GENOME SEQUENCING OF CORONAVIRUS IN SEWAGE CAN HELP DETECT LOCAL VARIANTS

A Project Report Submitted to Amrita School of Engineering in partial fulfilment of the Requirements for the Degree of Bachelor of Technology in Computer Science and Engineering (Artificial Intelligence)

INTELLIGENCE OF BIOLOGICAL SYSTEMS - 2

Submitted by

Anbazhagan E – 20001

Ramya Polaki – 20053

Trinaya Kodavati – 20069

Smithin Reddy – 20031

Abhiram Kunchapu - 20035

Under the Guidance of

Dr.Oviya I R

Amrita School of Engineering

Amrita Vishwa Vidyapeetham

Chennai – 601 103,

Tamil Nadu, India.

August 2021

ACKNOWLEDGEMENT

I offer my sincere pranams at the lotus feet of Universal guru, **Mata Amritanandamayi Devi** who blessed me with her grace to make this a successful project.

I express my deep sense of gratitude to **Dr. P Shankar,** Principal, Amrita School of Engineering, Chennai for his kind support. I would like to extend my gratitude and heartfelt thanks to **Dr. A Manikandan,** Chairperson, Department of Electrical and Electronics Engineering, for his constant help, suggestions and inspiring guidance. I am grateful to my guide **Dr. Oviya I R,** Department of Biology, for his invaluable support and guidance during the course of the major project work. I am also indebted to **Dr. Prasanna Kumar**, Chairperson of Department of AI, ASE, Chennai for his guidance.

I specially thank our director **Sri I B Manikandan** who supported us in every possible manner.

ABSTRACT

SARS stands for severe acute respiratory syndrome. Coronavirus 2 (SARS-CoV-2) is often shed in faeces during illness and viral RNA has recently been identified in sewage in certain countries. The viral genome enhanced the sensitivity of RT-qPCR tests and created nucleotide sequences with known sequence polymorphisms among SARS-CoV-2 isolates. We can see variations in viral RNA sequences over time, which corresponds to the rising worldwide dominance of the Spike protein G614 pandemic variant. Low amounts of viral RNA were also found in the catchment region of the sewage facility.

The concentration of SARS-CoV-2 RNA rose in March and April, then decreased sharply in May, demonstrating the effectiveness of lockdown measures. We conclude that viral RNA sequences may be utilised to track SARS-CoV-2 transmission, identify virus variations, and detect virus imports.

Additional variants detected in wastewater have only been identified in genomes from patients, indicating that wastewater sequencing can provide evidence for recent introductions of viral lineages before they are detected by local clinical sequencing. These results demonstrate that epidemiological surveillance through wastewater sequencing can aid in tracking exact viral strains in an epidemic context.

Keywords: COVID-19; wastewater; SARS-CoV-2 RNA; next-generation sequencing; variant G614; virus evolution; alter surveillance system

INTRODUCTION

A global pandemic of coronavirus illness (COVID-19) is now underway, driven by a novel beta-coronavirus known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The outbreak was discovered in Wuhan (China) in December 2019 and quickly expanded to 213 countries/territories, with 33.50 million confirmed cases and 1,004,421 fatalities as of September 30th, 2020. While the majority of infections cause little or moderate symptoms, some can develop to severe respiratory illness, multi-organ failure, and death.

Although there is evidence of the early detection of viral RNA in sewage even before COVID-19 cases were reported, it remains to be determined how well virus found in sewage represents virus circulating in humans and whether Environment surveillance can aid in the early detection of peaks in virus transmission for a timely effective public response.

As COVID-19 infection rates continue to grow, the need for rapid, sensitive, and dependable diagnostics to monitor patients and limit COVID-19 spread becomes more critical. Even as vaccinations are being distributed across the world, new COVID-19-related issues continue to emerge. Scientists and public health authorities have been particularly concerned about new strains of the SARS-CoV-2 coronavirus that have appeared in the United Kingdom (B.1.1.7), South Africa (B.1.351), Brazil (P.1), California (B.1.429), and New York (B.1.526).

Each variation is distinguished by a distinct collection of known changes in its genomic sequence, some of which might improve the virus's transmissibility, pathogenicity, or capacity to avoid immune responses. To properly comprehend and respond to these changes, researchers must be able to accurately determine where the variations are, when they emerge, and how frequent they become over time.

As of now we can’t go to bring waste water sample and proceed for the project

METHODOLODY

Targeted wastewater surveillance entails sampling wastewater from upstream in the wastewater network (e.g., lift stations, interceptors, manholes). Targeted wastewater surveillance may provide a better understanding of how SARS-CoV-2 infections are distributed within a sewer shed. However, there are currently little data demonstrating the application of this approach.

When deciding whether targeted wastewater surveillance would be useful for public health action, it is important to consider the following:

- SARS-CoV-2 RNA concentrations are more variable upstream from the wastewater treatment plant than at the plant intake because upstream wastewater has had less time to mix and contains faces from fewer people.
- Access to sewer lines serving only the intended target population may require infrastructure alterations or may not be possible.
- Depending on the size of the target population, conducting effective targeted wastewater surveillance may be more costly and logistically challenging than case surveillance.

How often to sample

Wastewater sampling frequency depends on how the data will be used for public health and the prevalence of COVID-19 in the community. With sufficient testing frequency, wastewater testing may be used to track trends over time. Single samples or very infrequent (e.g., monthly) sampling will likely not be informative for establishing trends, but could be used for establishing presence of COVID-19 in a community.

If the goal of wastewater surveillance is to screen for the presence of SARS-CoV-2 in wastewater, sampling once per week may be adequate. If the goal is early indication of infection trends, at least three sampling points are needed within a trend period of interest for surveillance. There are little data available describing how rapidly wastewater concentrations may change under various epidemic scenarios.

Consider the following when determining sample frequency at a specific location:

- **A minimum of three samples is required to detect wastewater trends over time.** The time between consecutive wastewater samples determines the minimum length of time over which a trend may be detected. For example, if samples are collected twice per week, 8 days is the minimum timespan over which a trend can be confirmed.
- Laboratory testing capacity and supply chain shortages may limit the maximum sampling frequency.
- One-time sampling will not provide actionable data beyond presence of SARS-CoV-2 infection within the sewer shed.

Sample type is an important consideration for collecting representative samples and will depend on the sample collection location and factors specific to the wastewater treatment plant. Closely consult with treatment plant staff to determine appropriate sample types that will best represent the target population. Samples should be collected at locations that precede addition of chemicals or mixing of waste streams at the wastewater treatment plant.

There are two wastewater surveillance sample types

1. **Untreated wastewater:** Untreated wastewater includes waste from household or building use (e.g., toilets, showers, sinks), which contains human faecal waste, as well as waste from non-household sources (e.g., rainwater, industrial use). Untreated wastewater may be sampled from wastewater treatment plant influent (prior to primary treatment) or upstream in the wastewater collection network. In most cases, untreated wastewater will likely require concentration prior to RNA extraction.
2. **Primary sludge:** Primary sludge comprises suspended solids that settle out of wastewater during the first solids removal ("sedimentation") process at a wastewater treatment plant. Primary sludge is distinct from secondary sludge following primary treatment. Do not use secondary sludge for wastewater surveillance. Changes in SARS-CoV-2 RNA concentrations in primary sludge samples have been shown to correlate with trends in reported cases. An advantage of primary sludge samples compared to untreated wastewater is that SARS-CoV-2 concentrates in sludge, which reduces the sample volume required to detect the virus and may eliminate the need to concentrate the sample prior to quantification.

Selecting a sample type Untreated wastewater and primary sludge are both acceptable community wastewater surveillance sample types. For upstream targeted wastewater surveillance, only untreated wastewater samples are available. If laboratory methods are available, sludge sampling is recommended to evaluate infection presence within a sewer shed with few known case patients because the virus will be more concentrated in sludge. Untreated wastewater samples are recommended when wastewater treatment plants apply disinfectant before sludge can be sampled, sludge testing demonstrates high assay inhibition.

There are two sample collection methods for wastewater surveillance

1. **Grab:** Grab samples can be collected rapidly and do not require automated equipment. However, grab samples may be less representative of community faecal contributions than composite samples. For untreated wastewater and sludge, grab samples represent a single moment in time and are highly influenced by daily fluctuations in wastewater flow and composition.
2. **Composite:** Composite samples are collected by pooling multiple grab samples at a specified frequency over a set time period – typically 24 hours for wastewater surveillance. You can collect composite samples of untreated wastewater manually.

Selecting a sample volume

The volume of sample to collect will depend on the sample type (wastewater or sludge). A 1 litre (L) composite wastewater sample or 100 millilitre (ml) grab sludge sample volume should be adequate for testing. The maximum amount of sludge solids that may be directly extracted is typically around 2 grams. The remaining sample volume (if any) can be used for repeat measurement or to assess biological variability.

The volume of sample that is concentrated and quantified will determine the lowest amount of SARS-CoV-2 RNA that can be detected. Concentrating more than 1 L of wastewater may result in poor recovery or viral signal inhibition. If using grab samples, consult with wastewater treatment plant staff to collect representative samples that capture peak times of human faecal loading and to understand the solids residence time for sludge.

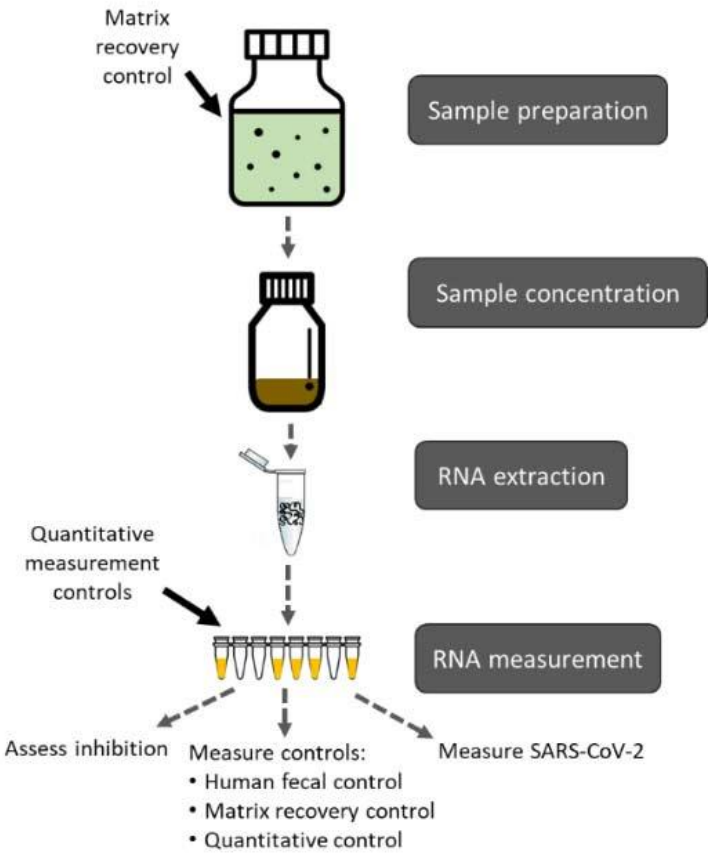
Sampling safety: There is no evidence to date that anyone has become sick with COVID-19 because of exposure to wastewater. Standard practices associated with wastewater treatment plant operations should be sufficient to protect wastewater workers from SARS-CoV-2. These standard practices can include engineering and administrative controls, handwashing, specific safe work practices, and personal protective equipment normally required when handling untreated wastewater. Beyond CDC recommendations for how to protect against COVID-19, no additional COVID-19–specific protections are recommended for workers managing wastewater, including those at wastewater treatment facilities.

Storage: Never store samples at temperatures higher than refrigeration (4°C). Refrigerate samples during the collection process. If possible, process samples within 24 hours of collection, as effective actionable wastewater surveillance relies on rapid data collection. Remaining samples can be frozen at -70°C for archiving. Avoid more than one freeze-thaw cycle. Preliminary data have shown potential loss of signal following freezing.

Shipping: When sending samples to laboratories, CDC recommends packing samples with cold packs (4°C) and using same-day or overnight shipping.

Testing methods

Multiple testing methods and laboratory workflows are used to quantify SARS-CoV-2 in wastewater. Laboratory controls can ensure that results are comparable by accounting for method performance and data quality. Based on the levels of SARS-CoV-2 in wastewater, methods can be adapted to higher or lower detection limits as needed. For example, if levels of SARS-CoV-2 RNA are sufficiently high in wastewater, small volumes of wastewater (e.g., 1 ml) may be tested without additional concentration processes. Testing methods include sample processing steps, use of laboratory controls, and implementation of biosafety measures to ensure that data can be interpreted for public health use.



Overview of wastewater sample processing and testing for SARS-CoV-2

After sample collection, the first step in SARS-CoV-2 wastewater testing is sample preparation. A matrix recovery control should be spiked into the sample during this step. The second step is sample concentration. The third step is RNA extraction from the concentrated wastewater sample. The final step is RNA measurement. Along with measurement of SARS-CoV-2 RNA in this step, several laboratory controls should also be measured, including matrix recovery controls, human faecal normalization, quantitative measurement controls, and controls to assess molecular method inhibition.

Sample processing

Sample processing for measuring SARS-CoV-2 RNA in wastewater involves sample preparation, sample concentration, RNA extraction, and RNA measurement methods. Methods selected at each step must be tailored for use with wastewater, which is a chemically and biologically complex and variable mixture. Evaluate the performance of these wastewater sample processing procedures using appropriate laboratory controls. Proper biosafety protocols for processing wastewater samples that may contain SARS-CoV-2 should be followed and are described later on this webpage.

Sample preparation

Properly storing and preparing wastewater samples helps ensure that SARS-CoV-2 RNA wastewater measurements are accurate.

- **Storage**: Refrigerate samples at 4°C immediately after collection and, if possible, process them within 24 hours to reduce SARS-CoV-2 RNA degradation and increase surveillance utility. If you cannot process samples within 24 hours after collection, you should spike a matrix recovery control into the sample prior to refrigerating it at 4°C or freezing it at -20°C or -70°C.
- **Homogenization**: Both liquid wastewater and primary sludge samples should be well-mixed prior to removing portions of collected wastewater for downstream processing. Mix by inverting samples several times (for liquid samples) or by mechanical mixing. Homogenizing samples can also include procedures to break up wastewater solids and disaggregate virus particles, such as by sonication.
- **Sample clarification**: Clarifying liquid wastewater samples by removing large solids can aid subsequent filtration-based concentration steps if they are used for sample concentration. However, removing solids will also remove SARS-CoV-2 RNA adhered to those solids. You can clarify samples using large pore size filters (5 µm or larger) or centrifugation.

Sample concentration

Concentrating wastewater samples can improve detection of SARS-CoV-2 RNA. Concentration may be more important for untreated wastewater samples than primary sludge samples. See What to Sample under ‘Developing a Wastewater Sampling Strategy’ for more information on selecting a sample type.

Concentration approaches evaluated to date that yield adequate recovery for SARS-CoV-2 detection in wastewater include:

- Ultrafiltration
- Filtration through an electronegative membrane with sample pre-treatment by addition of $MgCl_2$ or acidification
- Polyethylene glycol (PEG) precipitation

- Skim milk flocculation
- Ultracentrifugation

Consider the following factors when selecting a virus concentration method:

- **Sample type**: For untreated wastewater samples, several filtration and precipitation methods, listed above, are available. For primary sludge samples, centrifugation is the most effective way to concentrate solids.
- **Sample volume**: Large untreated wastewater sample volumes may require dividing the sample prior to membrane filtration (due to slow filtration rate) or PEG precipitation (due to centrifuge volume constraints). Sample volumes greater than 5 L may require pre-concentration by methods designed to concentrate large volume, such as large cartridge ultrafiltration.
- **Potential supply chain issues**: Methods that require commercial filtration products, such as membrane filters or ultrafiltration cartridges, may be more sensitive to supply chain issues than other methods.
- **Sample processing time**: Concentration method selection will be constrained by method processing time and availability of laboratory personnel. Membrane filtration of turbid wastewater samples may take several hours.
- **Availability of laboratory equipment**: Centrifuge volumes and force capacity, as well as availability of membrane filtration units, will also constrain method selection.

RNA extraction

Nucleic acid extraction and purification is an essential step in isolating SARS-CoV-2 RNA from the sewage mixture. Sewage is a complex mixture with materials known to interfere with molecular viral quantification methods, so consider the following when selecting an extraction method:

- Select an extraction protocol designed to produce highly purified nucleic acid extracts from environmental samples. Commercial kits are available for environmental sample extraction.
- Use an extraction kit or a protocol designed specifically to purify RNA and that includes RNase denaturants prior to lysis.

- Avoid degradation of extracted RNA due to multiple freeze-thaw cycles by aliquoting extracts into separate tubes and storing them at -70°C or below.

RNA measurement

Detection methods: Quantify SARS-CoV-2 RNA in wastewater using either RT-qPCR (reverse transcription-quantitative polymerase chain reaction) or RT-ddPCR (RT-droplet digital PCR; other forms of digital PCR are also possible but less common). Each method can be performed as either a 1-step reaction, in which RT and PCR occur in the same reaction mixture, or a 2-step reaction, in which RT and PCR are performed in separate, sequential reactions. A 1-step RT-ddPCR protocol is advantageous for wastewater because RT is performed in individual droplets, which can reduce RT inhibition compared to RT in bulk solution, as in a 2-step process and in RT-qPCR.

SARS-CoV-2 Whole-Genome Sequences Used for Nucleotide Sequence Analyses

Whole-genome SARS-CoV-2 sequences were downloaded from NCBI database (national centre for biotechnology information). Only sequences >29,000 nt in length were used in our analysis.

The RNA is reverse transcribed to DNA using a specific enzyme. Scientists then add additional short fragments of DNA that are complementary to specific parts of the transcribed viral DNA. If the virus is present in a sample, these fragments attach themselves to target sections of the viral DNA. Some of the added genetic fragments are used for building DNA strands during amplification, while the others are used for building the DNA and adding marker labels to the strands, which are then used to detect the virus.

Nested RT-PCR (nPCR) Amplification

Whole-genome SARS-CoV-2 viral RNA sequences were downloaded from NCBI database to identify suitable genetic markers to be used in our sequence analyses, specifically we looked at sequence variations observed between viral RNA sequences from England. MATLAB software was used for all nucleotide sequence analyses. Whole-genome sequences were aligned to a reference sequence (Wuhan-Hu-1 strain) with National Centre for Biotechnology Information (NCBI) accession no. NC_045512 and the frequency of sequence variation at each

nucleotide position was determined by inbuilt command for Needleman Wunsch Algorithm and smith waterman Algorithm MATLAB software default settings.

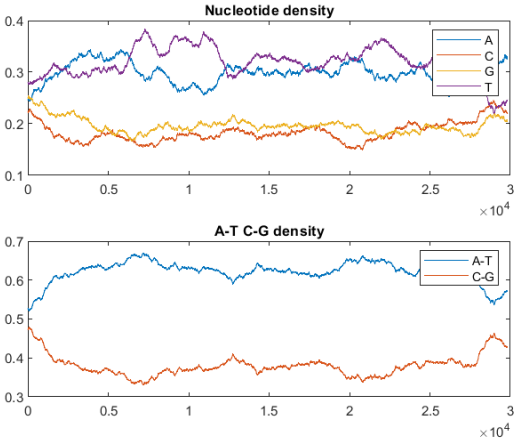
RT-PCR fragments corresponding to different regions across the SARS-CoV-2 genome were amplified from purified viral RNAs by one-step RT-PCR with Platinum Taq DNA Polymerase. Genome location and nucleotide sequences of primer sets used for the PCR reactions. Amplification conditions were: 50 ◦C for 30 min followed by 94 ◦C for 2 min plus 40 cycles of 94 ◦C for 15 s, 55 ◦C for 30 s and 68 ◦C for 8 min with a final extension step of 68 ◦C for 5 min. Following the first PCR reaction, 1 µL of amplified product was used for the second PCR reaction with the same amplification conditions used for the first PCR step. Final amplified products were purified and ready for Sanger and next-generation sequencing analysis. Primers were tested using serial dilutions of purified RNA.

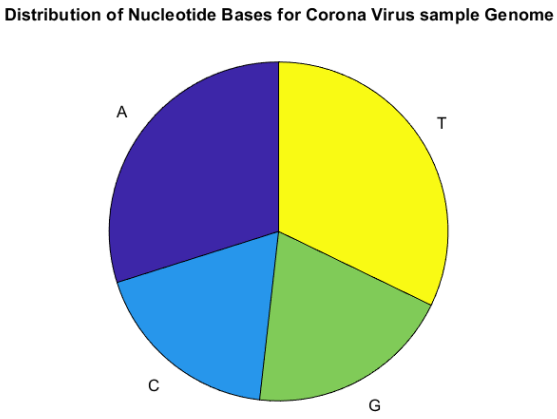
RNA extraction and no template controls were included in every assay and were always found to be negative. Primers used in this study did not closely match viral RNA sequences from seasonal coronavirus that had been circulating worldwide the last several years. Besides this, published nucleotide sequences of seasonal coronavirus serotypes in the PCR regions amplified, are at least 30% different to those from SARS-CoV-2 isolates, which means that full-sequence analysis can unequivocally demonstrate that the sequenced nPCR products from this study were from SARS-CoV-2 and not seasonal coronavirus. All nucleotide sequences of nPCR products in this study were identical or nearly identical to sequences from COVID-19 isolates from England and none resembled those from seasonal coronaviruses.

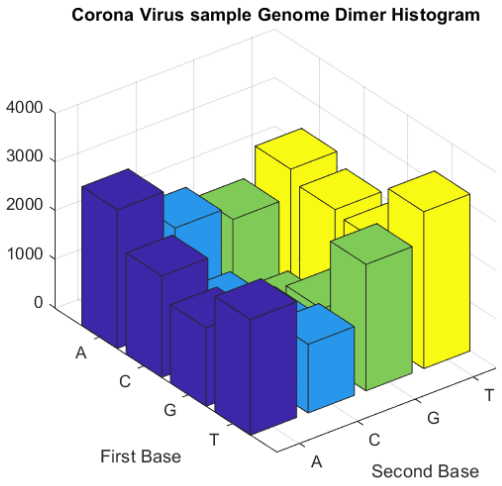
After PCR amplification the product is being used for sanger analysis and it is sequenced with the help of SNP’s.

RESULT AND DISSCUSION

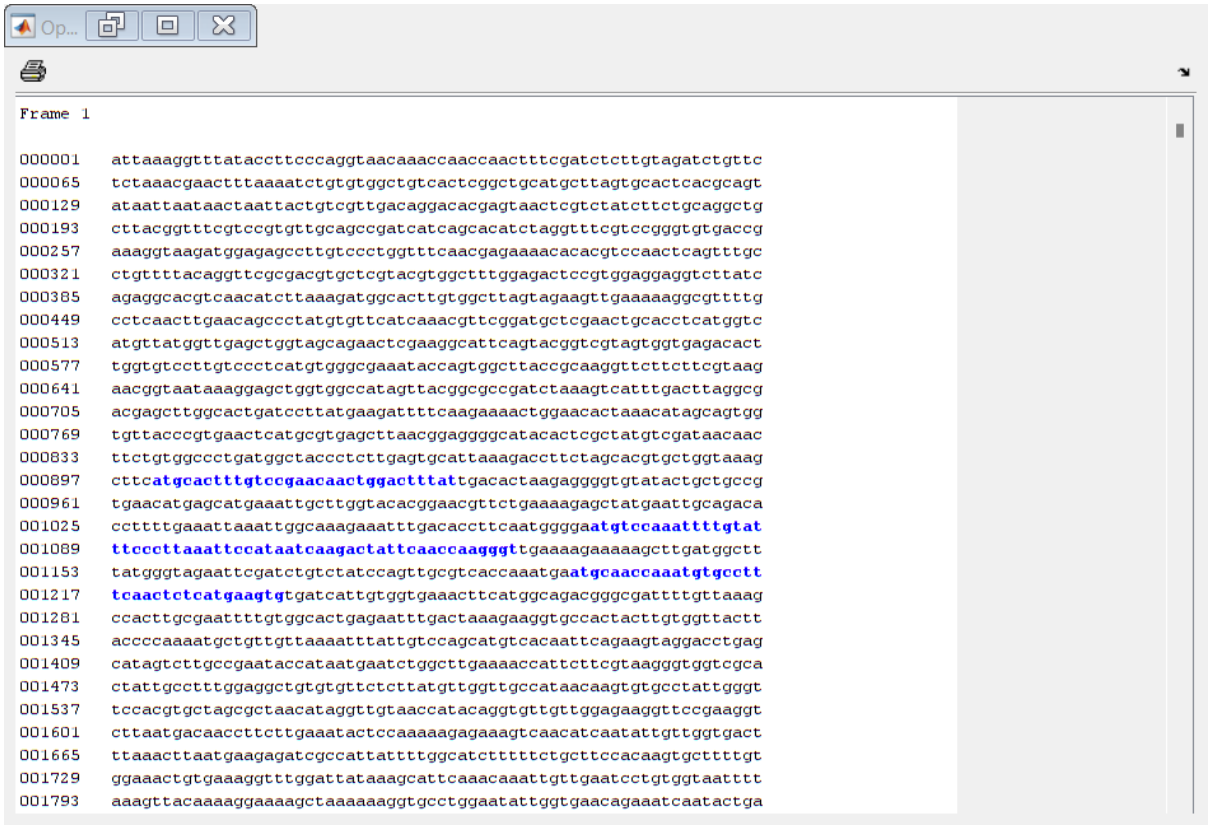
Following concentration of raw sewage, we tested a minimum of five replicate RNA samples from at least two independent wastewater concentration processes for each sample. Further replicate RNAs were tested for positive samples to obtain more accurate viral RNA quantification. SARS-CoV-2 RNA in wastewater samples was quantified using a real-time quantitative polymerase chain reaction (RTqPCR) assay targeting the RNA-dependent RNA Polymerase (RdRP) gene However, it was clear that there was a large reduction of SARS-CoV-2 RNA concentration in sewage]. Positive and negative results were independently confirmed using a second PCR. Integrity of process was confirmed through use of previous experience with enteroviruses. This was demonstrated both by detection of enteroviral RNA and recovery of infectious virus in cell cultures from all wastewater concentrates.





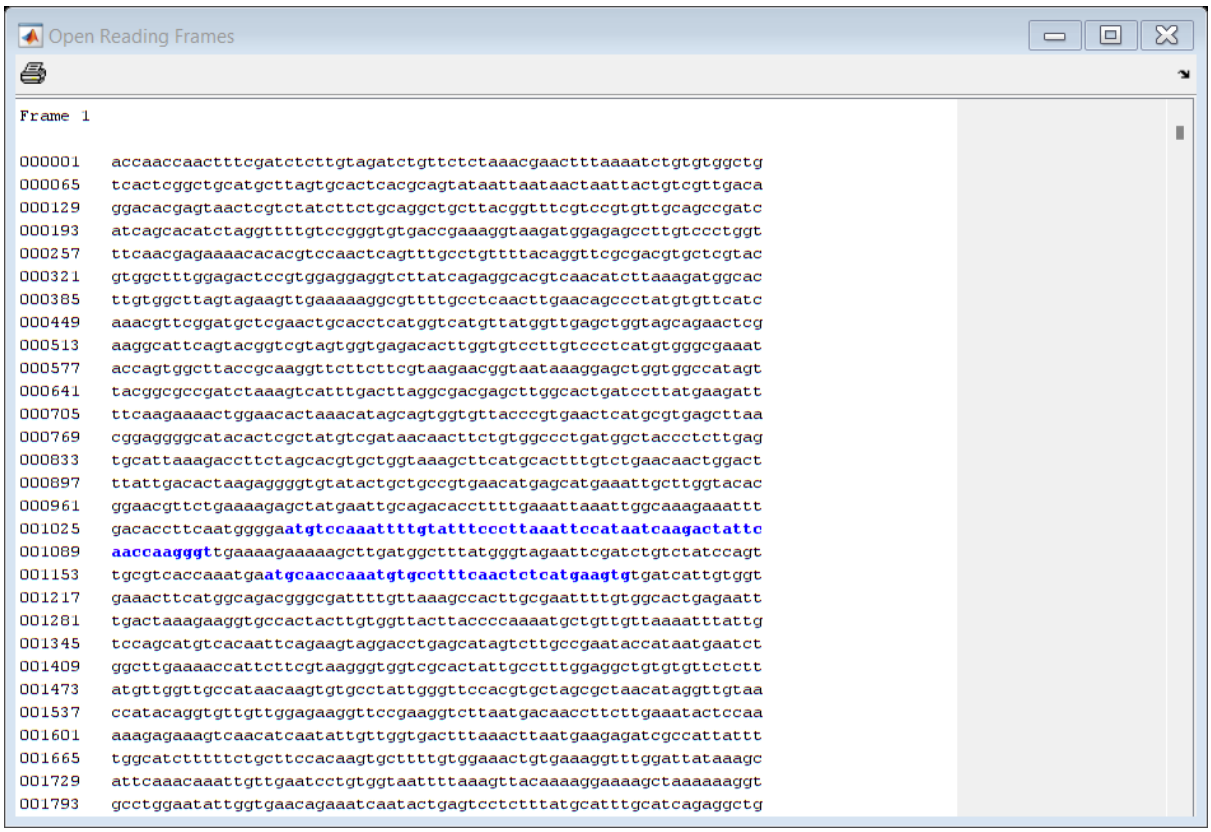


OU398982(sample) nucleotide density and its distribution



Frame 1

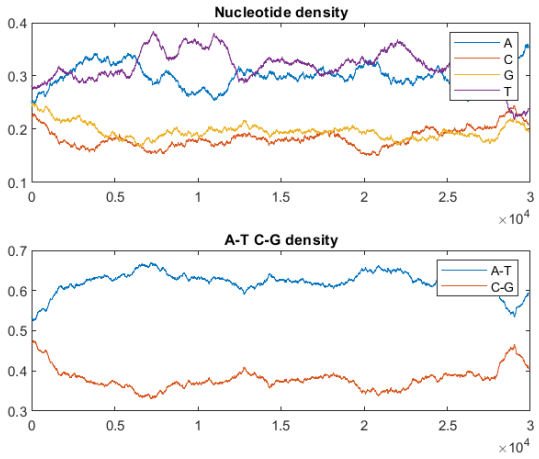
```
000001   attaaaggtttataccttcccaggtaacaaaccaaccaactttcgatctcttgtagatctgttc
000065   tctaaacgaactttaaaatctgtgtggctgtcactcggctgcatgcttagtgcactcacgcagt
000129   ataattaataactaattactgtcgttgacaggacacgagtaactcgtctatcttctgcaggctg
000193   cttacggtttcgtccgtgttgcagccgatcatcagcacatctaggtttcgtccgggtgtgaccg
000257   aaaggtaagatggagagccttgtccctggtttcaacgagaaaacacacgtccaactcagtttgc
000321   ctgttttacaggttcgcgacgtgctcgtacgtggctttggagactccgtggaggaggtcttatc
000385   agaggcacgtcaacatcttaaagatggcacttgtggcttagtagaagttgaaaaaggcgttttg
000449   cctcaacttgaacagccctatgtgttcatcaaacgttcggatgctcgaactgcacctcatggtc
000513   atgttatggttgagctggtagcagaactcgaaggcattcagtacggtcgtagtggtgagacact
000577   tggtgtccttgtccctcatgtgggcgaaataccagtggcttaccgcaaggttcttcttcgtaag
000641   aacggtaataaaggagctggtggccatagttacggcgccgatctaaagtcatttgacttaggcg
000705   acgagcttggcactgatccttatgaagattttcaagaaaactggaacactaaacatagcagtgg
000769   tgttacccgtgaactcatgcgtgagcttaacggagggggcatacactcgctatgtcgataacaac
000833   ttctgtggccctgatggctaccctcttgagtgcattaaagaccttctagcacgtgctggtaaag
000897   cttcatgcactttgtccgaacaactggactttattgacactaagaggggtgtatactgctgccg
000961   tgaacatgagcatgaaattgcttggtacacggaacgttctgaaaagagctatgaattgcagaca
001025   ccttttgaaattaaattggcaaagaaatttgacaccttcaatggggaatgtccaaattttgtat
001089   ttcccttaaattccataatcaagactattcaaccaagggttgaaaagaaaaagcttgatggctt
001153   tatgggtagaattcgatctgtctatccagttgcgtcaccaaatgaatgcaaccaaatgtgcctt
001217   tcaactctcatgaagtgtgatcattgtggtgaaacttcatggcagacgggcgattttgttaaag
001281   ccacttgcgaattttgtggcactgagaatttgactaaagaaggtgccactacttgtggttactt
001345   accccaaaatgctgttgttaaaatttattgtccagcatgtcacaattcagaagtaggacctgag
001409   catagtcttgccgaataccataatgaatctggcttgaaaaccattcttcgtaagggtggtcgca
001473   ctattgcctttggaggctgtgtgttctcttatgttggttgccataacaagtgtgcctattgggt
001537   tccacgtgctagcgctaacataggttgtaaccatacaggtgttgttggagaaggttccgaaggt
001601   cttaatgacaaccttcttgaaatactccaaaaagagaaagtcaacatcaatattgttggtgact
001665   ttaaacttaatgaagagatcgccattattttggcatctttttctgcttccacaagtgcttttgt
001729   ggaaactgtgaaaggtttggattataaagcattcaaacaaattgttgaatcctgtggtaatttt
001793   aaagttacaaaaggaaaagctaaaaaaggtgcctggaatattggtgaacagaaatcaatactga
```

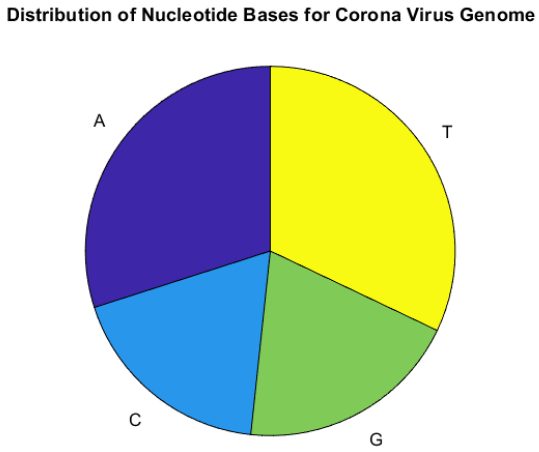


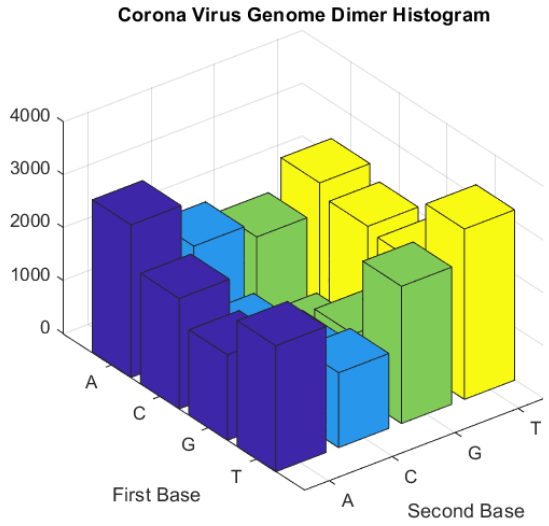
Open Reading Frames

Frame 1

```
000001   accaaccaactttcgatctcttgtagatctgttctctaaacgaactttaaaatctgtgtggctg
000065   tcactcggctgcatgcttagtgcactcacgcagtataattaataactaattactgtcgttgaca
000129   ggacacgagtaactcgtctatcttctgcaggctgcttacggtttcgtccgtgttgcagccgatc
000193   atcagcacatctaggttttgtccgggtgtgaccgaaaggtaagatggagagccttgtccctggt
000257   ttcaacgagaaaacacacgtccaactcagtttgcctgttttacaggttcgcgacgtgctcgtac
000321   gtggctttggagactccgtggaggaggtcttatcagaggcacgtcaacatcttaaagatggcac
000385   ttgtggcttagtagaagttgaaaaaggcgttttgcctcaacttgaacagccctatgtgttcatc
000449   aaacgttcggatgctcgaactgcacctcatggtcatgttatggttgagctggtagcagaactcg
000513   aaggcattcagtacggtcgtagtggtgagacacttggtgtccttgtccctcatgtgggcgaaat
000577   accagtggcttaccgcaaggttcttcttcgtaagaacggtaataaaggagctggtggccatagt
000641   tacggcgccgatctaaagtcatttgacttaggcgacgagcttggcactgatccttatgaagatt
000705   ttcaagaaaactggaacactaaacatagcagtggtgttacccgtgaactcatgcgtgagcttaa
000769   cggaggggcatacactcgctatgtcgataacaacttctgtggccctgatggctaccctcttgag
000833   tgcattaaagaccttctagcacgtgctggtaaagcttcatgcactttgtctgaacaactggact
000897   ttattgacactaagaggggtgtatactgctgccgtgaacatgagcatgaaattgcttggtacac
000961   ggaacgttctgaaaagagctatgaattgcagacacctttgaaattaaattggcaaagaaattt
001025   gacaccttcaatggggaatgtccaaattttgtatttcccttaaattccataatcaagactattc
001089   aaccaagggttgaaaagaaaaagcttgatggctttatgggtagaattcgatctgtctatccagt
001153   tgcgtcaccaaatgaatgcaaccaaatgtgcctttcaactctcatgaagtgtgatcattgtggt
001217   gaaacttcatggcagacgggcgattttgttaaagccacttgcgaattttgtggcactgagaatt
001281   tgactaaagaaggtgccactacttgtggttacttaccccaaaatgctgttgttaaaatttattg
001345   tccagcatgtcacaattcagaagtaggacctgagcatagtcttgccgaataccataatgaatct
001409   ggcttgaaaaccattcttcgtaagggtggtcgcactattgcctttggaggctgtgtgttctctt
001473   atgttggttgccataacaagtgtgcctattgggttccacgtgctagcgctaacataggttgtaa
001537   ccatacaggtgttgttggagaaggttccgaaggtcttaatgacaaccttcttgaaatactccaa
001601   aaagagaaagtcaacatcaatattgttggtgactttaaacttaatgaagagatcgccattattt
001665   tggcatctttttctgcttccacaagtgcttttgtggaaactgtgaaaggtttggattataaagc
001729   attcaaacaaattgttgaatcctgtggtaattttaaagttacaaaaggaaaagctaaaaaaggt
001793   gcctggaatattggtgaacagaaatcaatactgagtcctctttatgcatttgcatcagaggctg
```

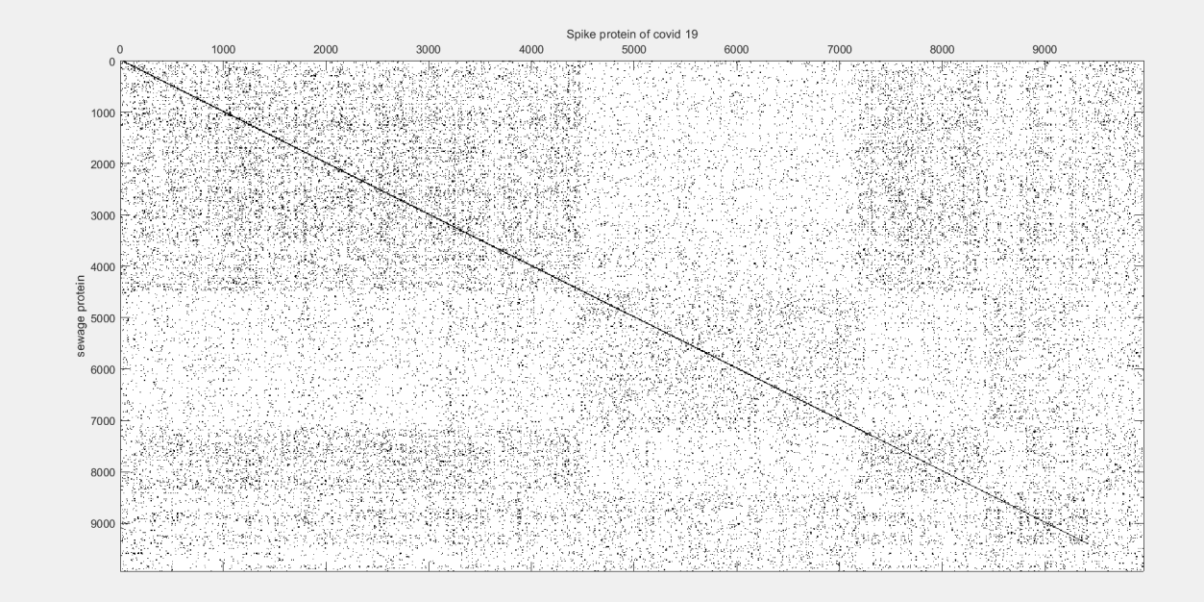




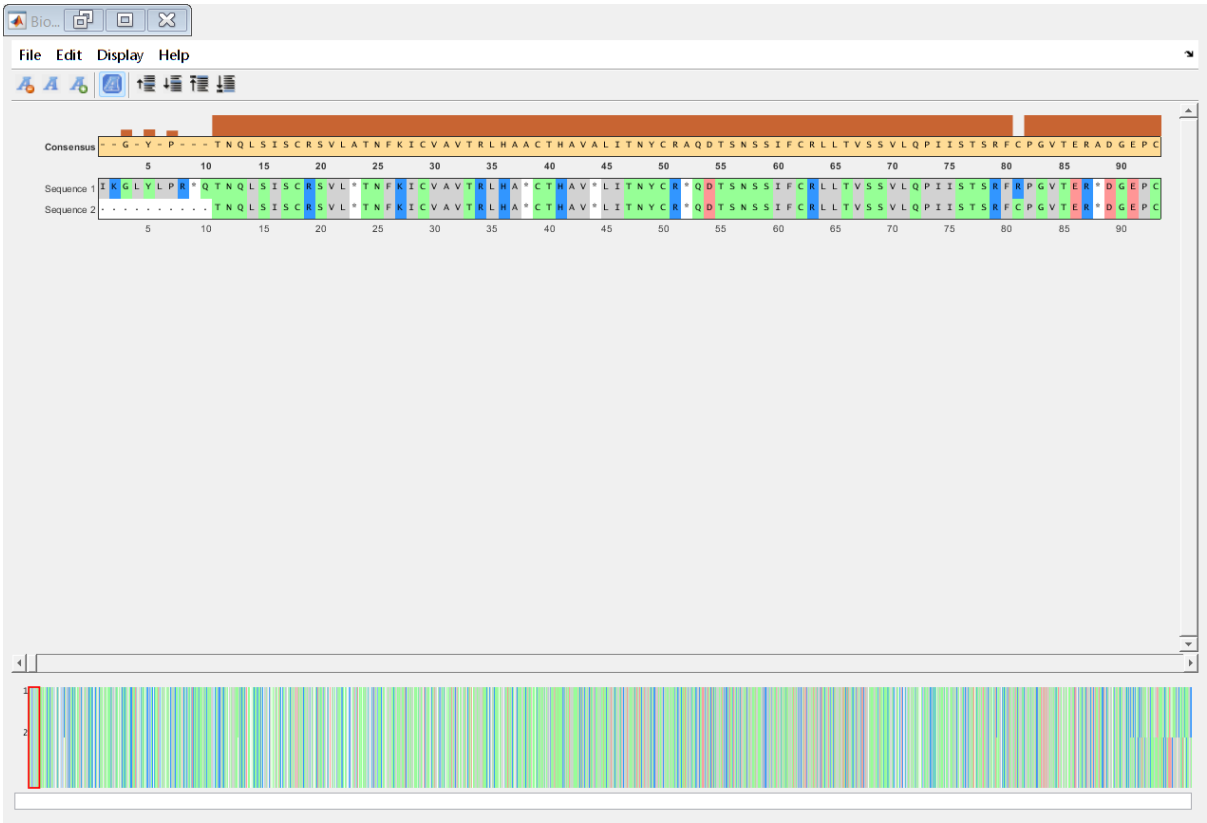


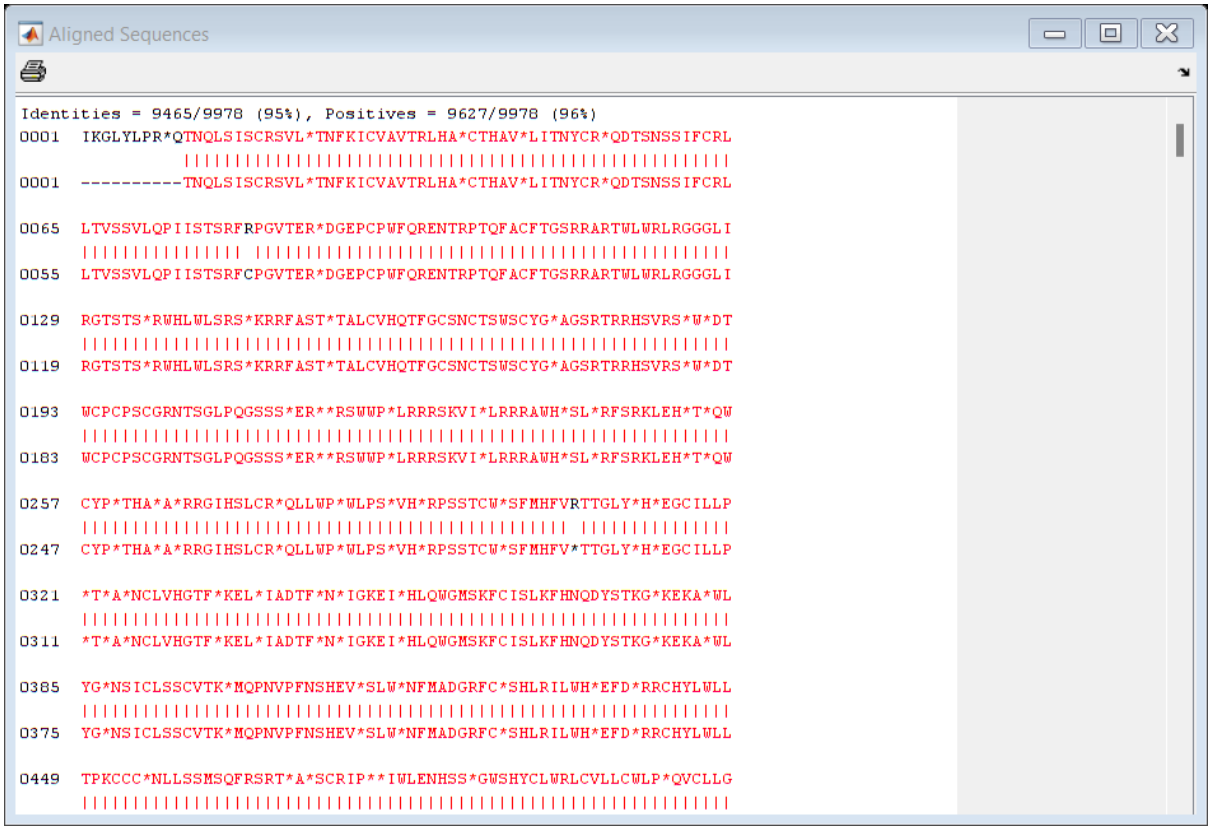
NC_045512(reference) nucleotide density and its distribution

We considered two genome sequences Accession number OU398982(Sample) and NC_045512(original sequence for reference). By using the MATLAB software, the sample and reference is been sequenced and with the inbuilt commands of bioinformatics toolbox its amino acid sequence is generated. With Needleman Wunsch and smith waterman Algorithms we aligned those two amino acids sequenced synthesised from sample and reference DNA. Needleman Wunsch algorithm will return us the global alignment, which is the comparison of whole genomic sequence whereas smith waterman algorithm will return us a local alignment which is a comparison only among the paired sequence. So, indeed the later will return a higher percentage of matches whereas the prior will be of some approximated percentage. From this analysis of MATLAB, we can say the variant's matching probability with the original sequence of Sars COV2 virus.

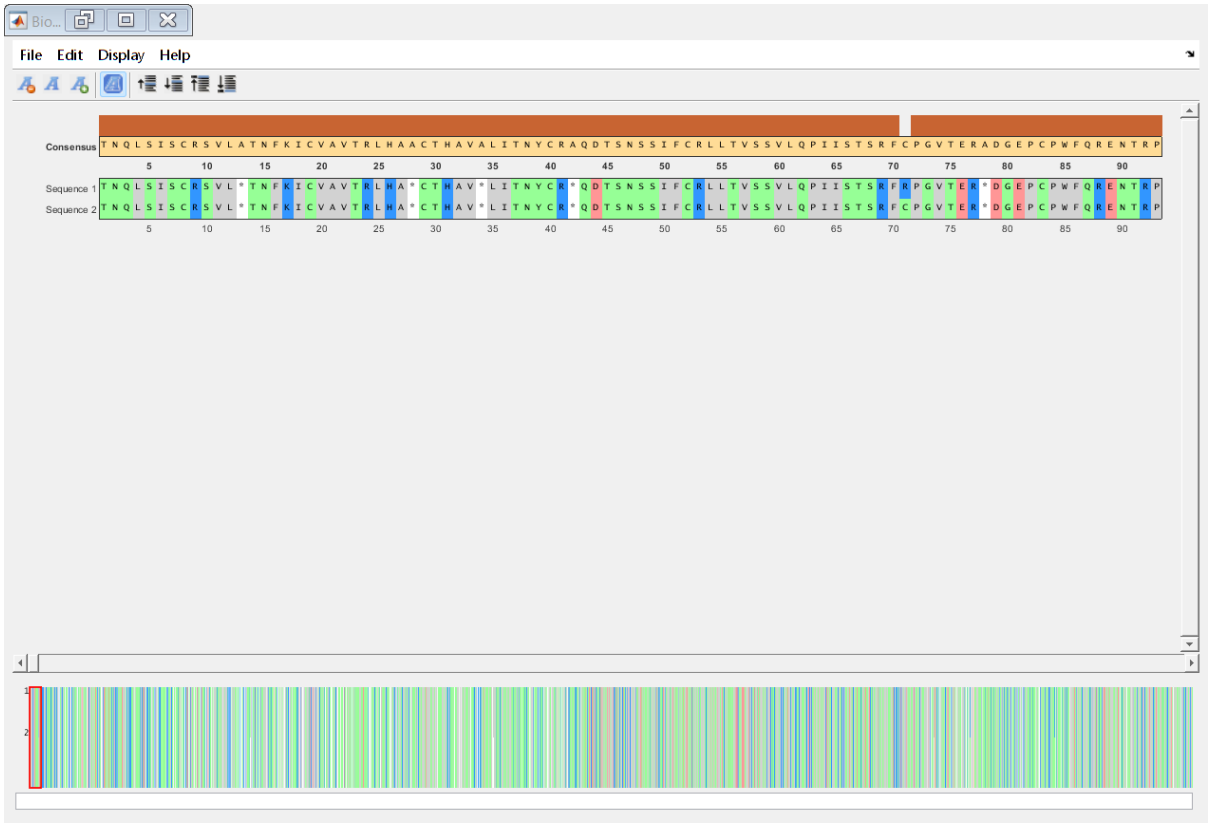


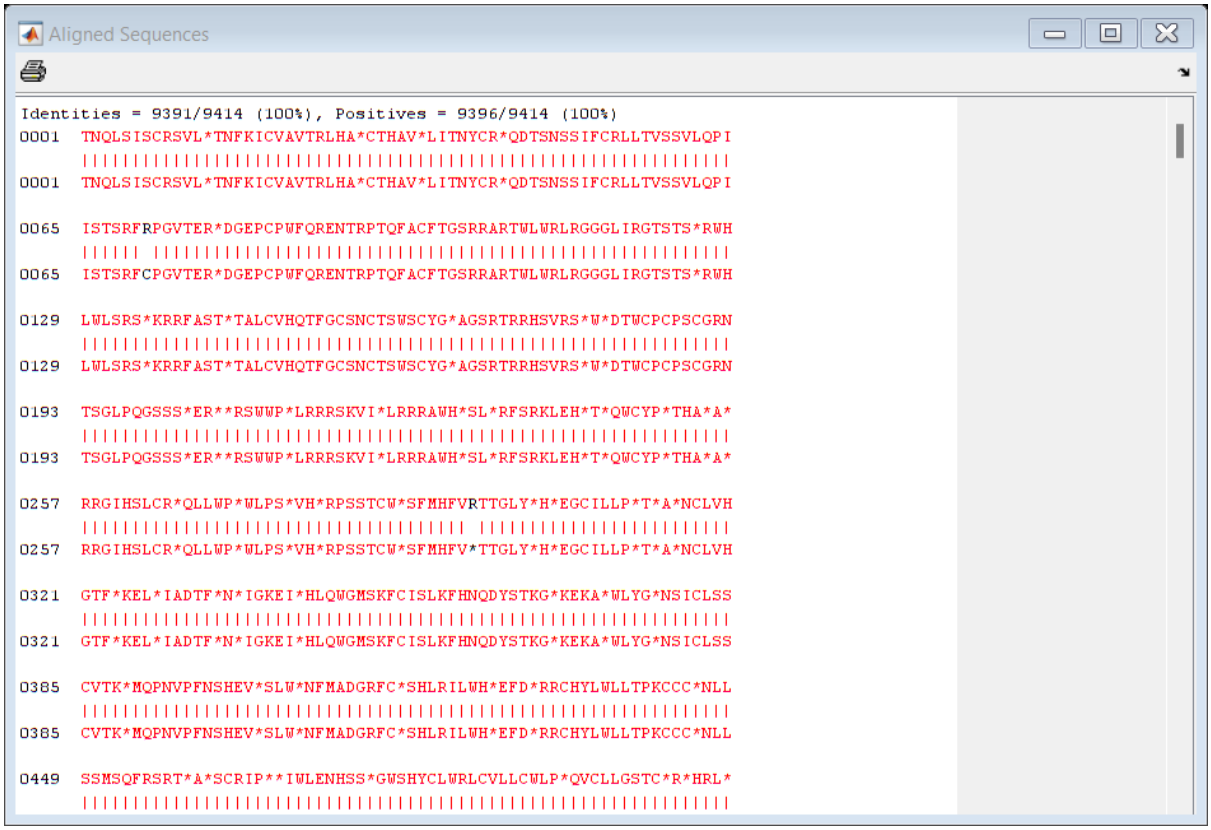
GLOBAL ALIGNMENT





LOCAL ALIGNMENT





CONCLUSION

We have shown that environmental surveillance can be used to monitor SARS-CoV-2 transmission detecting virus variants specifically circulating around the world and identifying changes in virus variant predominance known to have occurred during the COVID-19 pandemic. Environmental surveillance could be used for the early detection of peaks in virus transmission for public health interventions to be timely implemented.

one of the biggest obstacles the public is still facing is detecting tracking and tracing the spread of covid19 there's increasing realization in the medical and scientific communities that they're just not going to be able to test every individual the logistical barriers are just too high so for an effective surveillance method. we need a way to detect covid19 in communities that is both accurate and easily deployable this is where wastewater testing comes in imagine being able to test a dormitory with dozens of hundreds a town or city with thousands of millions of people all in one pooled sample.

our wastewater system allows us to do this the analysis of wastewater looking for biological or chemical compounds is known as wastewater-based epidemiology which has been used already looking for drugs for pharmaceutical and industrial waste and bacteria and viruses as pathogens in wastewater and water supplies.

The goal of doing wastewater testing or wastewater surveillance for covid19 is to create an early warning system in order to better stop the spread of the virus. within a week of the outbreak for the first patients since then several studies have actually been published talking about some of the trade-offs the pluses wastewater testing.

REFERENCES

- **file:///C:/Users/hi/Downloads/Genome_Sequencing_of_Sewage_Detects_Regionally_Pre%20(2).pdf**
- **https://www.news-medical.net/news/20200827/Two-viral-gene-targets-needed-in-SARS-CoV-2-detection-by-PCR-say-researchers.aspx**
- **https://www.ncbi.nlm.nih.gov/nuccore/NC_045512**
- **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7384408/**
- **https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/testing-methods.html**
- **https://www.promega.in/applications/virus-detection-assay-coronavirus-detection-covid-19-sars-cov-2/wastewater-based-epidemiology-covid19/**