Micro B3 EU-US Bioinformatics Course: Day 2
Command line BLAST, parsing output

1. **Go back to the home directory and start an interactive torque job**

   ```
   $ cd ~
   $ int
   ```

2. **Create a few output directories**

   ```
   $ mkdir output_repository
   $ cd output_repository
   $ mkdir blastn
   $ mkdir blastp
   $ mkdir tblastn
   $ mkdir extracted
   $ cd ~
   ```

3. **Performing a BLASTn**

   ```
   $ blastn -db ~/microb3-shared/blast/databases/small.nuc
   -query ~/microb3-shared/blast/queries/small.nuc.read.fasta
   -out ~/output_repository/blastn/small.raw —outfmt 0
   ```

4. **Checking out the results**

   ```
   $ cd output_repository/blastn
   $ less small.raw
   ```

5. **Another format for output (more compact format)**

   ```
   $ cd ~
   $ blastn -db ~/microb3-shared/blast/databases/small.nuc
   -query ~/microb3-shared/blast/queries/small.nuc.read.fasta
   -out ~/output_repository/blastn/small.btab —outfmt 6
   ```

6. **Checking out the results**

   ```
   $ cd output_repository/blastn
   $ less small.btab
   ```

7. **Performing a BLASTp**

```
$ cd ~
$ blastp -query ~/microb3-shared/blast/queries/small.pep.orf
.fasta -db ~/microb3-shared/blast/databases/small.pep
-out output_repository/blastp/small.raw -outfmt 0
```

8. Checking out the results

```
$ cd output_repository/blastp
$ less small.raw
```

9. Bigger BLASTp

```
$ cd ~
$ blastp -query ~/microb3-shared/blast/queries/dna_polymerase_I.fasta
-db ~/microb3-shared/blast/databases/cbay.orf.pep
-out ~/output_repository/blastp/dnapola.btab -outfmt 6
```

10. Checking out the results

```
$ cd output_repository/blastp
$ less dnapola.btab
```

11. How many hits?

```
$ wc -l dnapola.btab
```

12. How many unique query sequences, unique subject sequences?

```
$ cut -f1 dnapola.btab | sort -u | wc -l

$ cut -f2 dnapola.btab | sort -u | wc -l
```

13. Extracting query sequences

```
$ extract_blast_hits.pl -m

$ extract_blast_hits.pl --btab ~/output_repository/blastp/dnapola.btab
--fasta ~/microb3-shared/blast/databases/cbay.orf.pep
--out ~/output_repository/extract.subj.fasta --subj
```

14. Using other BLAST flavors (tBLASTn)

```
$ cd ~
```

```
$ tblastn -db ~/microb3-shared/blast/databases/cbay.reads -
query ~/microb3-shared/blast/queries/dna_polymerase_I.fasta -
out ~/output_repository/tblastn/cbay.dnapol.btab -outfmt 6

$ tblastn -db ~/microb3-shared/blast/databases/cbay.reads -
query ~/microb3-shared/blast/queries/dna_polymerase_I.fasta -
out ~/output_repository/tblastn/cbay.dnapol.raw -outfmt 0
```

15. Tinkering with word size

```
$ exit

$ cd ~

$ less ~/microb3-shared/blast/queries/crispr.spacers.gos.fasta

$ less ~/microb3-shared/bin/blastn-crispr.qs

$ qsub ~/microb3-shared/bin/blastn-crispr.qs

$ wc -l ~/output_repository/blastn/crispr.btab

$ less ~/microb3-shared/bin/blastn-crispr-word.qs

$ qsub ~/microb3-shared/bin/blastn-crispr-word.qs

$ wc -l ~/output_repository/blastn/crispr.7.btab
```