:44676-44787 q1  q2  NOTEST  0 0 0 0 1 1 no
4  XLOC_000042 XLOC_000042 AT1G01580 1:209394-213041 q1  q2  OK  1.59512 0 -inf
     nan 5e-05 0.0096703 yes
5  ```
```

- [ ] Locus XLOC_000002 appears upregulated in condition 2. The log fold change for XLOC_000002 is 1.62404

This should not be selected

- [x] Locus XLOC_000002 is located on the genomic axis '1', between positions 5927 and 8737

Un-selected is correct

- [x] Locus XLOC_000002 has FPKM 1.1302 in condition 1 and 3.48406 in condition 2

Un-selected is correct

- [ ] There are too many alignments for testing for differential expression at locus XLOC_000004

This should be selected