Module 3 Quiz

10/10 points (100%)

Quiz, 10 questions

Congratulations! You passed!

Next Item

1 / 1 points

1.
Which of the following statements is FALSE:

- A polymorphism is a genetic variant that occurs in more than 1% of the population.
- In a normal cell, each locus in the genome can have at most two alleles.
- SNP refers to a Single Non-defined Polymorhism

Correct

- SNVs encompass single nucleotide insertions, deletions and substitutions.

---

1 / 1 points

2.
Which of the following statements is FALSE:

- The BAM format is a binary compressed representation for alignments of next generation sequencing reads.
- The VCF format shows the changes in amino acid resulting from the nucleotide mutation, in column 3.

Correct

- The mpileup format has either 6 or 7 columns.

Module 3 Quiz

Quiz, 10 questions

10/10 points (100%)

- The genotype fields in VCF provide information about the variant in each sample.

1 / 1 points

3.
What program can be used to generate a list of candidate sites of variation in an exome data set:

- bcftools

Correct

- bedtools
- samtools
- mkdir

1 / 1 points

4.
In a comprehensive effort to study genome variation in a patient cohort, you sequence and call variants in the exome, whole genome shotgun and RNA-seq data from each patient. Which of the following is FALSE when comparing these three types of resources:

- RNA-seq allows detection of intronic variants.
- RNA editing can confound the detection of variants from RNA-seq data.
- All of the three methods can identify variants located in the introns, albeit to different degrees.
- Exome sequencing comprehensively captures variants in the 3' and 5' UTRs of genes.

Correct

Module 3 Quiz

Quiz, 10 questions

10/10 points (100%)

1 / 1 points

5.

Which of the following options can be used to allow bowtie2 to generate partial alignments?

- -D
- --sensitive
- --local

Correct

- –ignore-quals

1 / 1 points

6.

Select the correct interpretation for the snippet of 'mpileup' output below.

```
1  ```
2  Chr3  11700316  C 8 .$....... 8C@C;CB3
3  Chr3  11951491  G 16  AAAA,......aA..A  C2@2BCBCCCAC2CC4
4  ```
```

- Only site 1 shows potential variation;

  the alternate letter for site 1 is C;

  site 1 has 9 supporting reads, and site 2 has 16

- Only site 1 shows potential variation;

  the alternate letter for site 1 is $;

  site 1 has 8 supporting reads, and site 2 has 16

- Only site 2 shows potential variation;

  the alternate letter for site 2 is A;

  site 1 has 8 supporting reads, and site 2 has 16

Correct

Both sites show potential variation;

the alternate letter for site 1 is '.', and for site 2 is A;

site 1 has 9 supporting reads, and site 2 has 16

Module 3 Quiz

Quiz, 10 questions

10/10 points (100%)

---

1 / 1 points

7.

Given the set of variants described in the VCF excerpt below, which of the following is FALSE?

```
```
##INFO=<ID=DP,Number=1,Type=Integer,Description="Raw read depth">
##INFO=<ID=MQ,Number=1,Type=Integer,Description="Average mapping quality">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FORMAT=<ID=PL,Number=G,Type=Integer,Description="List of Phred-scaled genotype likelihoods">
Chr3 11966312   . G A 15.9 . DP=5;MQ=15   GT:PL   1/1:43,9,0
Chr3 11972108 . TAAAA TAAA 32.8 . INDEL;IDV=7;IMF=0.636364;DP=11;MQ=22 GT:PL 0/1:66,0,2
Chr3 13792328 rs145271872 G T 5.5 . DP=1;MQ=40 GT:PL 0/1:32,3,0
```
```

The sample contains only the alternate allele for variant 1

The sample contains both alleles for variant 2

Average mapping quality for variant 3 is 40

The sample contains only the alternate allele for variant 3

Correct

---

1 / 1 points

8.

What does the following code do:

```
```
bowtie2 -x species/species -U in.fastq | grep -v "^@" | cut -f3 | sort | uniq -c
```
```

Run bowtie2 with a set of single-end reads, reporting up to 5 alignments per read;

then determine the number of matches with unmapped mates

Module 3 Quiz

Quiz, 10 questions

10/10 points (100%)

Run bowtie2 with a set of single-end reads, allowing for local matches;

then determine the number of matches with unmapped mates

Run bowtie2 with a set of single-end reads, reporting the best alignment only;

then determine the number of matches on each genomic sequence

Correct

Run bowtie2 with a set of single-end reads, reporting the top 5 alignments for a read;

then list the number of matches containing insertions and deletions, respectively

---

1 / 1 points

9.

What does the following snippet of code do NOT do:

```
1 ```
2 samtools mpileup -O -f genome.fa  in.bam | cut -f7
3 ```
```

Report in the intermediate mpileup output the qualities of all read bases aligned at that position

Require a sorted BAM file

Report an empty column

Correct

Produce a 7-column intermediate mpileup file that is piped to 'cut'

---

1 / 1 points

10.

Module 3 Quiz

Quiz, 10 questions

10/10 points (100%)

What does the following code do NOT do:

```
```
bcftools call -v -c -O z -o out.vcf.gz in.vcf.gz
```
```

- [x] Skip indels

Correct

- [ ] Call variants in a single sample
- [ ] Report variant sites only
- [ ] Report output in compressed VCF format