Title: The role of SARS-CoV-2 genomic surveillance and innovative analytical platforms for informing public health preparedness in Bengaluru, India.

Authors

Pulleri Kandi Harsha1, Chitra Pattabiraman1, Anson Kunjumon George1, Madhusudhan J1, Saumitra Mardikar1, Mudasir Nazaar4, Srivathsan Adimoolam1, Divya Priya A1, Jayee Pitale1, Mahesh Nagarajan1, Srikant Sridharan1, Yasodha Kannan1, Anand Janakiraman1, Vamsi Veeramachaneni1, Ramesh Hariharan1,2, Vishal US Rao3,5, Mohammed Shariff7, Thrilok Chandra4, K. Sudhakar6, Randeep Dev5, Vasanthapuram Ravi5, Vijay Chandru1,2,5*

*Corresponding author: Vijay Chandru, vijaychandru@iisc.ac.in

1. Strand Life Sciences Private Limited, Bengaluru, India
2. Indian Institute of Science, Bengaluru, India
3. Centre for Academic Research, Health Care Global Cancer Hospital, Bengaluru, India.
4. Bruhat Bengaluru Mahanagara Palike (BBMP), Bengaluru, India
5. Genomic Surveillance Committee, Karnataka, Health & Family Welfare Services, Karnataka, India
6. Minister for Medical Education & Health & Family Welfare, Government of Karnataka, India
7. Department of health and family welfare services, Karnataka

Abstract

A comprehensive SARS-CoV-2 genomic surveillance programme that integrates logistics, laboratory work, bioinformatics, analytics, and timely reporting was deployed through a public-private partnership in the city of Bengaluru, Karnataka in India. As a result, 12461 samples have been sequenced and reported to the Karnataka State public health officials as time-sensitive, decision support during the last one year and uploaded in global public databases in a timely manner. This programme has developed an analytics platform for studying SARS-CoV-2 sequences and their epidemiological context. Continuous sequencing effort enabled timely detection of emergence of Omicron variant in India and the subsequent spread of the same and its sub-lineages with more logistic growth (BA.10, BA.12 and BA.5) in Bengaluru. Our data also helped to provide timely information on variants to determine which of the Variants of Concern

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
tracked globally, were observed in Bengaluru, ensuring targeted efforts and reducing unwarranted fear. This effort highlights the importance of, and the urgent need to, increase genomic surveillance to support the states with limited sequencing and bioinformatics capacity. We describe the development and deployment of this end-to-end solution for genomic surveillance of SARS-CoV-2 in the city of Bengaluru.

**Keywords:** Genomic surveillance, SARS-CoV-2, COVID-19, analytical platforms, pandemic preparedness

**Introduction:**

Following the first pandemic wave of coronavirus disease 2019 (COVID-19), the introduction and spread of SARS-CoV-2 variants have resulted in new waves of infections across the globe. Some SARS-CoV-2 variants disappeared immediately, whereas others characterized by several key mutations adapted well, enabling their rapid spread (1). Genomic surveillance has shaped our understanding of the SARS-CoV-2 variants and their transmission dynamics (2).

The most important contribution of sequencing efforts worldwide has been the detection of Variants of Concern (VoC), including Alpha, Beta, Delta, and more recently Omicron (3). These are variants of SARS-CoV-2 that have been reported to have either increased transmission, increased evasion of immune response (from natural infection or existing vaccines) or associated with severe disease (4). The rapid expansion of Omicron into multiple lineages and the unprecedented community spread is likely to give rise to other variants that need to be monitored for their abilities to cause severe disease, evade immunity and transmit better (5).

Viruses mutate all the time but only some mutations affect their ability to spread or evade prior immunity from vaccination or infection, or the severity of disease they cause (4). Currently there are over 70 Omicron sub-lineages detected all over the world. The existing data suggests BA.2 and sub-lineages are more transmissible than other Omicron lineages (6,7), and are predominant in India.
The mutation rate and the impacts of mutation on viral dynamics within and between individual hosts determine the emergence and spread of viral variants and the evolution of pandemics (8). This warrants continued genomic surveillance particularly in densely populated areas. Staying abreast of virus mutations and lineages is a critical aspect of public health management.

There are now over 11 million sequences of SARS-CoV-2 in the global GISAID database (9). These sequences have helped understand how the virus is introduced into a population/location, how it spreads and whether there are mutations or changes that affect properties of the viruses (10–12). The early detection of a VoC is important for public health response as advance warning can prepare health systems for a surge.

In the early stages of the pandemic in India, Karnataka and its capital city Bengaluru (population 13.1 million) as well as Delhi, the national capital region, were exceptional with early surveillance of the SARS-CoV2 virus samples that were coming in mostly through international travellers (13,14). Contact tracing and line lists were readily available to the researchers who carried out fairly detailed genomic epidemiology of the early onset of infections and diffusion to the locations around the state. This early momentum faded however and through the deadly second wave of the “Delta” variant, the sequencing data was lagging behind the infection and the alacrity that was needed for surveillance and public health response was missing. Therefore by the middle of 2021 when the second wave subsided, the scaling up genomic sequencing capacity was a key requirement for public health officials in Karnataka. The state appointed COVID-19 Genomic Surveillance Committee for Karnataka began its meetings in June 2021 and the qualification of Strand Life Sciences, a private laboratory of repute in the city to accelerate the sequencing efforts was tabled on June 25th 2021.

Strand Life Sciences was qualified soon after by the surveillance committee and recognized by the Indian SARS-CoV-2 Genomics Consortium (INSACOG) as a private laboratory authorized to participate in genomic surveillance. Strand Life Sciences proceeded to interact with the city and state health authorities. Philanthropic funding was secured to execute the large-scale sequencing and data analysis at the Strand Labs. This ambitious project successfully enabled monitoring at around 10% of positive cases in the city with a turnaround time of 7 to 10 days from sample collection to dashboard presentations for the Karnataka State public health officials, who were the
key stakeholders for this project. The logistics of sample collection, sequencing, informatics, analytics and timely reporting presented daunting challenges that needed continued support for successful execution. This article describes the details of this effort that largely covered the surveillance of the entire “third wave” of COVID-19 infections in the city of Bengaluru in Karnataka. It was significant that this was accomplished in a timely and cost-effective manner (at under USD 80 per sample for logistics, sequencing, data analysis and decision support informatics) by Strand Life Sciences, a private sector lab, in conjunction with private philanthropy and with direct support of the stakeholders from the state and national health authorities thereby suggesting a new paradigm for public health interventions in LMICs.

**Materials and Methods:**

Technical Advisory Committee (TAC) of the State for Pandemic response mandated Strand Life Sciences to initiate the genomic Surveillance of SARS-CoV-2 in various parts of Bengaluru city (Ref: CHO (PH)/PR/P-103/2021-22), to detect trends in mutations and to avoid the impact seen in the second wave of COVID-19 in Karnataka.

Sequencing was carried out on SARS-CoV-2 RT PCR positive samples with Ct value < 30 from the various parts of the city of Bengaluru, India. Samples were collected from multiple wards across the city. Sampling was opportunistic – sequencing of samples that are available and fit the inclusion criteria with the overall aim of high-density sampling within Bengaluru City every week (continuous over time). The focus was to include sub-groups – such as children and hospitalized individuals and also collect information on location and vaccination status.

SARS-CoV-2 genomes are recovered by amplicon-based sequencing strategy (Illumina/ Qiagen sequencing kits) followed by consensus generation using mapping-based methods. Lineage identification is performed using the Pangolin tool (v 4.0.6 PLEARN v1.6) and classification system and mutation annotations via Nextclade tools (v 1.11.0).

Statistical analysis and visualization of sequences in their epidemiological context was performed using the CoviSurve dashboard from Strand Life Sciences Pvt Ltd as a tool for genomic surveillance to enable the interpretation of sequencing data in conjunction with associated
epidemiological data. Strand COVID Mutation miner software was used to track the specific COVID-19 mutations of interest (https://covidmm.mystrand.org/).

Reports were generated on a regular basis and summary data and key highlights shared with local authorities, the Genomic Surveillance Committee for SARS-CoV-2 and the COVID task force in the state. The entire workflow is outlined in the figure (1).

Figure 1: The work flow of SARS-CoV-2 genomic surveillance conducted at Strand Life Sciences Pvt Limited

Results:

A total of 12461 RT PCR positive COVID-19 samples were sequenced at Strand Life Sciences during the time period between July 2021 and May 2022, of which 12165 sequences with lineage assignments were obtained (Figure 2). Of these 12165 samples, 5405 (44.4%) were Delta and sub-lineages and 6723 (55.2%) were Omicron and sub-lineages. The 57.5% (3065) of the Delta sequences were from males and 42.5% (2265) of them were from females (75 sequences did not have gender information). Similarly, 53.7% (3438) of the Omicron lineages detected in men and 46.3% (2959) detected in females (326 sequences did not have gender information) (Figure 3). All these samples were collected from the various wards across Bengaluru city as shown in the map (Figure 4).
Most of the Delta (55.2%) and Omicron (62.5%) lineages were detected from the age group between 19 and 45. The children below 12 years were more affected by Delta (11.5%) when compared to Omicron (5.7%) (Figure 5).

**Figure 2:** Month-wise distribution of SARS-CoV-2 Variants of Concern (VoCs) and Variants of Interest (VoIs) in Bengaluru from July 2021 to May 2022. X-axis represents the month and Y-axis represents the total number of sequences and the proportion of variants. The Delta to Omicron takeover is clearly visible in the figure.

**Figure 3:** Gender wise distribution of Delta and Omicron lineages detected in this study
Figure 4: Bengaluru city map highlighting regions (in dark green) where samples were collected.

Figure 5: Age-wise distribution (for each variant) of Delta (A) and Omicron (B) lineages.
We have identified 75 Delta sub-lineages and 26 Omicron sub-lineages in Bengaluru during this study period (Supplementary Figure 1). Of these Delta sub-lineages, over 70% were B.1.617.2. Whereas predominant Omicron sub-lineages identified were BA.1 and BA.2. The origin and spread of Omicron sub-lineages are given in Table-1.

![Figure 6: Relative frequencies of Delta (A) and Omicron (B) sub-lineages detected in Bengaluru during the study period.](image)

The first Omicron case in Bengaluru was detected in November 2021 and Omicron grew from <5% in December 2021 to 100% by the end of January 2022 (Figure 2). Of Omicron, BA.1, BA.2, BA.3 and BA.5 and their sub-lineages have been detected in Bengaluru till May 2022.
Sub-lineages of Omicron and their spread: A total of 6723 Omicron cases were detected in the samples sequenced from January to May 2022. We found a total of 26 Omicron sub-lineages during this time period. BA.1 has 13 sub-lineages and BA.2 has 10 sub-lineages. Of these 6723 Omicron cases, 5985 (89.0%) are BA.2 and sub-lineages, 628 (9.3%) are BA.1 and sub-lineages, 115 (1.7%) are BA.3. We detected two BA.5 cases in April 2022. The top four BA.2 sub-lineages found in Bengaluru during this time period were BA.2.10, BA.2, and BA.2.10.1and BA.2.12. The top three BA.1 sub-lineages found in Bengaluru during this time period were BA.1.1, BA.1, BA.1.1.7 and BA.1.18 (Table 1).

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Sub-lineages</th>
<th>Nov-21</th>
<th>Dec-21</th>
<th>Jan-22</th>
<th>Feb-22</th>
<th>Mar-22</th>
<th>Apr-22</th>
<th>May-22</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>BA.1.1.529</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>BA.1</td>
<td>1</td>
<td>20</td>
<td>156</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>BA.1.1</td>
<td>0</td>
<td>7</td>
<td>240</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>BA.1.1.1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>BA.1.1.14</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>BA.1.1.16</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>BA.1.1.7</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>BA.1.13.1</td>
<td>1</td>
<td>5</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>BA.1.1.4</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>BA.1.15</td>
<td>0</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>BA.1.1.7</td>
<td>0</td>
<td>1</td>
<td>100</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>BA.1.1.7.2</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>BA.1.1.8</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>BA.1.1.9</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>BA.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>BA.2</td>
<td>0</td>
<td>10</td>
<td>1071</td>
<td>75</td>
<td>87</td>
<td>62</td>
<td>73</td>
</tr>
<tr>
<td>2</td>
<td>BA.2.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>BA.2.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>BA.2.10</td>
<td>0</td>
<td>6</td>
<td>4187</td>
<td>211</td>
<td>35</td>
<td>9</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>BA.2.10.1</td>
<td>0</td>
<td>0</td>
<td>77</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>BA.2.11</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>BA.2.12</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>BA.2.12.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>BA.2.13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>BA.2.31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>BA.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>BA.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Details of Omicron sub-lineages detected from November 2021 to May 2022
Discussion:

Before the COVID-19 pandemic, genomic surveillance was reserved mainly for conducting small studies of antibiotic-resistant bacteria, investigating outbreaks and monitoring influenza strains (15–21). As new coronavirus VoC continue to emerge, genomic surveillance has an important role to play in helping bring the pandemic under control (22,23). This genomic surveillance programme was initiated to build an end-to-end solution that addresses challenges of SARS-CoV-2 genomic surveillance and integrates logistics, laboratory work, informatics/analytics and reporting in a timely and seamless manner. This was expected to lead to early detection and identification of both local as well as imported VoC, provide reliable baseline data for the assessment of local VoC and timely analysis of outbreaks that can inform public health policy.

Continuous genomic surveillance is always coupled with numerous challenges in all parts of the world (24). The main challenges in effective genomic surveillance in Bengaluru were extreme variations in COVID-19 incidence which affected the sample collection and cost-effective batching, sequencing a small portion of the growing number of cases made it difficult to pick VoC early and the samples which are not uniformly spread in space and time made interpretation difficult. The unavailability of metadata made interpretation of results with relevant clinical and epidemiological data (vaccination, hospitalization status, age etc.) difficult.

The impacts of our surveillance programme on the genomic surveillance efforts are discussed below.

Sequencing, timely reporting and data sharing

Our sequencing efforts successfully sequenced a total of 12461 samples from the city of Bengaluru between July, 2021 and May 2022 with continuous engagement with local authorities at BBMP and the Genomic Surveillance Committee for Karnataka. The timely reporting of the sequencing allowed the health officials to begin preparedness for the pandemic response very early. This sequencing frequency allowed us to capture the entire delta wave at high resolution – giving us the ability to look at expansion of different sub-lineages of Delta and introduction and spread of Omicron (Figure 2-3) in the city of Bengaluru.

As of May 31, 2022, a total of 30,066 (including 11500 from Strand Life Sciences) sequences have been deposited into openly available database Global Initiative on Sharing All Influenza Data.
(GISAID) that facilitated rapid and open sharing of SARS-CoV-2 genome sequences, from the Indian State of Karnataka (9). This data suggests that the 38.2% of the sequences has so far contributed to GISAID by genomic sequencing efforts of Strand Life Sciences in collaboration with local government bodies and with support by IN COVID Support, FZE LLC. Along with detection, this provided an added advantage of understanding the genetic epidemiology of the outbreak over a period of time. Our data demonstrated that the Delta and Omicron affected males more compared to females (Figure 2), however a large-scale study conducted by Wang et al in the USA states that the females are more infected by these two variants (25). We have detected over 100 lineages of Delta and Omicron lineages circulating in Bengaluru during the study period and reported the changing landscape of SARS-CoV-2 variants in the city (Figure-6).

**Introduction and spread of Omicron and its sub-lineages in Bengaluru**

The variant Omicron (lineage B.1.1.529 and its sub-lineages BA.1 to BA.5) was designated as a Variant of Concern by the WHO on 26th Nov 2021, very soon after reports of the variant and associated surge in cases was reported from South Africa (26). This variant has a large number of mutations in the spike protein, conferring properties of immune escape (27,28).

Strand Life Sciences played a key role in detection of Omicron in returning international travelers to Karnataka. Continuous monitoring of SARS-CoV-2 in Bengaluru city allowed early detection of the spread of Omicron in the community. Sequencing of returning international travelers from South Africa was initiated in late November, even as the world was beginning to appreciate the high transmission of the Omicron variant. The first sequence of Omicron from Karnataka was reported in collaboration with the National Centre for Biological Sciences on Dec 02, 2021 as submission to the GISAID database.

Of Omicron and its sub-lineages, BA.1, BA.2 and BA.3 and BA.5 and their sub-lineages have been detected in Karnataka (Table 1). It has also been found to displace and transmit faster than Delta, the then circulating VoC in many parts of the world. Initially, a sub-variant BA.1 was the most common circulating version of Omicron. But a genetically distinct sub-variant BA.2 and its sub-lineages BA.2.10 and BA.2.12 now account for majority of new cases and have become the dominant coronavirus variant around the world (29). Our data suggest that BA.2.10 and BA.2 were responsible for the COVID-19 wave in January and February 2022.
Timely information that AY.4.2 was not a variant of concern in Bengaluru

AY.4.2 is a sub-lineage of the Variant of Concern Delta (3). It was highlighted as a potential Variant of Interest in the UK in late October 2021 (30). At this time, there were two reported sequences of this lineage from Karnataka in GISAID collected in July 2021. Due to the continuous sequencing of SAR-CoV-2 performed between July -October 2021 by Strand Life Sciences, we were able to show that in Bengaluru there had been no expansion of this sub-lineage. Only 1/948 sequences in October 2021 from Bengaluru belonged to this lineage. This information was shared with the relevant authorities in a timely manner.

**Key Contributions of this Initiative**

In order to address the main challenge of data integration, Strand Life Sciences has developed a dashboard for real-time integration of SARS-CoV-2 sequencing data with the clinical and epidemiological information of the cases – CoviSurve. CoviSurve eases the visualization and interpretation of sequencing data. When a variant is detected, CoviSurve can help answer the questions on who/how many had that variant and where and when it was detected and also help analyse the mutations and locations where a novel mutation is dominant – this can serve as a tool for early identification of novel variants. The inbuilt reporting into CoviSurve helps generate relevant and real-time reports, including analysis of variants in sub-groups such as vaccinated individuals and children.

**Conclusion**

It was essential that the surveillance efforts be kept up, and ideally expanded, so that variants, in particular those that can cause severe disease or evade immunity that is conferred by natural infections or existing vaccines, can be detected early on. This will allow for the most effective prevention and treatment options. It was imperative that this be done in order to stop the further spread of the disease. As part of the response to the pandemic, continuous monitoring of SARS-CoV-2 variants at a local scale, as well as timely analysis and reporting of potential variants of concern, must continue to be prioritised. In order to make a significant contribution, we have developed tools and scalable systems for genomic surveillance. These systems and tools have the potential to be utilised in the investigation of other types of pathogens in the future.
Data Availability Statement

All the sequences are available at GISAID open database.

Author Contributions


Funding

Support from ACT COVID Response Collective, IN COVID Support FZE LLC, as well as Institute of Virology Research & Innovation, Department of Biotechnology (INSACOG), India, is gratefully acknowledged.

Competing Interests

Authors who are employees of Strand Life Sciences Private Limited may hold shares in Strand Life Sciences Private Limited. The remaining authors declare no competing interests.

Acknowledgments

The help and support of COVID Task force, Bruhat Bengaluru Mahanagara Palike (BBMP), Genomic Surveillance Committee for Karnataka, Health and Family Welfare Services, Karnataka, and numerous private COVID-19 testing laboratories of Bengaluru are acknowledged. The Karnataka state leadership starting with the Chief Minister as well the Minister for Higher Education and Biotechnology were supportive of this new paradigm of public-private partnership in pandemic genomics surveillance and preparedness.
References:


