Documentation: **working_70_cdhit.fasta** used as query in all instances.
The following 2 run datasets were like test runs and are different from the bulk of the later runs.

Natural wetlands soil metagenome.
Files in 300 reads dataset folder.
>https://www.ebi.ac.uk/metagenomics/search#analyses

>https://www.ebi.ac.uk/ena/browser/view/PRJNA304204

No significant hits (significant as in E-value<0.01, qcovs>80%, pident>35%)

Westerlies Biome of marine viral metagenome. Files in BH folder.

>https://www.ebi.ac.uk/metagenomics/samples/ERS1871373

>https://www.ebi.ac.uk/metagenomics/analyses/MGYA00345647#download

6 hits, some with tail fibers.

WP_003117973.1 ERZ802819.51354

256,1382,1,239,719,1378,239,5.24e-65,43.515,93

WP_047680427.1 ERZ802819.51354

233,1382,9,220,749,1381,213,1.34e-57,45.540,91

WP_047730209.1 ERZ802819.51354

234,1382,7,219,743,1381,215,5.79e-48,42.326,91

WP_047780224.1 ERZ802819.51354

237,1382,8,220,746,1381,215,1.35e-50,41.860,90

WP_048731243.1 ERZ802819.164511

192,712,13,181,595,77,176,2.79e-22,35.227,88

WP_047720857.1 ERZ802819.51354

245,1382,9,221,749,1381,215,4.24e-41,40.465,87

Documentation for large set of environmental metagenomes in file Subject metagenomes.txt from https://www.ncbi.nlm.nih.gov/bioproject/?term=txid1070528[Organism:noexp]

Small Aral sea Metagenomic assembly
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA641503
ID: 641503
SRR12095147
Data in folder: Small Aral Sea
2 hits: both Rhizobium.
WP_047506852.1 SRR12095147.1115259.1
87,301,16,86,31,243,71,8.85e-15,43.662,82
WP_047506852.1 SRR12095147.650748.1
87,290,16,84,83,289,69,5.88e-14,43.478,79

Single Virus Genomics marine phages
BioProject Accession: PRJNA611689
ID: 611689
SRR11283656,SRR11283657,SRR11283658,SRR11283659,SRR11283660,SRR11283661,SRR11283662
Data in folder: Single Virus Genomics marine phages
No significant hits.

Virome of Andean tubers
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA576364
ID: 576364
SRR10244774
Data in folder: Andean Tubers
No significant hits

Lake Baikal Viromes Raw sequence reads
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA547700
ID: 547700
SRR9221221,SRR9217670
Data in folder: Baikal
Hits: Some nearly significant with E.coli and tail fibers
WP_047506852.1 SRR9221221.3132929.2
87,301,2,86,260,3,86,2.06e-14,40.698,98
WP_047506852.1 SRR9221221.1944889.1
87,301,11,86,295,65,77,2.07e-12,44.156,87
WP_045718260.1 SRR9221221.3596752.1
134,301,35,131,300,28,97,2.64e-05,34.021,72
WP_047668184.1 SRR9221221.3065453.1
155,301,46,154,292,23,109,3.49e-14,43.119,70
WP_047506852.1 SRR9217670.381825.2
87,301,3,79,234,1,78,8.06e-12,41.026,89
WP_047506852.1 SRR9217670.2733908.1
87,301,3,73,225,10,72,1.40e-13,45.833,82
WP_047652789.1 SRR9217670.2027147.1
124,301,20,115,5,283,97,1.88e-05,30.928,77

Contigs from Dean Provided week of July 13-17
Data in folder: Contigs
2 significant results (including E. coli tail fiber) as well as some hits slightly below threshold.
WP_048373798.1 QUINN_env_virus_005_075207
312,10961,37,310,9933,10724,281,2.18e-04,20.996,88
WP_048242643.1 QUINN_env_virus_005_185231 (E. coli)
183,33712,23,181,25389,24871,179,6.27e-10,28.492,87
WP_047653121.1 QUINN_env_virus_005_034742
147,6638,5,116,3248,3571,114,7.50e-01,27.193,76
WP_047506852.1 QUINN_env_virus_005_122589
87,1332,22,86,928,1122,65,6.15e-07,41.538,75
WP_047403075.1 QUINN_env_virus_005_102054
197,16268,1,142,4295,3849,161,3.20e-01,24.224,72
WP_047506852.1 QUINN_env_virus_005_031449
87,7928,22,84,3179,2994,63,2.25e-07,44.444,72

Yangshan harbor viral metagenome Metagenomic assembly
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA610033
ID: 610033
SRR11097769
Data in folder: Yangshan Harbor
Many hits but none significant, highest qcovs was 60%.

Viral metagenome in Jiangsu river water samples
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA471935
ID: 471935
SRR7288019
Data in folder:Jiangsu River Metagenome
1 significant hit with others slightly below.
WP_047506852.1 SRR7288019.442045.2
87,246,16,87,221,6,72,1.62e-13,43.056,83

Genomes from the Global Ocean Virome v1
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA477650
ID: 477650
Data in folder: Global Ocean Virome
Protein hits: 37 unique hits, some with tail fibers
WP_001297572.1 WP_047965459.1 WP_048532374.1 WP_048731244.1
WP_048731244.1 WP_003117973.1 WP_016249384.1 WP_017765307.1
WP_047541866.1 WP_047680427.1 WP_047730209.1 WP_047925204.1
WP_048025268.1 WP_048248406.1 WP_048358120.1 WP_001133967.1
WP_047670712.1 WP_047680427.1 WP_047698164.1 WP_047714565.1
WP_047668184.1 WP_047780224.1 WP_048535907.1 WP_048663532.1
WP_048765510.1 WP_047570873.1 WP_047668184.1 WP_048251468.1
WP_048663532.1 WP_047720857.1 WP_047730209.1 WP_047780224.1
WP_045190581.1 WP_047720857.1 WP_048731243.1 WP_003117973.1
WP_047977684.1 WP_048251465.1 WP_048242643.1 WP_017218926.1
WP_047662788.1 WP_047885557.1 WP_048731243.1 WP_000072167.1
WP_047675600.1 WP_048236900.1 WP_048532374.1
Similar hits from dna sequence.

Marine Bacteriophage Metagenome (S-Australia)
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA516067
ID: 516067
SRR8466803,SRR8466804,SRR8466805,SRR8466806,SRR8466807,SRR8466808
Data in folder: Marine Bacteriophage Metagenome
Some significant hits with Rhizobium and other environmental microbes, some with S. flexneri and E. coli just below threshold.

Sichuan takin Metagenome
Organism: viral metagenome (Taxonomy ID 1070528)
BioProject Accession: PRJNA421859
ID: 421859
SRR6365042
Data in folder: Sichuan Takin Metagenome
No significant hits.

Search only metagenomes with good length, >= 300bp reads.
Visualize metagenomic search findings.
Record video for symposium and review presentation materials.
Write SULI abstracts, make presentation, and outline final paper.

NCBI Genome Workbench Version 3.4.1.
CD HIT Server cd-hit accessed 6/22/2020.
NCBI Magic-Blast Version 1.5.0.
R STUDIO Version 1.3.959.
MAFFT Version 7.467.
FigTree Version 1.4.4.
MEGA X Version 10.1.8 build 10200331.
AliView Version 1.26.
HMMER Version 3.3.
NCBI Batch Entrez Accessed from 5/18/2020 to 8/3/2020.
NCBI TBLASTN Accessed from 6/20/2020 to 8/3/2020.