

Fabiola Cruz-López¹, Karolane González¹, Vianca Aponte-Rolón¹, Andrea Firpo¹, Fabiola Fontanet-Jaime¹, Natalie Machargo¹, Tania Mitwalli¹, Patricia Serrano¹, Ernesto Vázquez¹, Martin González², Kendall Billig³, Nathan Grubaugh³, David E. Rivera-Aponte¹

¹Universidad Central del Caribe, Bayamón, PR; ²Universidad de Puerto Rico- Recinto de Río Piedras, SJ; ³Yale University, CT

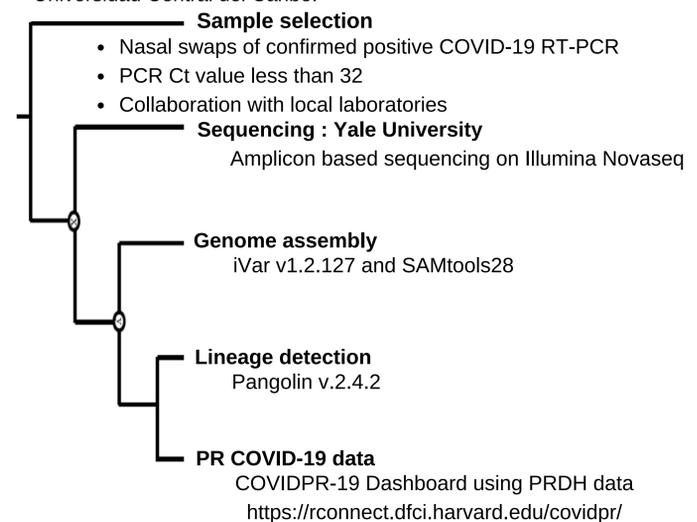
Introduction

The emergence of SARS-CoV-2 caused a significant pandemic that drove scientists to use advanced technology such as Next Generation Sequencing (NGS) to develop vaccines, treatments for mitigation, and to study the evolution in real-time to identify variants of interest (VIC) or concern (VOC) to public health. The SARS-CoV-2 virus is spread mainly by respiratory droplets and contaminated surfaces and causes coronavirus disease -19 (COVID-19). Infected individuals mainly develop respiratory symptoms like shortness of breath, fever, and pneumonia. Moreover, the severity of COVID-19 lead to hospitalization, death, long-term symptoms persisting for more than six months, and is a risk factor for the development of cardiovascular diseases, obesity, and other chronic diseases. From 2020 to 2021, the virus SARS-CoV2 caused more than 200 million cases and 4 million deaths. Also, more than ten variants of SARS-CoV-2 have been monitored due to its impact on the epidemic behavior and severity. Thus, increasing the importance of implementing genomic surveillance as a public health tool, not only for identification and comprehension of the transmission pattern but also for preventing the dissemination of the virus.

By February 2022, Puerto Rico had reported nearly 270,000 SARS-CoV-2 confirmed cases and more than 4,000 deaths. Although sequencing efforts started early in March 2020, genomic surveillance is still poor, with spatiotemporal bias and high turnaround times (TAT) compared to other high-income countries. The data available is helpful for retrospective molecular epidemiology studies but is scarce for real-time genomic surveillance. Therefore, public health policies for managing the pandemic in Puerto Rico mainly consider other epidemiological parameters such as positivity rates, incidence, and hospital capacity. A strengthening of the genomic surveillance system will provide Puerto Rico with a tool for the early detection of VOIs and VOCs, helping to establish a timely measure to control virus transmission before cases begin to increase. Thus, in our study, we established a small-scale genomic surveillance system in Puerto Rico through a collaborative effort between Universidad Central del Caribe, Yale University, and local laboratories. Here, we will be describing the results and contributions of this effort to strengthen genomic surveillance on the island.

Materials and Methods

This study followed the approved IRB protocol 2021-21 from the Universidad Central del Caribe.



Results

Our small-scale genomic surveillance sequenced a total of 682 genomes collected between May 5, 2021, and October 5, 2021. The genomic surveillance officially started on June 15, 2021, with weekly sampling collection. However, in the first batch of samples, we included 72 genomes collected from May 5 to June 14, 2021, as part of an onboarding process.

Enhancing the capacity of genomic surveillance: turnaround time, percentage of confirmed cases sequenced, and spatial representation

Table 1. Description of sequences from Puerto Rico reported in GISAID: TAT and percentage of confirmed cases sequenced.

Period	March 13, 2020 - June 14, 2021	June 15 (Week 24) - October 5 (Week 40), 2021 (All genomes)	June 15 (Week 24) - October 5 (Week 40), 2021 (without UCC-Yale)	June 15 (Week 24) - October 5 (Week 40), 2021 (UCC-Yale)
Number of sequences	1825 genomes	1722 genomes	1115 genomes	607 genomes
TAT median	40 days	27 days	29 days	21 days
TAT mean	74 days	29 days	33 days	22 days
Percent age of sequences with a TAT ≤ 21 days (# of sequences with a TAT ≤ 21/total # genomes)	(376/1825) = 20.6 %	(560/1722) = 32.5%	(196/1115) = 17.6%	(364/607) = 60%
Percent age of confirmed cases represented by sequences with a TAT ≤ 21 days (# of sequences with a TAT ≤ 21/total # confirmed cases)	(376/133,829) = 0.28%	(560/29,215) = 1.9%	(196/29,215) = 0.7%	(364/29,215) = 1.2%

Figure 1. Percentage of cases sequence in Puerto Rico reported in GISAID with TAT<21 days, during summer 2021

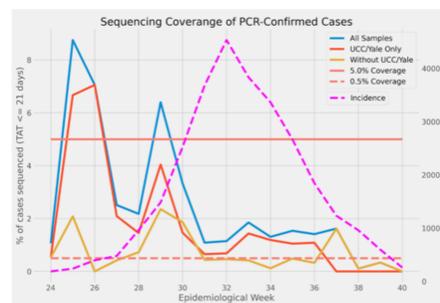


Figure 2. Weekly prevalence of SARS-CoV-2 variants in Puerto Rico, during summer 2021

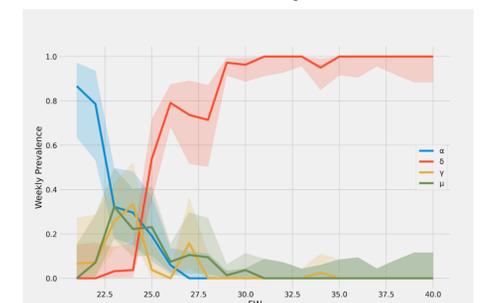
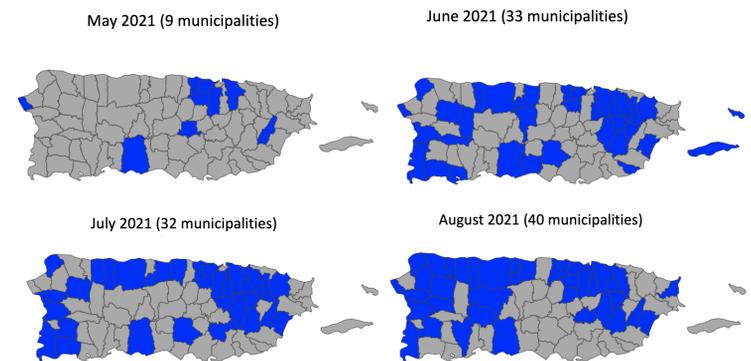


Figure 3. Spatiotemporal representation of genomes sequenced in Puerto Rico from May to August 2021.



Discussion

Our study was the first effort in Puerto Rico to monitor the fixation of the Delta variant in real-time and improve the genomic surveillance response. We successfully reduced the median TAT by two days (from 29 to 27), and our data alone achieved the median of 21 days. An optimistic scenario should have a turnaround time (TAT) median of < 21 days and 0.5% of confirmed reported cases. These parameters provide a 20% probability of identifying a new variant before it reaches 100 cases. Also, at least 60% of our sequences were reported within the 21 days compared to only 17.6% from other efforts.

The application of sequencing technology to study the genetic diversity and characteristics of SARS-CoV-2 is a powerful tool in identifying and analyzing epidemiologic patterns for the rapid development of strategies to manage and control possible outbreaks. However, we must have sequenced at least 0.5% of confirmed cases, and more than 5% should be expected in countries with resources. Our sample represented 1.2% of the cases reported during the study period compared to 0.7% from all other efforts. Although we did not implement the small-scale system at the time for the early detection of Delta, we were able to monitor its establishment and its effects on the pandemic behavior on the island. Lastly, in GISAID, our samples were the only ones containing information about the municipality. We covered at least 57 of 78 municipalities during the period. There is still a need to increase accessibility to molecular testing to provide a better geographical distribution of genomes for VOI/VOC.

The collaboration between Universidad Central del Caribe, Yale University, and local laboratories was crucial to enhancing the island's capacity for genomic surveillance. We provided real-time data that helped identify the establishment and prevalence of Delta early in June and alert the public health authorities in the first week of July about the possibility of a sudden increase in cases. By week 32, Delta represented more than 90% of the sequenced cases. The increase of cases could be explained by the establishment of Delta and other events such as vaccine immunity waning. This variant has increased transmissibility compared to different variants, including Alfa, which already had increased transmission compared to the original strain. Early detection of these patterns was critical in the following step strategies taken by the Puerto Rico Department of Health (PRDH).

Conclusion

High-Income Countries (HIC) are more likely to have robust genomic surveillance for SARS-CoV-2 due to the preexisting infrastructure of laboratory detection and surveillance capacity for other viruses such as influenza. However, in Puerto Rico, classified as HIC, the genomic surveillance capacity is still suboptimal. Before this study, local efforts mainly focused on other genomic activities, such as describing the emergence of past viral lineages in retrospective analyses.

For four months, our small-scale genomic surveillance succeeded at weekly reporting genomes to PRDH to guide immediate public health strategies, reduce TAT and improve coverage of cases during the establishment of Delta. The translation of our project into a substantial organized effort to fund academic entities and local laboratories will facilitate obtaining real-time genomic surveillance results and ensure rapid identification of new variants of concern to public health in Puerto Rico.

Acknowledgments

We want to thank our collaborators at Yale School of Public Health, Dr. Nathan Grubaugh's lab. Also, we thank the data scientists, Mr. Roberto A. Calzadilla, for the assistance in the analysis. Finally, to our funding institutions: Yale University, Universidad Central del Caribe and INBRE (P20 GM103475-15).