**TEXT in GREEN is a specific argument for each function/parameter**

**Installing the NCBI e-utlies command through CLI**

sh -c "$(wget -q ftp://ftp.ncbi.nlm.nih.gov/entrez/entrezdirect/install-edirect.sh -O -)"

**Download a mRNA sequence in fasta format from NCBI using accession number**

esearch -db nucleotide -query "NC\_001552" | efetch-format fasta > output.fasta

**Download a Protein sequence in fasta format from NCBI using accession number**

esearch -db protein -query "NP\_001277121.1" | efetch -format fasta > protein.fasta

**Download a Genomic sequence in fasta format from NCBI using accession number**

esearch -db nucleotide -query "NC\_035444.2" | efetch -format fasta > output.fasta

**Batch retrieval for all proteins for taxon ID. This example will download all proteins for viruses in fasta format.**

esearch -db "protein" -query " txid10239[Organism]" | efetch -format fasta > output.fasta

**Download sequences in fasta format from NCBI using edirect using isolate info**

esearch -db nucleotide -query "GDST008" | efetch -format fasta > output.fasta

**Download sequences from NCBI using edirect using bioproject accession or ID**

esearch -db bioproject -query "PRJNA285593" | elink -target nuccore | efetch - format fasta

esearch -db bioproject -query "285593[id]" | elink -target nuccore | efetch - format fasta

**Get all Coding sequence from a genome**

esearch -db protein -query 302315370 | elink -target nuccore | efetch -format fasta| grep -A 4 --no-group-separator CDS

esearch -db protein -query 302315370 | elink -target nuccore | efetch -format fasta\_cds\_na | grep YP\_003815423.1

**Get taxonomy ID from protein accession number**

esearch -db protein -query "NP\_066243" | elink -target taxonomy | efetch -format uid

#use gi instead

esearch -db protein -query "10314000" | elink -target taxonomy | efetch -format uid

**Get taxonomy ID from accession number using esummary**

esearch -db nucleotide -query "NC\_001552" | esummary | grep TaxId

**Download all refseq protein sequences for viruses**

esearch -db "protein" -query "txid10239[Organism] AND refseq[filter]"| efetch -format fasta > refseq\_protein\_viruses.fa

**Download reference genome sequence from taxonomy ID**

**Note: Using efilter command**

esearch -db taxonomy -query "txid11053 [Organism]" | elink -target nuccore | efilter -query "refseq" | efetch -format fasta >output.fa

**Get all proteins from a genome accession**

elink -db nucleotide -id JN420361.1 -target protein | efetch -format DocSum

**#Downloads protein sequences in fasta format**

elink -db nucleotide -id JN420361.1 -target protein | efetch -format fasta

**#Get GI for these protein**

elink -db nucleotide -id JN420361.1 -target protein | efetch -format uid

**#Retrieve accession numbers of the protein**

elink -db nucleotide -id JN420361.1 -target protein | efetch -format acc

**Note: Following command would work with protein accession and GIs used as -id parameter in elink command.**

elink -db protein -id 817524604 -target nuccore | efetch -format acc

**#Use protein accession instead of GI as -id**

elink -db protein -id YP\_009134732.1 -target nuccore | efetch -format acc