SARS-CoV-2 Genome Characterization by Next Generation Sequencing in Bamrasnaradura Infectious Diseases Institute: Implication for vaccine tools

Sumonmal Uttayamakul[1], Paninee Mongkolsuk[1], Sarinee Reawrang[1], Pongpun Sawatwong[2], Thidathip Wongsurawat[3], Piroon Jenjaroenpun[3], Beth A Skaggs[2], Navin Horthongkham[4], Weerawat Manosuthi[1]

[1]Bamrasnaradura Infectious Diseases Institute, Department of Disease Control, Ministry of Public Health
[2]Division of Global Health Protection, Thailand Ministry of Public Health-US. CDC, Nonthaburi
[3]Division of Bioinformatics and Data Management for Research, Department of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[4]Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

Background

The genetic diversity of SARS-CoV-2 has been reported. SARS-CoV-2 strain investigation is important for anti-viral drug design and to assess vaccine efficacy. Nanopore sequencing Technology is included in the WHO guideline (2021) as an accepted method to obtain SARS-CoV-2 genome data

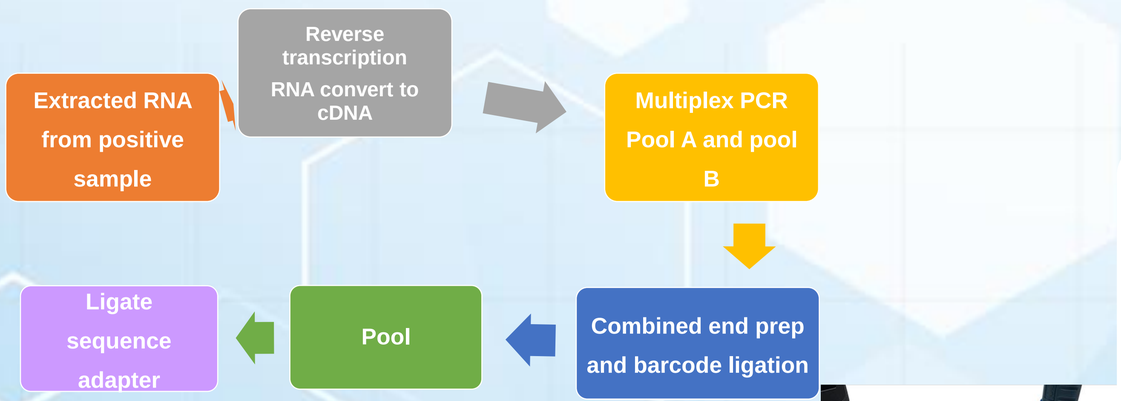
Objective: To investigate the SARS-COV-2 variant strains by Nanopore Next Generation Sequencing

Materials/Methods:

A total of 83 COVID-19 cases from the first (February to April 2020, n=51) and second (January to April 2021, n=32) waves of the COVID-19 outbreak were investigated at BIDI. Viral RNA from nasopharyngeal/throat swabs was extracted by MagNaPure Compact and full-length SARS-CoV-2 genome sequences were determined using Oxford Nanopore Technologies (UK) platform and ARTIC protocol.

Library preparation & whole-genome sequencing

Extracted RNA was reverse transcribed as described in the PCR tiling of COVID-19 protocol (vPTC_9096_v109_revF_06Feb2020) published by the ARTIC Network primer V3 (https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye). The libraries were prepared using a ligation-based sequencing kit (SQK-LSK109 kit; ONT) and barcoding using NBD104 reagent Kit (ONT), loaded onto a MinION flow cell (106 version), and sequenced with the MinION Mk1B device (ONT).



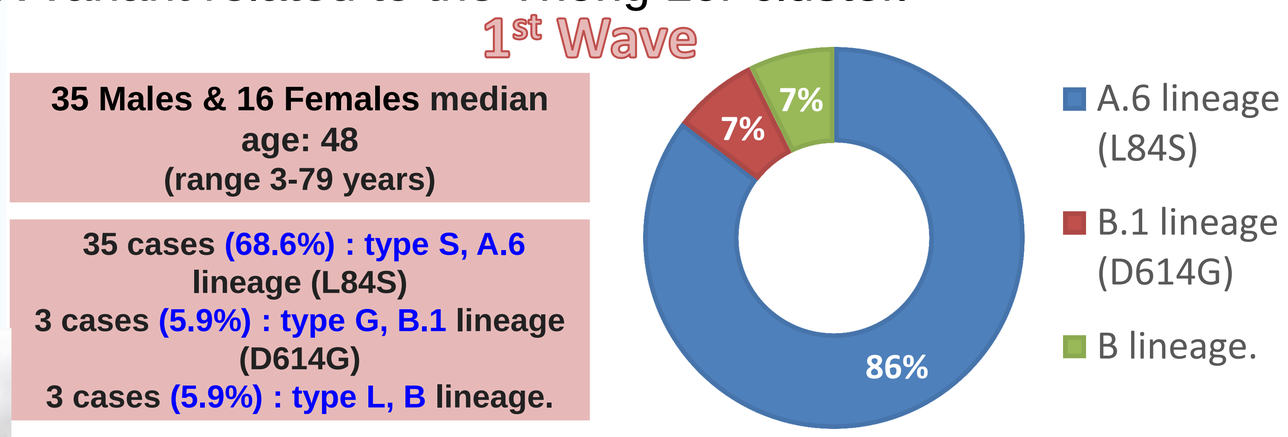
Base-calling and demultiplexing by Guppy software

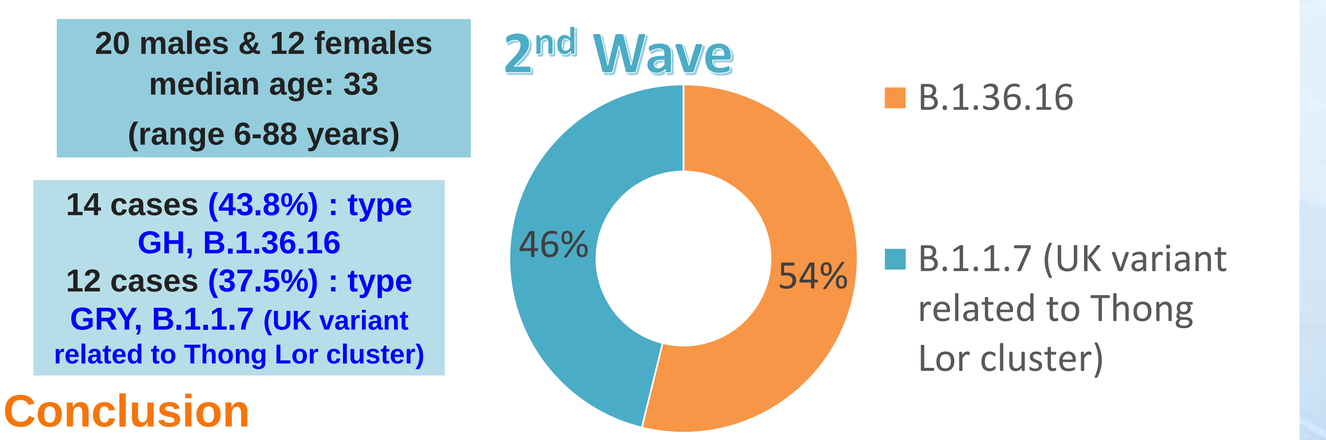
Reference assembly by using ARCTIC bioinformatics pipeline

Bioinformatics

Base calling of the resulting FAST5 files was performed in real time using Guppy (v4.5.2) (1) (ONT) with high accuracy mode (dna_r9.4.1_450bps_hac). RAMPART software (v1.0.5) from the ARTIC Network (https://github.com/artic-network/rampart) was used to real-time monitor sequencing. (https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html) was used, followed by a reference assembly using the MinION script with Medaka polishing against the sequence of the Wuhan-Hu-1 isolate (GenBank accession number MN908947.3).

Results. Fifty-one COVID-19 cases included 35 males and 16 females with median age of 48 (range 3-79 years) were studied in the first wave. The majority of SARS-CoV-2 genomes analyzed (35 cases, 68.6%) were classified as type S, A.6 lineage (with L84S) followed by three cases (5.9%) classified as type G, B.1 lineage (with D614G) and three cases (5.9%) of type L, B lineage. During the second wave, 32 COVID 19 cases included 20 males and 12 females (median age of 33, range 6-88 years); 14 cases (43.8%) were classified as type GH, B.1.36.16 lineage followed by 12 cases (37.5%) classified as type GRY, B.1.1.7 lineage which is the UK variant related to the Thong Lor cluster.





Conclusion

The most prevalent SARS-CoV-2 strain in the first wave was S type, A.6 lineage which related to a boxing stadium cluster. The variant in the most recent outbreak was GH type, B.1.36.16 lineage. This information could inform COVID-19 vaccine efficacy monitoring.

Contact: sumonmal@health.moph.go.th