Bioinformatics
CS300
Working With Blast:
manual and "automatic" use

Spring 2021
Oliver BONHAM-CARTER

Blast:
Find your sequence

GenBank ▾

Send to: ▾

Change region shown

Customize view

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

NCBI Reference Sequence: NC_045512.2

FASTA   Graphics

Go to: ▽

```
LOCUS       NC_045512               29903 bp ss-RNA     linear   VRL 18-JUL-2020
DEFINITION  Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1,
            complete genome.
ACCESSION   NC_045512
VERSION     NC_045512.2
DBLINK      BioProject: PRJNA485481
KEYWORDS    RefSeq.
SOURCE      Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)
  ORGANISM  Severe acute respiratory syndrome coronavirus 2
            Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes;
            Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae;
            Betacoronavirus; Sarbecovirus.
```

Analyze this sequence

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

NCBI Virus

Retrieve, view, and download SARS-CoV-2 coronavirus genomic and protein sequences.

Blast:
Enter accession number of sequence

Blast:
Fill in nucleotide parameters

Blast Initiation:
Running the analysis

Blast Results:
Resembling sequences

Blast:
Making sense of scores

- **Max Score**
  - The score of the best matching segment for local alignment, not global
- **Total Score**
  - The total scores of all matching segments found (same as max score if there is only one matching segment)
- **Query Coverage**
  - The percentage of the query sequence that aligned to some part of the match.
- **E-Value**
  - A statistical measure evaluating how likely it is that a match this good could occur by chance. Lower e-scores indicate that both sequences are truly similar and are not similar by chance alone. Identical sequences have e-scores of zero.
- **Max Indent**
  - The percentage of nucleotides that are identical between the query and the target sequences within the matching regions.

Reference:
https://www.ncbi.nlm.nih.gov/books/NBK62051/

Blast:
Phylogenic trees to show relatedness

Blast is Cool!
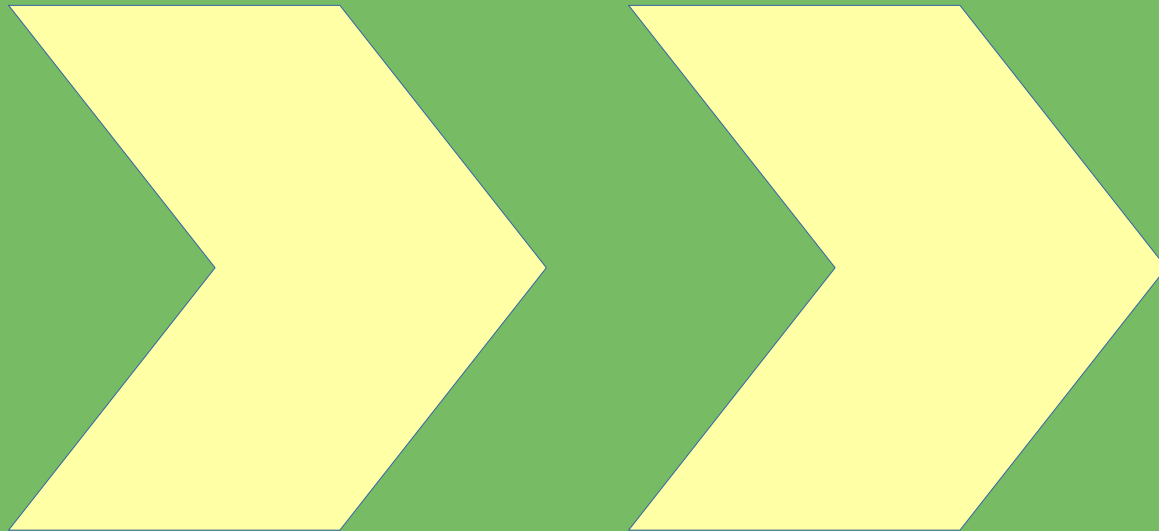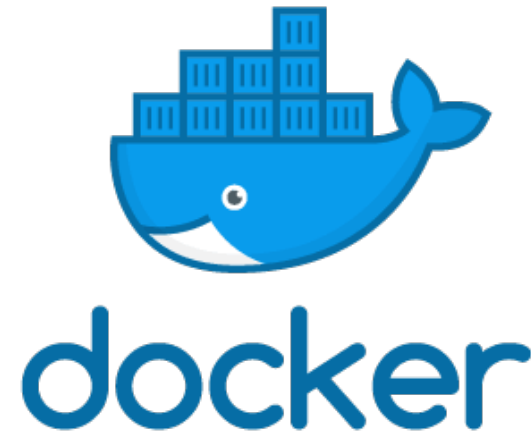
Bring the Code!

Up Next!

Please run your container!!

Dockerfile
(for Docker Desktop)

We will be using BioPython

```
FROM ubuntu:20.04

RUN apt-get update && apt-get -y
install git htop vim python3 python3-pip
clustalw


RUN \
pip3 install --upgrade pip \
pip install bioPython

WORKDIR /root

CMD ["bash"]
```

Activity 06

- Activity 06
  - https://forms.gle/1aKbnUTz11P937fm8
- Blast Link:
  - https://blast.ncbi.nlm.nih.gov/Blast.cgi