
Laboratory Journal

Cody Glickman

cody.glickman@ucdenver.edu

Beginning 27 February 2018

Contents

Week of 26 February 2018	1
1 Identifying Virulence Factors in Phages	1
Week of 5 March 2018	2
1 Identifying Virulence Factors in Phages	2
2 Lysogenic Pan Genome	2
Week of 12 March 2018	4
1 This shows a sample table	4
2 example2	4

Week of 26 February 2018

1 Identifying Virulence Factors in Phages

Downloaded phage protein data searching by taxonomy from Uniprot/TrEMBL. Downloaded three datasets Claudioviruses, Ligamenvirales, and Unclassified. Downloaded viral contigs from vHMM [Earth Virome Project](#).

Metagenomic gene prediction through prodigal downloaded from conda "conda install prodigal"

```
prodigal -i mVGs_sequences_v2.fna -o my_genes -a my_proteins.faa -p meta
```

The myproteins.faa file contains the translated predicted genes. This set is applied to the VF HMM and similar to the other datasets, returned no hits with hmmsearch.

Creating virulence factor blast database to blast against viral contigs. Choosing a E value threshold:

Goals Finish written GRAB and Abstract for NLM Done (3/2/18)

```
makeblastdb -in Combined_VF.faa -dbtype prot -out Combined_VF -title "Combined_VF"
```

```
blastp -db /Users/stronglab2/blastdb/Combined_VF/Combined_VF -out results.txt  
-outfmt 6 -query phage_proteins.faa
```

```
## Hello World  
print(x)
```

Week of 5 March 2018

Goals

- Prophage Annotation and VF/ARG Pipeline (Wednesday)
- ML Pipeline Active Prophage
- CAMI Data With Conda for Reproducibility

1 Identifying Virulence Factors in Phages

Establishing the BASH pipeline

1. Prophage Prediction

Input: Contigs.fasta

Output: Prophage Zip Folder

2. Gene Prediction

Input: Sequences of Identified Prophages

Output: Protein Fasta

3. Virulence Factor Identification

Input: Protein Fasta

Output: Proteins called Virulence Factors

2 Lysogenic Pan Genome

Downloaded phage table from [PhageDB](#). Parse temperate phages from Graham Hatful's List and those that infect Mycobacterium. Save Genbank ID numbers as Numbers.txt in script PhageDBProcessing.Rmd.

Calling GBK Files from nuccore.

```
## Load GenBankIds and Remove Whitespace
with open("Numbers.txt") as f:
    content = f.readlines()
content = [x.strip() for x in content]
```

```
## Call Entrez for Genbank_Ids
for i in range(0,len(content)):
    handle = Entrez.efetch(db="nucleotide", id=content[i], rettype="gb")
    filename = 'genbank_files/' + content[i] + '.gbk'
    out_handle = open(filename, "w")
    out_handle.write(handle.read())
    out_handle.close()
    handle.close()
    print("Saved " + filename)
```

Run Core Genome Analysis

Convert to GFF3

```
bp_genbank2gff3.pl --dir pathtofiles
```

Run Roary

```
roary -e --mafft -p 8 *.gff
```

Week of 12 March 2018

1 This shows a sample table

Groups	Treatment X	Treatment Y
1	0.2	0.8
2	0.17	0.7
3	0.24	0.75
4	0.68	0.3

Table 1: The effects of treatments X and Y on the four groups studied.

Table 1 shows that groups 1-3 reacted similarly to the two treatments but group 4 showed a reversed reaction.

2 example2

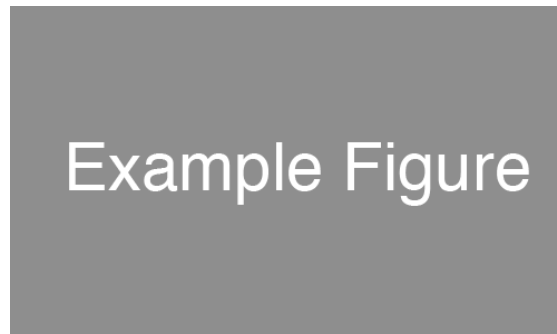


Figure 1: Example figure.

Formulae and Media Recipes

Media

Media 1

Compound	1L	0.5L
Compound 1	10g	5g
Compound 2	20g	10g

Table 1: Ingredients in Media 1.

Formulae

Formula 1 - Pythagorean theorem

$$a^2 + b^2 = c^2$$