

Covid-19 Analysis

Genom Analysis



Use Biopython To Identity And
Perform Some Basic
Characterization Of A
Coronavirus Genome
Sequence



What is Covid-19?

Coronaviruses are a family of viruses

Coronavirus disease 2019 (COVID-19) is a contagious disease caused by a virus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China.



625,108,290

Recovered Around the world

646,687,764

Infection Around the world

6,637,957

Deaths Around the world



With the great development of technology and data analysis, which had a significant impact on keeping pace with this disease, as analysis of the virus geneom helped limit its spread and the discovery of vaccine and antibodies.

Our Services

We will help you find out



Sequence Length &

Protein Frequency



Blast search



Nucleotide Distribution





Amino Acid

Molecular Weight



Comparing & Alignments



Visualization & fasta file

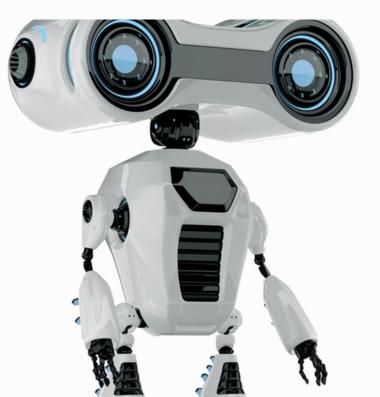
Project Tools that we use!











3. Algorithem

Python Languages

DataSet and Packages

DataSet

We analyze Covide-19 Genome Sequance from NCBI Gene Bank Database, Fasta File of SRS,MERS and covid.

Packages

We use Biopython package in python language to Analyze Sequance And Mateplotlib in Visualization our Results.



Our Results

OUR WEBSITE

Home Page



OUR WEBSITE

Home Page

Our Services

Use Biopython To Identity And Perform Some Basic Characterization Of A Coronavirus Genome Sequence

Sequence Length & GC

Finding The nucleotides number of Covid-19 genome and GC Content

Molecular Weight

Finding the Molecular Weight of Covid-19 genome

OUR WEBSITE

Analysis Results Page





Analysis Results

Use Biopython To Identity And Perform Some Basic Characterization Of A Coronavirus Genome Sequence

Information	Value
Genome Length	29903
Molecular Weight	9241219.214400413
GC Content	37.97277865097148
Count Nucleotides	{'A': 8954, 'T': 9594, 'C': 5492, 'G': 5863}
Amino Acid	9193
Proteins	409
Polypeptide	775

Covid-19 DNA Transcription

Most ten Common Amino Acids

Amino Acid	Frequency
L	886
S	810
Т	679
С	635
F	593
R	558
V	548
Υ	505
N	472
I	436

Information	Value
Genome Length	29903
Molecular Weight	9241219.214400413
GC Content	37.97277865097148
Count Nucleotides	{'A': 8954, 'T': 9594, 'C': 5492, 'G': 5863}
Amino Acid	9193
Proteins	409
Polypeptide	775



Covid-19 DNA Translation

IKGLYLPR*QTNQLSISCRSVL*TNFKICVAVTRLHA*CTHAV*LITNYCR*QDTSNSSIFCRLLTVSSVLQPIISTSR
FRPGVTER*DGEPCPWFQRENTRPTQFACFTGSRRARTWLWRLRGGGLIRGTSTS*RWHLWLSRS*KRRFAST
*TALCVHQTFGCSNCTSWSCYG*AGSRTRRHSVRS*W*DTWCPCPSCGRNTSGLPQGSSS*ER**RSWWP
*LRRRSKVI*LRRRAWH*SL*RFSRKLEH*T*QWCYP*THA*A*RRGIHSLCR*QLLWP*WLPS*VH*RPSSTCW*
SFMHFVRTTGLY*H*EGCILLP*T*A*NCLVHGTF*KEL*IADTF*N*IGKEI*HLQWGMSKFCISLKFHNQDYSTK
G*KEKA*WLYG*NSICLSSCVTK*MQPNVPFNSHEV*SLW*NFMADGRFC*SHLRILWH*EFD*RRCHYLWLLT
PKCCC*NLLSSMSQFRSRT*A*SCRIP**IWLENHSS*GWSHYCLWRLCVLLCWLP*QVCLLGSTC*R*HRL*P
YRCCWRRFRRS**QPS*NTPKRESQHQYCW*L*T**RDRHYFGIFFCFHKCFCGNCERFGL*SIQTNC*ILW*F*
SYKRKS*KRCLEYW*TEINTESSLCICIRGCSCCTINFLPHS*NCSKFCACFTEGRYNNTRWNFTVFTETH*CYD
VHI*FGY*QSSCNGLHYRWCCSVDFAVAN*HLWHCL*KTQTRP*LA*REV*GRCRVS*RRLGNC*IYLNLCL*N
CRWTNCHLCKGN*GECSDIL*ACK*IFGFVC*LYHYWWS*T*SLEFR*NICHALKGIVQKVC*IQRRNWPTHAS
KSPKRNYLLRGRNTSHRSVNRGSCLENW*FTTIRTTY**SC*SSIGWYTSLY*RAYVARNQRHRKVLCPCT*YDG
NKQYLHTQRRCTNKGYFW**HCDRSARLQECEYHF*T**KD**ST**EVLCLYS*TRYRSK*VRLCCGRCCHKN
FATSI*ITYTTGH*FR*VEYGYILLI**VW*V*IGFTYVLFFLPSR*G*RRR*L*RRRV*AINSI*VWY*R*LPR*TFGIWC

Covid-19 DNA Transcription

AUUAAAGGUUUAUACCUUCCCAGGUAACAACCAACCAACUUUCGAUCUCUUGUAGAUCUGUU
CUCUAAACGAACUUUAAAAUCUGUGUGGCUGUCACUCGGCUGCAUGCUUAGUGCACUCACGCA
GUAUAAUUAAUAACUAAUUACUGUCGUUGACAGGACACGAGUAACUCGUCUAUCUUCUGCAGG
CUGCUUACGGUUUCGUCCGUGUUGCAGCCGAUCAUCAGCACAUCUAGGUUUCGUCCGGGUGU
GACCGAAAGGUAAGAUGGAGAGCCUUGUCCCUGGUUUCAACGAGAAAACACACGUCCAACUC
AGUUUGCCUGUUUUACAGGUUCGCGACGUGCUCGUACGUGGCUUUGGAGACUCCGUGGAGGA
GGUCUUAUCAGAGGCACGUCAACAUCUUAAAGAUGGCACUUGUGGCUUAGUAGAAAGUUGAAAA
AGGCGUUUUGCCUCAACUUGAACAGCCCUAUGUGUUCAUCAAACGUUCGGAUGCUCGAACUG
CACCUCAUGGUCAUGUUAUGGUUGAGCUGGUAGCAGAACUCGAAGGCAUUCAGUACGGUCGU
AGUGGUGAGACACUUGGUGCCCUCAUGUGGCCGAAAUACCAGUGGCUUACCGCAA
GGUUCUUCUUCGUAAGAACGGUAAUAAAGGAGCUGGUGGCCAUAGUUACGGCGCCGAUCUAAA
GUCAUUUGACUUAGGCGACGAGCUUGGCACUGAUCCUUAUGAAGAAAACUGGAA
CACUAAACAUAGCAGUGGUGUUACCCGUGAACUCCUUAACGAGAGAAUUUUCAAGAAAACUGGAA
CACUAAACAUAGCAGUGGUGUUACCCGUGAACUCAUGCGUGAGCUUAACGGAGGGCAUACA
CUCGCUAUGUCGAUAACAACUUCUGUGGCCCUGAUGGCUACCCUCUUGAGUGCAUUAAAGAC



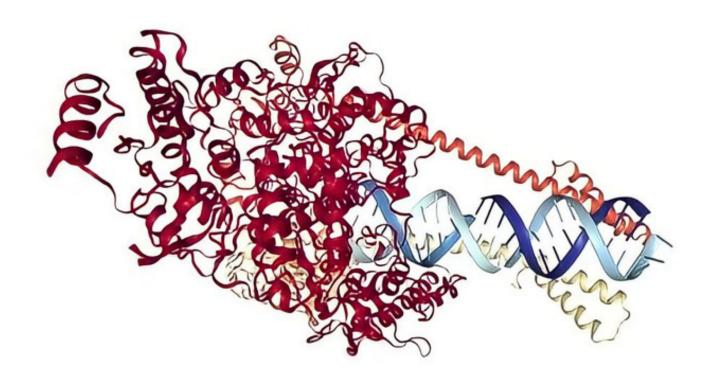
Blast Search

description:
Chain A, nsp12 [Severe acute respiratory syndrome coronavirus 2]
E value:
0.0
Bit Score:
1938.31
alignment:
Alignment with 2 rows and 925 columns
FKRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTVLQ unnamed
LNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTVLQ pdb 6YYT A
Sequence ID:pdb 6XEZ A
description:
Chain A, RNA-directed RNA polymerase [Severe acute respiratory syndrome c
E value:
0.0

First Five Polypeptide

ID	Protein
1	IKGLYLPR
2	QTNQLSISCRSVL
3	TNFKICVAVTRLHA
4	CTHAV
5	LITNYCR

Structure Of Replicating SARS-COV-2 Polymerase



Longest Protein

CTIVFKRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLIDSYFVVKRHTF SNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDTLKEILVTYNCC DDDYFNKKDWYDFVENPDILRVYANLGERVRQALLKTVQFCDAMRNAGIVGVLTLDNQDLNGNWYDFGDFI QTTPGSGVPVVDSYYSLLMPILTLTRALTAESHVDTDLTKPYIKWDLLKYDFTEERLKLFDRYFKYWDQTYHPNCV NCLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLVY AADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDG NAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKWGKARLYYDS MSYEDQDALFAYTKRNVIPTITQMNLKYAISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGATVVIGTSK FYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLARKHTTCCSLSHRFYRLANECAQVLS EMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRD VDTDFVNEFYAYLRKHFSMMILSDDAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGP HEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFHLY LQYIRKLHDELTGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIRRPFL CCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPPISFPLCANGQVFGLYKNTC



Sequence identity (%) 80 70 60 50 40 30 -20 -10 SARS/COV2 MERS/COV2 MERS/SARS

Comparison Between Covid & SARS & MERS

SARS/COV Similarity (%): 83.33837518066619

MERS/COV Similarity (%): 69.39141405757164

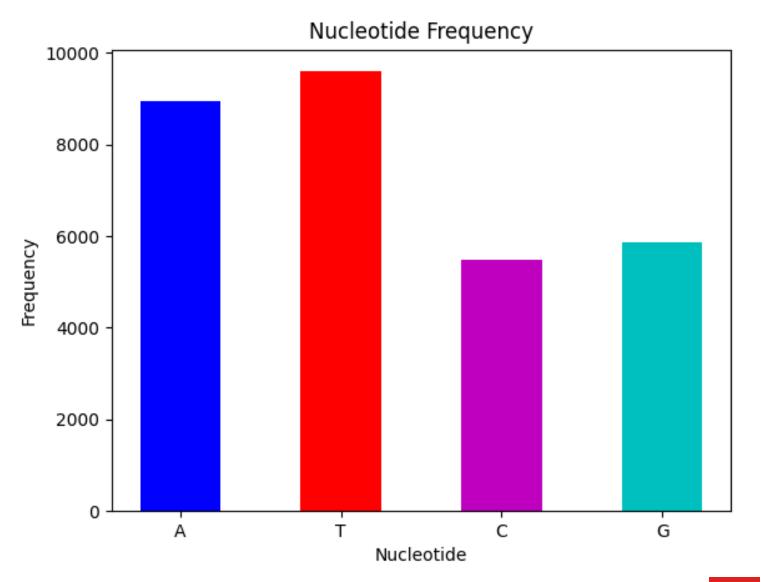
MERS/SARS Similarity (%): 69.93714496991697



Amino Acid Frequency

Protein Sequence Frequency 800 600 Frequency 200 IKGLYPRQTNSCVFAHDEWM Amino Acid

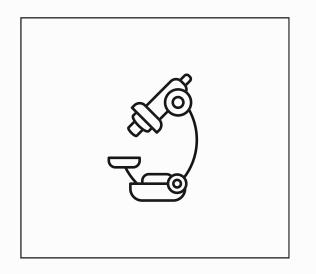
Nucleotide Frequency

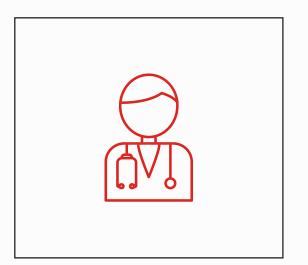


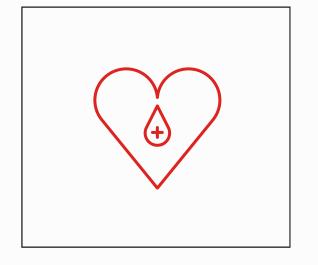


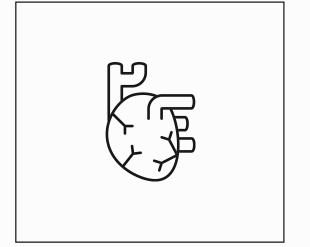
Our Source Code

https://drive.google.com/drive/folders/1m5 4GBmHwlhC4t5vvVU77RRP84SvAhPK7? usp=sharing











Team Info

- Caroline Emad
- Mary Elias
- Mariam Nader
- Kerolos Mamdouh
- Shery Rafaat
- Martina Romany
- Abdelrhman Ahmed
- Eyad Mahmoud

