Molecular characterisation and tracking of the severe acute respiratory syndrome coronavirus 2, Thailand, 2020–2022

Jiratchaya Puenpa
  Chulalongkorn University Faculty of Medicine

Patthaya Rattanakomol
  Chulalongkorn University Faculty of Medicine

Nutsada Saengdao
  Mahidol University Faculty of Medicine Siriraj Hospital

Jira Chansaenroj
  Chulalongkorn University Faculty of Medicine

Ritthideach Yorsaeng
  Chulalongkorn University Faculty of Medicine

Kamol Suwannakarn
  Mahidol University Faculty of Medicine Siriraj Hospital

Somchai Thanasitthichai
  Ministry of Public Health, Thailand

Sompong Vongpunsawad
  Chulalongkorn University Faculty of Medicine

Yong Poovorawan (✉ yong.p@chula.ac.th)
  Chulalongkorn University   https://orcid.org/0000-0002-2337-6807

Research Article

Abstract

The global COVID-19 pandemic caused by a novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) was first detected in China in December 2019. To date, there have been approximately 3.4 million reported cases and over 24,000 deaths in Thailand. This study investigated the molecular characterisation and the evolution of the SARS-CoV-2 identified during 2020–2022 in Thailand. Two hundred and sixty-eight SARS-CoV-2 strains, collected mostly in Bangkok from COVID-19 patients, were characterised by partial genome sequencing. Moreover, 5,627 positive SARS-CoV-2 samples were identified as viral variants [B.1.1.7 (Alpha), B.1.617.2 (Delta), B.1.1.529 (Omicron/BA.1) and B.1.1.529 (Omicron/BA.2)] by multiplex real-time reverse-transcription polymerase chain reaction (RT-PCR) assays. The results revealed that B.1.36.16 caused the predominant outbreak in the second wave (December 2020–January 2021), B.1.1.7 (Alpha) in the third wave (April–June 2021), B.1.617.2 (Delta) in the fourth wave (July–December 2021), and B.1.1.529 (Omicron) in the fifth wave (January–March 2022). The evolutionary rate of the viral genome was 2.60×10^{-3} (95% highest posterior density [HPD], 1.72×10^{-3} to 3.62×10^{-3}) nucleotide substitutions per site per year. Continued molecular surveillance of SARS-CoV-2 is crucial for monitoring emerging variants with the potential to cause new COVID-19 outbreaks.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection causes the COVID-19 pandemic as declared by World Health Organization (WHO) on 11 March 2020 and continues to impact public health and world economies [1]. As of 26 December 2021, there were over 278 million confirmed COVID-19 cases, with over 5 million fatalities globally [2]. SARS-CoV-2 is an enveloped virus with a positive-sense, single-stranded RNA genome of approximately 27–32 kb in size [3]. The genome encodes four structural proteins [spike (S), envelope (E), membrane (M), and nucleocapsid (N)] and 16 non-structural proteins (nsp1-nsp16) involved in viral function and replication. The massive circulation of SARS-CoV-2 worldwide and inequitable global vaccine distribution led to the evolutionary pressure on the virus and the emergence of new variants [4, 5]. An abundance of SARS-CoV-2 genome sequences has been generated rapidly and deposited in the global archive, namely the Global Initiative on Sharing All Influenza (GISAID) database. As of 2 March 2021, significant mutations on the viral genome led to the viral classification into nine clades (S, L, V, G, GH, GK, GR, GV, and GRY) [6]. The epidemiologically relevant phylogenetic cluster of SARS-CoV-2 is further defined as a lineage by the Phylogenetic Assignment of Named Global Outbreak Lineages (PANGOLIN) tool [7]. Based on the enhanced transmissibility, increase in virulence, and decrease in the natural infection- and vaccine-mediated neutralizing antibody attributed to significant amino acid substitutions, several SARS-CoV-2 variants are classified as Variants of Concern (VOC) [8]. To assist public communication and avoid stigmatisation, the Greek alphabet, i.e. Alpha, Beta, Gamma, Delta, Omicron, is used to designate SARS-CoV-2 variants [8].

The first confirmed COVID-19 case was reported on 12 January 2020 in a traveller from China [9]. The number of the infected individuals rapidly surged during March–May 2020 because of transmission linked to boxing events and entertainment venues in Bangkok, which then spread throughout Thailand

and was considered the first COVID-19 wave [10]. To effectively contain the viral spread, public health and social measures including wearing masks, physical distancing, movement restriction, workplace and school closures, and city lockdowns were implemented. The third wave began in early April 2021 with the upsurge of COVID-19 cases linked to the entertainment venue at Thonglor district in Bangkok [11]. This severe and deadly wave was driven by the emergence of a more transmissible B.1.1.7 (Alpha) SARS-CoV-2 variant of concern (VOC), leading to rising hospitalisations and overwhelming healthcare facilities [12]. With the exception of the city lockdown, all protective measures implemented by the government were still in effect at this times. Moreover, field hospitals were set up to handle patient isolation. Amid the third wave, the mass vaccination campaign was rolled out on 7 June 2021 to slow down the transmission. However, the supply of vaccines was limited, and only two brands of vaccines, CorovaVac and Vaxzevria, were available at that time. As of 1 July 2021, there were 52,052 confirmed COVID-19 cases with 1,971 patients classified as having severe illness, of which 566 patients required ventilator support [13]. As of 8 July 2021, COVID-19 cases were reported in all 77 provinces of Thailand. The Centre for COVID-19 Situation Administration (CCSA) declared the emergence of the fourth wave of COVID-19 pandemic in Thailand from the highly contagious Delta variant, whose transmissibility is faster than that of the previous SARS-CoV-2 variant [13]. By December 2021, Thailand experienced its fifth COVID-19 wave, with approximately 3.4 million confirmed COVID-19 cases and over 24,000 deaths as of the end of March 2022 [14].

In our previous molecular epidemiological investigation of SARS-CoV-2 in Thailand during the first wave of the outbreak in 2020, 40 nasopharyngeal and/or throat swab specimens from 40 samples were found infected with SARS-CoV-2 types L, GH, GR, O, and S [15]. Here, we aimed to monitor and track emerging new variants of SARS-CoV2 circulating in Thailand between March 2020 and March 2022. The present study provides epidemiological patterns of SARS-CoV-2 in Thailand that could have implications for a more effective disease surveillance and public health preparedness.

Materials And Methods

Ethics statement

The study protocol was approved by the Institutional Review Board of the Faculty of Medicine of Chulalongkorn University (IRB number 178/64). Waiver for written informed consent was granted due to the samples' anonymity by the institutional review board of the Ethics Committee for human research. All methods were performed in accordance with the relevant guidelines and regulations.

Sample collection and RNA extraction

All nasopharyngeal swab samples were collected as part of outbreak investigations during the period of the first through fifth waves and state quarantine (SQ) from March 2020 to March 2022 (Fig. 1). During the study, 5,627 selected nasopharyngeal swab samples submitted to the collaborating hospitals and the Institute of Urban Disease Control and Prevention (IUDC) routinely tested positive for SARS-CoV-2 by multiplex real-time reverse transcription polymerase chain reaction (RT-PCR) assays described earlier [15].

Total nucleic acid was extracted from 200 μL of supernatant using magLEAD 12gC instrument (Precision System Science, Chiba, Japan) according to the manufacturer's instructions, and prior to analysis in our laboratory.

Genome sequencing

Briefly, the partial sequences of SARS-CoV-2 were amplified from the S, RNA-dependent RNA polymerase, and N gene amplification stages performed using the SuperScript III Platinum One-Step RT-PCR System (Invitrogen, Carlsbad, CA, USA) and 10 sets of oligonucleotide primers (Table S1). The following amplification conditions were used: one cycle of 3 min at 94°C, 40 cycles consisting of 30 s at 94°C, 30 s at 53°C, and 90 s at 68°C, and a final cycle of 7 min at 68°C. Amplicon purification with the HiYield Gel/PCR DNA Fragment Extraction kit (RBC Bioscience Co, Taipei, Taiwan) was followed by Sanger sequencing. The nucleotide sequences of SARS-CoV-2 were deposited in GenBank under accession numbers OK083891-OK084640, OM984745-OM984850, and OM996047-OM996083 (Table S2).

Multiplex real-time RT-PCR assay

The primers and probes specific to B.1.1.7 (Alpha), B.1.617.2 (Delta), B.1.1.529 (Omicron/BA.1) and B.1.1.529 (Omicron/BA.2) were designed to target the S gene. The sequences of primers and probes were selected from conserved regions of sequences available in GISAID database (http://www.gisaid.org/). To identify the lineage B.1.1.7 (Alpha), lineage B.1.617.2 (Delta) and B.1.1.529 (Omicron BA.1 and BA.2), one-step multiplex real-time RT-PCR was carried out in a total volume of 20 μl containing RNA template, the SensiFAST Probe One-Step RT-PCR system (Bioline, London, UK) including RNase inhibitor, Mg^{2+}, and two pairs of primers and probes. The multiplex sets of primers and *Taq*-Man probes are described in Table S3. One-step multiplex real-time RT-PCR was performed on the LightCycler 480 real-time PCR system (Roche, Mannheim, Germany). The thermocycler conditions included a reverse transcription step at 42°C for 30 min, and an activation hot start DNA *Taq* polymerase at 95°C for 10 min, followed by amplification that was performed during 45 cycles of denaturation at 95°C for 15 s and annealing/extension at 60°C for 30 s. Multiple fluorescent signals were obtained once per cycle upon completion of the extension step. Data acquisition and analysis of the real-time PCR assay were performed using LightCycler 480 SW1.5 software (Roche).

Phylogenetic analysis

The spike sequence dataset was constructed with the BioEdit v7.2.6 software [16] and aligned using CLUSTAL W on the European Bioinformatics Institute (EBI) webserver [17]. The diversity of SARS-CoV-2 lineages was analysed with the maximum-likelihood phylogenetic method available in the MEGA program (v7) [18]. The Tamura 3-parameter with gamma distribution was selected as the substitution model in the analyses. The bootstrap method determined the statistical consistency of tree nodes (500 random samplings).

A time-scaled phylogenetic tree for complete genome sequences was reconstructed with the BEAST version 1.10.4 program [19]. The Bayesian phylogenetic analyses used an uncorrelated log-normal prior distribution of nucleotide substitution rates among lineages. The general time-reversible (GTR) model was selected as the nucleotide substitution model. Bayesian Markov chain Monte Carlo analysis was run for 100 million steps, 10% of which were removed as burn-in and sampled every 1,000 steps from the posterior distribution. The Tracer version 1.7.1 tool (http://tree.bio.ed.ac.uk/software/tracer/) was used to assess for the convergence of all parameters (operator effective sample size of > 200). The maximum clade credibility (MCC) tree was calculated with the TreeAnnotator v1.10.4 tool (http://beast.bio.ed.ac.uk/treeannotator).

Results

Distribution of SARS-CoV-2 outbreaks in Thailand

From March 2020 to May 2021, 113 confirmed cases (1st wave; N = 8, 2nd wave; N = 40 and state quarantine; N = 65) were successfully genotyped by partial SARS-CoV-2 genome sequencing. In this study, the first wave of the outbreak (March–May 2020) in Thailand was characterised by two different lineages, A and B.1 (Figs. 1 and S1). During the period of SQ from May 2020 through May 2021, we received 65 clinical specimens obtained from travellers and Thais who presented with/without symptoms of COVID-19 and were admitted to a hospital/hotel in Bangkok and Chon Buri province. Among these, lineage B.1.1 was the most frequently detected genotype and accounted for 29.2% of the strains (19/65), followed by 48% (31/65) for lineage B.1. Of the remaining strains, 11 were classified as lineage B.1.1.7 (Alpha) and another three of lineage B.1.177 (Figure S1 and Table S2). All of these strains from SQ were imported from the Americas (12.3%), Asia (41.5%), Europe (23.1%), and unknown (23.1%). The results showed that lineage B.1.1.7 (Alpha) was imported from the United States, France, Slovenia, the United Kingdom (UK), and Northern Ireland. Lineage B.1.177 was imported from the UK and the United Arab Emirates (UAE). Lineage B.1.1 was predominantly imported from Asian countries (Qatar, India, Philippines, Japan, and Bahrain), the UK, and Italy. Lineage B.1 was imported from Asian and European countries. During the second wave of the outbreak (October 2020–March 2021), lineage B.1 became the predominant virus.

Multiplex real-time RT-PCR assays to differentiate variants of SARS-CoV-2

In Thailand, predominant variants were detected in different epidemic waves (Fig. 2). From March 2020 to 14 March 2022, 5,627 samples were identified as B.1.1.7 (Alpha), B.1.617.2 (Delta), and B.1.1.529 (Omicron BA.1 and BA.2) from SARS-CoV2 positive samples by using multiplex real-time RT-PCR. The results showed that clade B.1.1.7 (Alpha) was the most frequent variant in the third epidemic wave (1,510/5,627: 26.8%). B.1.617.2 (Delta) (2,382/5,627: 42.3%) was the predominant strain responsible for SARS-CoV-2 infection in the fourth epidemic wave. B.1.1.529 (omicron BA.1) was first detected in Thailand in mid-December 2021 (1,375/5,627: 24.4%), which caused a nationwide epidemic wave until

the end of February 2022. Since then, clade B.1.1.529 (omicron BA.2) emerged in Thailand at the end of January 2022 (360/5,627: 6.4%) and became the major variant in early March 2022.

Phylogenetic relationship

The complete spike nucleotide sequences of the 155 SARS-CoV-2 strains were investigated to evaluate the genetic relationship among Thai SARS-CoV-2 strains. A total of 155 positive strains consisting of clade B.1.1.7 (Alpha) (*N* = 20), B.1.617.2 (Delta) (*N* = 33), B.1.1.529 (omicron BA.1) (*N* = 48), and B.1.1.529 (omicron BA.2) (*N* = 54), were characterised by amplification of the partial S gene, which harbours the major antigenic sites in SARS-CoVs. Sequences determined in isolates from 113 samples during the first and second epidemic waves and SQ were also included. The phylogenetic relationships defined with the S sequences showed that clade B.1.1.529 (omicron BA.1 and BA.2) cluster together and were differentiated from other clades with bootstrap values > 95%.

The S gene analysis showed that the B.1.1.7 (Alpha) differed from the Wuhan-Hu1 strain by 7 amino acid substitutions: N501Y, A570D, D614G, P681H, T716I, S982A and D1118H. The B.1.177 variants contained L16F, D215H, A222V, N370S and D614G mutations. Several mutations were identified in the sequences of the B.1.617.2 (Delta) variants, namely, T17R, T93I, G140D, L452R, T478K, D614G, P681R, and D950N. Forty-two and 31 amino acid substitutions were detected in the B.1.1.529 (omicron BA.1) and the B.1.1.529 (omicron BA.2), respectively. In addition, one strain (Thailand_CU8056) carried an I1221T substitution in the S protein that was clustered in the B.1.1.529 (omicron BA.2) clade (Fig. 3).

The evolutionary history of the structural region of SARS-CoV-2 was investigated by performing Bayesian analysis with a SARS-CoV-2 S glycoprotein sequence data set. The mean evolution rate was 2.60×10^{-3} (95% highest posterior density [HPD], 1.72×10^{-4} to 3.62×10^{-4}) substitutions per site year (Fig. 4). The most recent common ancestor of all SARS-CoV-2 clades was dated in September 2019.

Discussion

Since early 2020, SARS-CoV-2 has been circulating in Thailand. The SARS-CoV-2 fourth wave outbreak is the largest known in Thailand, with over 800,000 recorded cases at the end of December 2021 [14]. This study characterised the partial genome sequences of 268 Thai strains, performed phylogenetic analyses, and analysed the molecular evolution to investigate their relationships with previously described viruses. The first outbreak began in Thailand in 2020, with the first imported case detected as having occurred in late January 2020, and spread to several other provinces, with Bangkok most severely affected [9]. This outbreak was attributed to the lineage A.6 (S) variants and was responsible for 67.5% of all SARS-CoV-2 [15]. This study showed that clade GH rapidly became the predominant variant throughout Thailand during the second wave of the COVID-19 epidemic (late December 2020). In November 2020, the B.1.1.7 (Alpha) variant emerged for the first time in the United Kingdom and caused higher mortality rate [20]. The B.1.1.7 (Alpha) variant subsequently spread through Europe, the United States and Asia over the next two months. Between April and June 2021, a third wave of the B.1.1.7 (Alpha) variant peaked in Thailand.

The emergence of the Delta variant in India in October 2020 caused large-scale outbreaks in that sub-continent [21]. Subsequent emergence of the B.1.617.2 (Delta) variant became the dominant variant worldwide since February 2021. In Thailand, the fourth epidemic peaks were observed during August–September 2021, mainly caused by the B.1.617.2 (Delta) variant. The B.1.1.529 (Omicron) variant was first identified in South Africa in November 2021 [22]. After its emergence, the B.1.1.529 (Omicron) variant has replaced the B.1.617.2 (Delta) variant and has circulated as a dominant variant in several countries since December 2021. In the mid December 2021, the B.1.1.529 (Omicron) variant was detected in Thailand. The highest recorded COVID-19 positive cases mainly caused by this variant occurred in March 2022 when COVID-19 case numbers surged.

A recent study in Malaysia showed that the lineage B.6 (O)-associated groups and B.1.524 (G) were the predominantly detected variants throughout the country during the second and third epidemic waves, respectively [23]. In Vietnam, there was a reported increase of two clusters of SARS-CoV-2 as the waves of virus infection progressed from its emergence during July 2020–February 2021, with a novel, shared mutation in nsp9 [24].

Several mutations in the S protein identified in SARS-CoV-2 isolates are variant of concern (VOCs), which could possibly enhance transmissibility and viral infectivity [25–26]. As is already known, the D614G mutation in the S protein increases the ability of the virus to replicate in the upper respiratory tract, causes a possible conformational change in the S1 subunit, and enhances furin cleavage efficiency at the S1/S2 site [27–28]. The D614G mutation on the S protein was first detected outside of China in the small outbreak in Germany in January 2020 [29]. Our analysis showed that almost 90% of the Thai variants contained the D614G mutation. Spike mutations N501Y and K417N are first recorded in the B.1.1.7 (Alpha) and B.1.351 (Beta) variants, respectively; the presence of these substitutions in the receptor-binding domain (RBD) on S enables increased binding affinity of the virus to the ACE2 receptor [30–32]. As recently reported, the E484K substitution was first detected in the B.1.351 (Beta) variants and was associated with antibody neutralisation escape by directly reducing antibody binding affinity [33–34]. The mutational analysis of the B.1.617.2 (Delta) variant showed that the variant contained T to G transversion at nucleotide position 22,917, corresponding to the L452R mutation in the S protein, also found in the B.1.617.1 (Kappa) and B.1.427 and B.1.429 (Epsilon). The L452R substitution is located on the RBD of SARS-CoV-2, which is involved in the reduction of antibody neutralising activity [35].

Our study showed that the mean evolutionary rate early in the epidemics was 2.60×10^{-3} nucleotide substitutions per site per year. This rate was approximately four times as high as that averaged from genomic sequences of SARS-CoV-2 strains from Pakistan outbreaks, at around 5.68×10^{-4} substitutions per site per year [36]. Our estimate was similar to those previously published report of SARS-CoV-2 (0.99–1.8×10^{-3} substitutions per site per year) [37–40].

A limitation of this study is that partial genome sequencing data were not available for all outbreak samples since S region typing is not yet routinely performed by our laboratory. We successfully obtained SARS-CoV-2 typing data by using a multiplex real-time RT-PCR assay between the third and fifth epidemic

waves occurring primarily in the last year of the study. The present study highlights the importance of molecular typing for a complete understanding of the diversity and circulation of SARS-CoV-2.

In summary, a SARS-CoV-2 outbreak has been ongoing in Thailand for more than 2 years, with a total of five epidemic waves. Clade B.1.36.16 (GH) predominated in the second epidemic wave (2021), clade B.1.1.7 (Alpha) in the third wave (2021), clade B.1.617.2 (Delta) in the fourth wave (2021), and clade B.1.1.529 (Omicron) in the fifth wave (2022), indicating a new epidemic wave occurring due to emerging strains. Continued molecular surveillance of SARS-CoV-2 is crucial for monitoring emerging variants to prevent possible new COVID-19 outbreaks.

Declarations

Acknowledgements

Author contributions

J.P. and P.R. analysed data and drafted the manuscript. P.R., N.S. and J.C. performed the laboratory experiments and collected data. R.Y. and N.S. collected specimens. K.S., S.T., S.V. and Y.P. supervised the study and finalized the manuscript. All authors reviewed the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Data availability

The datasets generated during the current study are available in the GenBank (https://www.ncbi.nlm.gov/genome) under accession numbers OK083891-OK084640, OM984745-OM984850, and OM996047-OM996083.

References

1. WHO (2020) Coronavirus disease 2019 (COVID-19). Situation Report-51. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf. Accessed 22 March 2022

2. WHO (2021) Weekly epidemiological update on COVID-19–28 December 2021. https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---28-december-2021. Accessed 5 February 2022

3. Wang MY, Zhao R, Gao LJ, Gao XF, Wang DP, Cao JM (2020) SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. Front Cell Infect Microbiol 10:587269. https://doi:10.3389/fcimb.2020.587269

4. Ahmad L (2021) Implication of SARS-CoV-2 Immune Escape Spike Variants on Secondary and Vaccine Breakthrough Infections. Front Immunol 12:742167. https://doi:10.3389/fimmu.2021.742167

5. Bui NN, Lin YT, Huang SH, Lin CW (2022) Haplotype distribution of SARS-CoV-2 variants in low and high vaccination rate countries during ongoing global COVID-19 pandemic in early 2021. Infect Genet Evol 97:105164. https://doi:10.1016/j.meegid.2021.105164

6. GISAID (2021) Clade and lineage nomenclature aids in genomic epidemiology studies of active hCoV-19 viruses. https://www.gisaid.org/references/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/. Accessed 5February2022

7. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, du Plessis L, Pybus OG (2020) A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 5:1403–1407. https://doi:10.1038/s41564-020-0770-5

8. WHO (2021) Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/. Accessed 15January2022

9. WHO (2020) Coronavirus disease 2019 (COVID-19). Situation Report-1. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf. Accessed 5January2022

10. WHO (2020) Coronavirus disease 2019 (COVID-19). Thailand Situation Report. https://www.who.int/docs/default-source/searo/thailand/2020-03-15-tha-sitrep-22-covid19.pdf. Accessed 20February2022

11. WHO (2021) Coronavirus disease 2019 (COVID-19). Thailand Situation Report-167. https://cdn.who.int/media/docs/default-source/searo/thailand/2021_04_05_eng_sitrep-167-covid19.pdf?sfvrsn=8b63d890_5. Accessed 20December2021

12. WHO (2021) Coronavirus disease 2019 (COVID-19). Thailand Situation Report-170. https://cdn.who.int/media/docs/default-source/searo/thailand/2021_04_09_eng_sitrep_170-covid19.pdf?sfvrsn=eee4c530_8. Accessed on 20December2021

13. WHO (2021) Coronavirus disease 2019 (COVID-19). Thailand Situation Report-191. https://cdn.who.int/media/docs/default-source/searo/thailand/2021_07_08_eng-sitrep-191-covid19-revision.pdf?sfvrsn=7c57b84_3. Accessed 20December2021

14. Ministry of public health (MOPH), Thailand (2022) Coronavirus disease 2019 (COVID-19) WHO Thailand Situation. https://ddc.moph.go.th/covid19-dashboard/ (In Thai). Accessed 15January2022

15. Puenpa J, Suwannakarn K, Chansaenroj J, Nilyanimit P, Yorsaeng R, Auphimai C, Kitphati R, Mungaomklang A, Kongklieng A, Chirathaworn C, Wanlapakorn N, Poovorawan Y (2020) Molecular epidemiology of the first wave of severe acute respiratory syndrome coronavirus 2 infection in Thailand in 2020. Sci Rep 610:16602. https://doi:10.1038/s41598-020-73554-7

16. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95−98

17. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636−W641. https://doi:10.1093/nar/gkz268

18. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725−2729. https://doi:10.1093/molbev/mst197

19. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 4:vey016. https://doi:10.1093/ve/vey016

20. Public Health England (2021) Coronavirus disease 2019 (COVID-19). Variants: distribution of cases data. https://www.gov.uk/government/publications/covid-19-variants-genomically-confirmed-case-numbers/variants-distribution-of-cases-data. Accessed 1June2022

21. Callaway E (2021) Delta Coronavirus Variant: Scientists Brace for Impact. Nature 595:17−18. https://doi:10.1038/d41586-021-01696-3

22. Poudel S, Ishak A, Perez-Fernandez J, Garcia E, León-Figueroa DA, Romaní L, Bonilla-Aldana DK, Rodriguez-Morales AJ (2022) Highly mutated SARS-CoV-2 Omicron variant sparks significant concern among global experts - What is known so far? Travel Med Infect Dis 45:102234. https://doi:10.1016/j.tmaid.2021.102234

23. Tan KK, Tan JY, Wong JE, Teoh BT, Tiong V, Abd-Jamil J, Nor'e SS, Khor CS, Johari J, Yaacob CN, Zulkifli MM, CheMatSeri A, Mahfodz NH, Azizan NS, AbuBakar S (2021) Emergence of B.1.524(G) SARS-CoV-2 in Malaysia during the third COVID-19 epidemic wave. Sci Rep 11:22105. https://doi:10.1038/s41598-021-01223-4

24. Phuong HVM, Tung TS, Trang UTH, Hang NLK, Son NV, Hien PT, Thanh LT, Cuong VD, Thanh TT, Nhan NTT, Duong TN, Nghia ND, Tu TA, Choisy M, Rabaa MA, van Rogier H, Anh DD, Mai LQ (2021) Novel Mutation of SARS-CoV-2, Vietnam, July 2020. Emerg Infect Dis 27:1519−1521. https://doi:10.3201/eid2705.210013

25. Yurkovetskiy L, Wang X, Pascal KE et al (2020) Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. Cell 183:739−751. https://doi:10.1016/j.cell.2020.09.032

26. Tang H, Gao L, Wu Z, Meng F, Zhao X, Shao Y, Shi X, Qiao S, An J, Du X, Qin FX (2021) Characterization of SARS-CoV-2 Variants N501Y.V1 and N501Y.V2 Spike on Viral Infectivity. Front Cell Infect Microbiol 11:720357. https://doi:10.3389/fcimb.2021.720357

27. Plante JA, Liu Y, Liu J et al (2021) Spike mutation D614G alters SARS-CoV-2 fitness. Nature 592:116−121. https://doi:10.1038/s41586-020-2895-3

28. Gobeil SM, Janowska K, McDowell S, Mansouri K, Parks R, Manne K, Stalls V, Kopp MF, Henderson R, Edwards RJ, Haynes BF, Acharya (2021) D614G Mutation Alters SARS-CoV-2 Spike Conformation and Enhances Protease Cleavage at the S1/S2 Junction. Cell Rep 34:108630. https://doi:10.1016/j.celrep.2020.108630

29. Böhmer MM, Buchholz U, Corman VM et al (2020) Investigation of a COVID-19 outbreak in Germany resulting from a single travel-associated primary case: a case series. Lancet Infect Dis 20:920–928. https://doi:10.1016/S1473-3099(20)30314-5

30. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, Navarro MJ, Bowen JE, Tortorici MA, Walls AC, King NP, Veesler D, Bloom JD (2020) Deep Mutational Scanning of SARS-CoV-2 Receptor BindingDomain Reveals Constraints on Folding and ACE2 Binding. Cell 182:1295–1310. https://doi:10.1016/j.cell.2020.08.012

31. Makowski L, Olson-Sidford W, W-Weisel J (2021) Biological and Clinical Consequences of Integrin Binding via a Rogue RGD Motif in the SARS CoV-2 Spike Protein. Viruses 13:146. https://doi:10.3390/v13020146

32. Fratev F (2021) N501Y and K417N Mutations in the Spike Protein of SARS-CoV-2 Alter the Interactions with Both hACE2 and Human-Derived Antibody: A Free Energy of Perturbation Retrospective Study. J Chem Inf Model 61:6079–6084. https://doi:10.1021/acs.jcim.1c01242

33. Jangra S, Ye C, Rathnasinghe R et al (2021) SARS-CoV-2 Spike E484K Mutation Reduces Antibody Neutralisation. Lancet Microbe 2:e283–e284. https://doi:10.1016/S2666-5247(21)00068-9

34. Andreano E, Piccini G, Licastro D, Casalino L, Johnson NV, Paciello I, Dal Monego S, Pantano E, Manganaro N, Manenti A, Manna R, Casa E, Hyseni I, Benincasa L, Montomoli E, Amaro RE, McLellan JS, Rappuoli R(2020) SARS-CoV-2 escape in vitro from a highly neutralizing COVID-19 convalescent plasma. bioRxiv [Preprint] 28: 2020.
12.28.424451
. https://doi:10.1101/2020.12.28.424451

35. Baral P, Bhattarai N, Hossen ML, Stebliankin V, Gerstman BS, Narasimhan G, Chapagain PP (2021) Mutation-Induced Changes in the Receptor-Binding Interface of the SARS-CoV-2 Delta Variant B. 1.617. 2 and Implications for Immune Evasion. Biochem Biophysic Res Comm 574:14–19. https://doi:10.1016/j.bbrc.2021.08.036

36. Ghanchi NK, Nasir A, Masood KI, Abidi SH, Mahmood SF, Kanji A, Razzak S, Khan W, Shahid S, Yameen M, Raza A, Ashraf J, Ansar Z, Dharejo MB, Islam N, Hasan Z, Hasan R (2021) Higher entropy observed in SARS-CoV-2 genomes from the first COVID-19 wave in Pakistan. PLoS ONE 16:e0256451. https://doi:10.1371/journal.pone.0256451

37. Duchene S, Featherstone L, Haritopoulou-Sinanidou M, Rambaut A, Lemey P, Baele G (2020) Temporal signal and the phylodynamic threshold of SARS-CoV-2. Virus Evol 6:veaa061. https://doi:10.1093/ve/veaa061

38. Li X, Wang W, Zhao X, Zai J, Zhao Q, Li Y, Chaillon A (2020) Transmission dynamics and evolutionary history of 2019-nCoV. J Med Virol 92:501–511. https://doi:10.1002/jmv.25701

39. Nie Q, Li X, Chen W, Liu D, Chen Y, Li H, Li D, Tian M, Tan W, Zai J (2020) Phylogenetic and phylodynamic analyses of SARS-CoV-2. Virus Res 287:198098. https://doi:10.1016/j.virusres.2020.198098

40. Ko K, Nagashima S, Ouoba EB, Akita S, Sugiyama T, Ohisa A, Sakaguchi M, Tahara T, Ohge H, Ohdan H, Kubo H, Kishita T, Kuwabara E, Takahashi M, Tanaka K J (2021) Molecular characterization and the mutation pattern of SARS-CoV-2 during first and second wave outbreaks in Hiroshima, Japan. PLoS ONE 16:e0246383. https://doi:10.1371/journal.pone.0246383

Figures

Figure 1

Timeline showing the SARS-CoV-2 epidemic wave in Thailand, 2020-2022 [9].

Figure 2

The timecourse of variant distribution in Thailand during 2020-2022.

Figure 3

Maximum likelihood estimates of the phylogenetic relationship for SARS-CoV-2 detected among Thai patients during 2020-2022 compared with Southeast Asian country sequences (red triangle). The tree was generated on the basis of the complete spike nucleotide sequence data. Scale bar indicates the number of nucleotide substitutions per site.

Figure 4

A Bayesian phylogenetic tree of complete spike gene of the 2020-2022 SARS-CoV-2 from Thailand. Posterior support for major nodes is shown.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.pdf