

1. Please go to the [link](#) and fill in the details in the table below.

Gene ID	1956
Official Symbol	
Gene Location	
Exon Count	
Gene Type	
Expression in Skin (RPKM)	

Answer:

Gene ID	1956
Official Symbol	EGFR
Gene Location	7p 11.2
Exon Count	32
Gene Type	Protein Coding
Expression in Skin (RPKM)	15.6

2. Use the following [link](#) for the nucleotide search using blastn.

Copy and paste the following nucleotide sequence as a query:

ATGCGTACTTTAAAGCATCGGGGGGA

Keep all the parameters as default as you click on the **BLAST**.

- (a) Which of the following species has the maximum query coverage?

- i. [Gobius niger](#) (Black goby)
- ii. [Andrena bicolor](#) (Gwynne's mining bee)
- iii. [Arabidopsis thaliana](#) (Thale cress)
- iv. [Petrosia ficiformis](#) (Stony sponge)

Answer: (ii), (iii)

- (b) Which of the following species has the highest total score?

- i. [Andrena bicolor](#) (Gwynne's mining bee)
- ii. [Stenoptilia bipunctidactyla](#) (Twin-spot plume)
- iii. [Herminia tarsipennalis](#) (Fan-foot)
- iv. [Petrosia ficiformis](#) (Stony sponge)

Answer: (i)

3. Use the following [link](#) for the multiple search alignment using Muscle:
Copy and paste the following FASTA sequences in the query:

- I.

```
>sp|P01308|INS_HUMAN Insulin OS=Homo sapiens OX=9606 GN=INS PE=1 SV=1
MALWMRLLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYTPKTRR
EAED LQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN
```
- II.

```
>sp|P01315|INS_PIG Insulin OS=Sus scrofa OX=9823 GN=INS PE=1 SV=2
MALWTRLLPLLALLALWAPAPAQAFVNQHLCGSHLVEALYLVCGERGFFYTPKARRE
AEN PQAGAVELGGGLGGLQALALEGPPQKRGIVEQCCTSICSLYQLENYCN
```
- III.

```
>sp|P01318|INS_SHEEP Insulin OS=Ovis aries OX=9940 GN=INS PE=1 SV=2
MALWTRLVPLLALLALWAPAPAHAFVNQHLCGSHLVEALYLVCGERGFFYTPKARRE
VEG PQVGALELAGGPGAGGLEGPPQKRGIVEQCCAGVCSLYQLENYCN
```
- IV.

```
>sp|Q6YK33|INS_GORGO Insulin OS=Gorilla gorilla gorilla OX=9595 GN=INS PE=3
SV=1
MALWMRLLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYTPKTRR
EAED LQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN
```
- V.

```
>sp|P01322|INS1_RAT Insulin-1 OS=Rattus norvegicus OX=10116 GN=Ins1 PE=1
SV=1
MALWMRFLPLLALLVLWEPKPAQAFVKQHLCGPHLVEALYLVCGERGFFYTPKSRR
EVED PQVPQLELGGGPEAGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN
```

Set the output format as ClustalW, keep all the parameters as Default, and then submit the job. Click on the tab **Phylogenetic tree**.

- (a) Which of the given sequence is shown to exhibit maximal similarity with human insulin?
- i. sp|P01322|INS1_RAT
 - ii. sp|P01315|INS_PIG
 - iii. sp|P01318|INS_SHEEP
 - iv. sp|Q6YK33|INS_GORGO

Answer: (iv)

- (b) Which of the following algorithms is used for the construction of the phylogenetic tree?
- i. Neighbour-Joining method
 - ii. Fitch-Margoliash method

- iii. Unweighted Pair Group Method with Arithmetic Mean (UPGMA)
- iv. Maximum Parsimony

Answer: (i)

4. Navigate to [BDMC of BrainProt](#) and select the disease [Meningioma](#).

(a) Which of the following gene is coming up as the top marker with respect to the BDMC score?

- i. NF2
- ii. TP53
- iii. EGFR
- iv. VIM

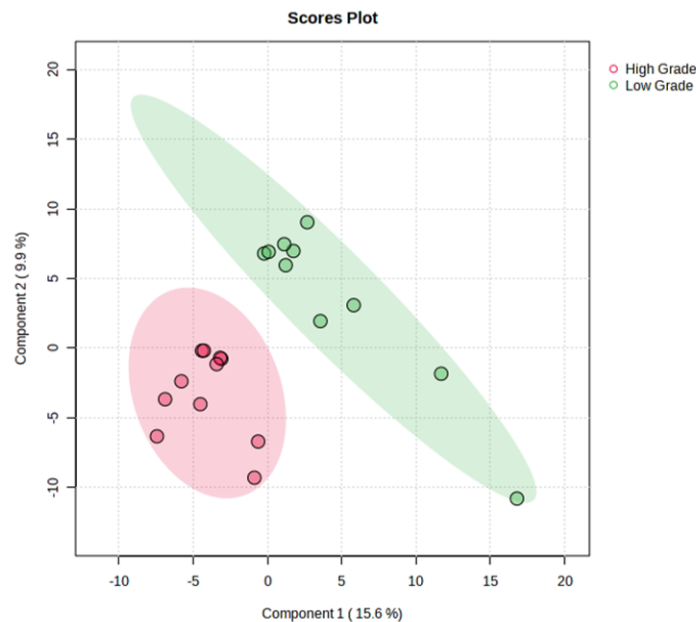
Answer: (i), (iii)

(b) What is the score of NF2 gene in PubPular?

- i. 1
- ii. 0.467
- iii. 0.497
- iv. 0.397

Answer: (i)

(c) Look at the following plot. Reema has performed a classification analysis taking high grade meningioma and low meningioma samples. Which of the following is true regarding the plot?



- i. The total cumulative variance is 25.5%
- ii. The total cumulative variance is 25.6%
- iii. The total cumulative variance is 5.6%
- iv. The total cumulative variance is 5.7%

Answer: (i)

(d) Which of the following plots have similar interpretation?

- i. Violin plot and Box plot
- ii. Bar chart and Pie chart
- iii. Violin plot and Bar chart
- iv. Bar chart and Box plot

Answer: (i)

5. Navigate to this [link](#) and go to the disease [Amyotrophic Lateral Sclerosis](#).

How many of the compounds are present in clinical trial phase 4.0?

- (a) 1
- (b) 2
- (c) 3
- (d) 4

Answer: (b)