

Stuart Kettler

#### Assignment 4 answers

1.

```
makeblastdb -in 'mystery_transcripts 2.fasta' -dbtype nucl -out mystery_transcripts_db
```

2.

```
makeblastdb -in 'mytery_proteins 2.fasta' -dbtype prot -out mystery_proteins_db
```

3.

```
blastn -query mystery_nucleotide_query.fasta -db mystery_transcripts_db -evalue 1e-3 -  
max_target_seqs 5 -outfmt '7 qseqid sseqid pident length mismatch gapopen qstart qend  
sstart send evalue bitscore' -out nucleotide_search_results.tab
```

```
4) mystery_nucleotide_query      T010027234997      91.9651481  115   4     1  
    1479  197   1675  0.0   2073
```

```
mystery_nucleotide_query T010003453160      91.3511480  126   2     1    1479  
    193   1671  0.0   2023
```

```
mystery_nucleotide_query T010027234995      91.3511480  126   2     1    1479  
    193   1671  0.0   2023
```

5.

```
awk '$3 > 90{print}' mystery_search.txt | awk '{print$2,$3}' >> most_similar.txt
```

```
blastp -db mystery_proteins -query 'mystery_protein_query 2.fasta' -outfmt 7
```

7. P010027090796, P010027090798, P010027073865, P010027077695

8. *Coffea arabica*

9. *Gardenia jasminoides*

10. putative catalase