Genomic surveillance of SARS-CoV-2 in Thailand reveals mixed imported populations, a local lineage expansion and a virus with truncated ORF7a

Elizabeth M Batty^{1,2,3,12}, Theerarat Kochakarn^{1,4,5,12}, Bhakbhoom Panthan^{1,6,12}, Krittikorn Kümpornsinsin^{1,7}, Poramate Jiaranai^{1,6}, Arporn Wangwiwatsin^{1,8}, Namfon Kotanan^{1,4}, Peera Jaruampornpan⁹, Treewat Watthanachockchai¹⁰, Kingkan Rakmanee¹⁰, Insee Sensorn^{1,6}, Somnuek Sungkanuparph¹¹, Ekawat Pasomsub^{1,10}, Thanat Chookajorn^{1,3} and Wasun Chantratita^{1,6}

¹COVID-19 Network Investigations Alliance (CONI), Bangkok, Thailand.

²Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand.

³Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

⁴Genomics and Evolutionary Medicine Unit (GEM), Center of Excellence in Malaria Research, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand.

⁵Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

⁶Center for Medical Genomics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

⁷Wellcome Sanger Institute, Hinxton, Cambridgeshire, United Kingdom.

⁸Department of Biology, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand.

⁹Virology and Cell Technology Research Team, National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani, Thailand.

¹⁰Division of Virology, Department of Pathology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

¹¹Chakri Naruebodindra Medical Institute, Faculty of Medicine Ramathibodi Hospital,
Mahidol University, Samut Prakan, Thailand.

¹²These authors contributed equally to the work

Correspondence: Thanat Chookajorn (thanat.cho@mahidol.edu) and Wasun Chantratita
(wasun.cha@mahidol.ac.th)

Keyword: COVID-19; Genomic Surveillance; SARS-CoV-2; Thailand

Abstract

Coronavirus Disease 2019 (COVID-19) is a global public health threat. Genomic surveillance of SARS-CoV-2 was implemented during March 2020 at a major diagnostic hub in Bangkok, Thailand. Several virus lineages supposedly originated in many countries were found, and a Thai-specific lineage, designated A/Thai-1, has expanded to be predominant in Thailand. A virus sample in the SARS-CoV-2 A/Thai-1 lineage contains a frame-shift deletion at ORF7a, encoding a putative host antagonizing factor of the virus.

Introduction

Coronavirus Disease 2019 (COVID-19) has reached the status of global pandemic. Genomic surveillance of its etiological virus, SARS-CoV-2, plays an important role in epidemiological investigations and transmission control strategies [1]. Genetic variation data of the virus could reveal transmission chains between infected individuals and could even map the connection between outbreak cohorts. Thailand has suffered from the spread of COVID-19 with the total number of confirmed cases over 3,000 and with more than 120,000 individuals screened as of May 2020. Since January 2020, when both imported and locally-transmitted COVID-19 cases were reported in Thailand, the country has implemented several measures to combat COVID-19 at a national scale [2-4].

Genomic surveillance could be a powerful tool in the implementation of the national COVID-19 control strategy in Thailand. ARTIC multiplex tiling PCR allows whole-genome sequencing with minuscule amount of material by generating genome-wide overlapping amplicons, which has led to its success during the Zika virus outbreak investigation in Brazil [5, 6]. Using leftover RNA samples from a standard RT-PCR diagnosis, the genomic information of SARS-CoV-2 can be decoded in less than a week. The data presented here provide an insight into the genetic repertoire, origins and viral lineages of SARS-CoV-2 in Thailand. The information is particularly important given the multiple introduction events into the country and the local expansion of the Thai-specific SARS-CoV-2 lineages.

Genomic surveillance of SARS-CoV-2 populations in Thailand

We sequenced 27 anonymized RT-qPCR positive samples from Ramathibodi Hospital in Bangkok during March 13-28, 2020 (Supplementary Table 1) [EC approval number: MURA2020/676]. The hospital acted as one of the major diagnostic hubs for COVID-19 in Bangkok during the study period. Enrichment and amplification steps were done according to

the ARTIC Network protocol with ARTIC primer version 2 [7]. The libraries were prepared using KAPA HyperPrep and KAPA Library Amplification kits and subsequently sequenced using with a MiSeq Reagent Kit v2 according to the manufactures' protocols. Variant calling was performed using the ncov2019-artic-nf pipeline (<https://github.com/connor-lab/ncov2019-artic-nf>). Consensus sequences were used to construct the maximum-likelihood and Bayesian phylogenetic trees with recommended representatives from various lineages worldwide utilizing IQ-TREE 2.0 and BEAST v1.10.4, respectively (Supplementary Table 2) [8-10]. Interestingly, Thailand appears to have had multiple introduction events of SARS-CoV-2 into the country, as evidenced by at least six separate clusters in the maximum-likelihood tree (Figure 1 and Supplementary Figure 1) [11]. Based on a Pangolin classification system (Database version 27 April 2020), they are grouped into A, B.1, B.1.5 and B.4 lineages (Figure 1 and Supplementary Figure 1) [9, 12]. Considering the origins and lineage branches, these SARS-CoV-2 lineages are likely to have recent ancestors outside Thailand. For example, six of the samples in the B.1 lineage in our collection are grouped tightly with virus samples commonly found in the United States of America and Europe and collected during the same period (March 2020) as visualized by Nextstrain Timetree (Supplementary Figure 2). The constructed Bayesian tree displayed the similar structure to that of the maximum-likelihood approach (Supplementary Figure 3).

It is worth noting the local expansion of a putative Thai specific lineage (Figure 2). This cluster of viruses has passed the criteria of a novel lineage as follows: (a) exhibits two shared nucleotide differences from the ancestral lineage [cut point ≥ 1], (b) contains 11 genomes with $> 95\%$ of the genome sequenced [cut point ≥ 5], (c) exhibits one shared nucleotide change (27,877G $\rightarrow$ U) among the ongoing transmission groups from March 2020 [cut point ≥ 1], and (d) has 99% bootstrap value for the lineage defining node [cut point $> 70\%$] [9]. This lineage, designated A/Thai-1 (Figure 1 and 2), descended from the original A

lineage (based on the maximum-likelihood based classification system), which was first reported in China before expanding into various countries in Asia, Europe, North America, South America and Australia [9]. This A/Thai-1 branch is separated from the rest of the original A lineage and subgroups. Upon visual inspection in Nextstrain, only one Malaysian sample (MKAK-CL-2020-5096) is the closest to A/Thai-1, but only with 63% bootstrap value and one shared lineage-specific nucleotide substitution (4,390G→U) (Supplementary Figure 4). Non-synonymous mutations unique to A/Thai-1 are 20,134G→U (ORF1b) and 24,047G→A (Spike protein) (the full mutation list is shown in Table 1). Among the changes, 20,134G→U mutation has been independently found in two samples in lineage B.1 from the Netherlands and USA. It remains to be determined with a larger sample size whether this is the result of convergent evolution or genetic recombination. This pattern of homoplasy was also hypothetically linked to putative RNA editing [13].

Thailand/Bangkok-0018, a sample in the A/Thai-1 lineage, contains a 4-nt frame-shift deletion at position 27,694-27,697, causing a premature truncation in ORF7a, which now contains five altered amino acid residues and loses the 16 original C-terminal residues (Figure 3). The deletion was confirmed by Sanger sequencing twice using two independent RT-PCR reactions (Supplementary Figure 5). The frame-shift mutation alters approximately one-sixth (21/121 residues) of the ORF7a protein. Based on protein homology to SARS-CoV, the missing region corresponds to a transmembrane helix and an ER retrieval motif, required for antagonizing a host antiviral factor [14, 15]. One sample from Arizona, USA also contains an 81-nt in-frame deletion in the ORF7a gene [16]. So far, only one sample in A/Thai-1 appears to have this frame-shift deletion. It is tempting to speculate on the relationship between ORF7a deletions and virus attenuation. However, further investigations by laboratory-based functional experiments are needed before reaching any conclusion on their biological and clinical significance.

When the analysis was extended to 22 additional genomes, independently deposited in GISAID by the Thai National Influenza Center and the Thai National Institute of Health, samples collected in January 2020 are grouped closely with the B lineage from China including the Wuhan-Hu-1 reference (Supplementary Table 3). The genetic repertoires from this additional collection also support the notion of multiple virus lineages introduced into Thailand. Nine of them also fall into A/Thai-1, making it the largest lineage in Thailand during the period of March 2020 (22/49 genomes).

Implication of the findings

Genomic surveillance is likely to be pivotal in the identification and the elimination of transmission cohorts and chains [17, 18]. The genetic composition presented here suggests the necessity for screening and monitoring international travelers during the period of COVID-19 pandemic. The local expansion of A/Thai-1 has created a new evolutionary branch unique to Thailand, which inevitably requires this lineage to be investigated for its compatibility to diagnosis and vaccine tools under development.

Acknowledgment

The work here was supported by Ramathibodi Foundation, TCELS, NRCT and Mahidol University. The authors acknowledge NSTDA Supercomputer Center (ThaiSC) for providing computing resources for this work. We are grateful for the comment and suggestion from P. Wilairat and Y. Yuthavong. We appreciate the contribution from K. Joonlasak, A. Huang, A.R. Jones, S. Fernandez and C. Klungthong (Armed Forces Research Institute of Medical Sciences). The computational aspects of this research were supported by the Wellcome Trust Core Award Grant Number 203141/Z/16/Z and the NIHR Oxford BRC. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Gudbjartsson, D.F., et al., *Spread of SARS-CoV-2 in the Icelandic Population*. N Engl J Med, 2020.
2. Okada, P., et al., *Early transmission patterns of coronavirus disease 2019 (COVID-19) in travellers from Wuhan to Thailand, January 2020*. Euro Surveill, 2020. **25**(8).
3. Pongpirul, W.A., et al., *Journey of a Thai Taxi Driver and Novel Coronavirus*. N Engl J Med, 2020. **382**(11): p. 1067-1068.
4. Hinjoy, S., et al., *A self-assessment of the Thai Department of Disease Control's communication for international response at early phase to the COVID-19*. Int J Infect Dis, 2020.
5. Quick, J., et al., *Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples*. Nat Protoc, 2017. **12**(6): p. 1261-1276.
6. Giovanetti, M., et al., *Genomic and Epidemiological Surveillance of Zika Virus in the Amazon Region*. Cell Rep, 2020. **30**(7): p. 2275-2283 e7.
7. Quick, J. *nCoV-2019 sequencing protocol*. 2020; Available from: <https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w>.
8. Drummond, A.J., et al., *Bayesian phylogenetics with BEAUti and the BEAST 1.7*. Mol Biol Evol, 2012. **29**(8): p. 1969-73.
9. Rambaut, A., E.C. Holmes, and O.G. Pybus, *A dynamic nomenclature for SARS-CoV-2 to assist genomic epidemiology*. bioRxiv 2020. Available from: <https://www.biorxiv.org/content/10.1101/2020.04.17.046086v1>
10. Minh, B.Q., et al., *IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era*. Mol Biol Evol, 2020. **37**(5): p. 1530-1534.

11. Hadfield, J., et al., *Nextstrain: real-time tracking of pathogen evolution*.
Bioinformatics, 2018. **34**(23): p. 4121-4123.
12. O'Toole, A.M., JT. *Phylogenetic Assignment of Named Global Outbreak LINEages*.
[cited 6 May 2020]; Available from: <https://github.com/hCoV-2019/pangolin>.
13. Simmonds, P., *Rampant C->U hypermutation in the genomes of SARS-CoV-2 and other coronaviruses – causes and consequences for their short and long evolutionary trajectories*. bioRxiv, 2020. Available from:
<https://www.biorxiv.org/content/10.1101/2020.05.01.072330v1>
14. Taylor, J.K., et al., *Severe Acute Respiratory Syndrome Coronavirus ORF7a Inhibits Bone Marrow Stromal Antigen 2 Virion Tethering through a Novel Mechanism of Glycosylation Interference*. J Virol, 2015. **89**(23): p. 11820-33.
15. Fielding, B.C., et al., *Characterization of a unique group-specific protein (U122) of the severe acute respiratory syndrome coronavirus*. J Virol, 2004. **78**(14): p. 7311-8.
16. Holland, L.A., et al., *An 81 nucleotide deletion in SARS-CoV-2 ORF7a identified from sentinel surveillance in Arizona (Jan-Mar 2020)*. J Virol, 2020.
17. Kumpornsin, K., T. Kochakarn, and T. Chookajorn, *The resistome and genomic reconnaissance in the age of malaria elimination*. Dis Model Mech, 2019. **12**(12).
18. Gardy, J.L. and N.J. Loman, *Towards a genomics-informed, real-time, global pathogen surveillance system*. Nat Rev Genet, 2018. **19**(1): p. 9-20.

Figure Legend

Figure 1. Cladogram based on a maximum-likelihood tree showing Thai populations of SARS-CoV-2. The cladogram represents a matching maximum-likelihood tree (1,000 bootstrap replicates) shown in Supplementary Figure 1. Genomes generated in this study are labelled in blue with remaining lineage representatives labelled in black. Bootstrap values (≥ 75) are shown at the nodes. Thai virus genomes independently generated by other groups and deposited in GISAID are labelled in green. Genomes collected from January 2020 are similar to that of reference Wuhan-Hu-1. Our data showed at least four independently lineages, and two additional events of lineage B from January 2020 and lineage B.6 (shown in green), could be recognized as potential introduction events. The A and B lineages are coloured in orange and cyan, respectively.

Figure 2. Maximum-likelihood tree containing A/Thai-1 subset. For visualization of A/Thai-1 lineage, the A/Thai-1 genomes from Figure 1 generated in this study are presented. The lineage defining node has a bootstrap value of 99 (1,000 replicates). Mutations that define each branch point are shown. 20,134G→U in Thailand_Bangkok-0030 could not be called.

Figure 3. Diagram depicting Thailand/Bangkok-0018 frame-shift deletion in ORF7a. The upper diagram shows the gene organization of the SARS-CoV-2 genome. The functional domains of ORF7a are indicated based on SARS-CoV. Sequences from SARS-CoV, the SARS-CoV-2 Wuhan-Hu-1 reference, the Arizona sample with an 81-nucleotide deletion (EPI_ISL_424669) and Thailand/Bangkok-0018 are aligned to demonstrate the altered region in red box.

Figure 1

Figure 2

Figure 3

Table 1. Mutations present in A/Thai-1 in comparison with Wuhan-Hu-1.

Position	Ref	Alt	Syn/ Non-syn	Gene	Samples	Note
895	A	G	S	ORF1a	0018, 0019, 0025, 0026, 0028, 0029, 0030, 0034, 0035, 0037, 0040, 0041	Unique to A/Thai-1
2942	C	U	S	ORF1a	0018, 0019, 0025, 0026, 0028, 0029, 0030, 0034, 0035, 0037, 0040, 0041	Unique to A/Thai-1
4390	G	U	NS	ORF1a	0018, 0019, 0025, 0026, 0028, 0029, 0030, 0034, 0035, 0037, 0040, 0041	Shared with Malaysia/MKAK-CL-2020-2096 (A lineage)
8782	C	U	S	ORF1a	0018, 0019, 0025, 0026, 0028, 0029, 0030, 0034, 0035, 0037, 0040, 0041	Shared among ancestral A lineages
9598	C	U	S	ORF1a	0018, 0019, 0025, 0026, 0028, 0029, 0030, 0034, 0035, 0037, 0040, 0041	Shared with Scotland/EDB2081 (B.1 lineage)
14212	G	A	NS	ORF1b	0025	-
20134	G	U	NS	ORF1b	0018, 0019, 0025, 0026, 0028, 0029, 0034, 0035, 0037, 0040, 0041	Cannot be called in 0030; Shared with Netherlands/NA_300 and USA/UN-UW-5172 (B.1 lineage)
21859	C	U	S	S	0035	-
24047	G	A	NS	S	0018, 0019, 0025, 0026, 0028, 0029, 0030, 0034, 0035, 0037, 0040, 0041	Unique to A/Thai-1
25047	C	U	NS	S	0030	-
25510	U	C	NS	ORF3a	0040	-
27694-27697	TTT C	-	Frame-shift deletion	ORF7a	0018	-
27877	G	U	NS	ORF7b	0018, 0019, 0026, 0028, 0030	Shared with Taiwan/170 and USA/UT-00514 (B.1 lineage)
28144	U	C	NS	ORF8	0018, 0019, 0025, 0026, 0028, 0029, 0030, 0034, 0035, 0037, 0040, 0041	Shared among ancestral A lineages

Supplementary Figure 2. Nextstrain Timetree presenting the grouping of the Thai SARS-CoV-2 B1 samples with the virus genomes from other countries during the same period. The data was retrieved on 6 May 2020.

Supplementary Figure 3. Bayesian phylogenetic tree of the SARS-CoV-2. Tree plotting was done with the same sample sets used for the maximum-likelihood tree.

Supplementary Figure 4. Nextstrain Timetree representing the A/Thai-1 lineage using the data from GISAID. The Malaysian virus genome MKAK-CL-2020-5096 is the closest one in the tree. The data was retrieved on 6 May 2020.

Supplementary Figure 5. Sanger sequencing result confirming the ORF7a deletion in BKK-0018. A sequencing chromatogram from the DNA regions (upper panel) corresponding to the deletion site (red arrow) is shown. The 4-nt deletion site marked in the red box causes a premature stop codon.

SARS-CoV-2_Wuhan-Hu-1	ATGAAAATTATTCTTTCTTGGCACTGATAACACTCGCTACTTGTGAGCTTTATCACTAC	60
SARS-CoV-2_BKK-0018	ATGAAAATTATTCTTTCTTGGCACTGATAACACTCGCTACTTGTGAGCTTTATCACTAC	60
SARS-CoV-2_Wuhan-Hu-1	CAAGAGTGTGTTAGAGGTACAACAGTACTTTTAAAGAACCTTGCTCTTCTGGAACATAC	120
SARS-CoV-2_BKK-0018	CAAGAGTGTGTTAGAGGTACAACAGTACTTTTAAAGAACCTTGCTCTTCTGGAACATAC	120
SARS-CoV-2_Wuhan-Hu-1	GAGGGCAATTCACCATTTTCATCCTCTAGCTGATAACAAATTTGCACTGACTTGCTTTAGC	180
SARS-CoV-2_BKK-0018	GAGGGCAATTCACCATTTTCATCCTCTAGCTGATAACAAATTTGCACTGACTTGCTTTAGC	180
SARS-CoV-2_Wuhan-Hu-1	ACTCAATTTGCTTTTGCTTGCTGCTGACGGCGTAAACACGCTCTATCAGTTACGTGCCAGA	240
SARS-CoV-2_BKK-0018	ACTCAATTTGCTTTTGCTTGCTGCTGACGGCGTAAACACGCTCTATCAGTTACGTGCCAGA	240
SARS-CoV-2_Wuhan-Hu-1	TCAGTTTCACCTAAACTGTTTCATCAGACAAGAGGAAGTTCAAGAACTTTACTCTCCAATT	300
SARS-CoV-2_BKK-0018	TCAGTTTCACCTAAACTGTTTCATCAGACAAGAGGAAGTTCAAGAACTTTACTCTCCAATT	300
SARS-CoV-2_Wuhan-Hu-1	TTTCTTATTGTTGCGGCAATAGTGTTTATAACACTTTGCTTCACACTCAAAGAAAGACA	360
SARS-CoV-2_BKK-0018	----TTATTGTTGCGGCAATAGTGTTTATAACACTTTGCTTCACACTCAAAGAAAGACA	356
	<i>original stop codon</i>	
SARS-CoV-2_Wuhan-Hu-1	GAATGA	366
SARS-CoV-2_BKK-0018	GAATGA	362

Supplementary table 1: List of samples used for genome

Name	GISAID ID	Collection date
Thailand/Bangkok-0017/2020	EPI_ISL_423039	2020-03-21
Thailand/Bangkok-0018/2020	EPI_ISL_450723	2020-03-20
Thailand/Bangkok-0019/2020	EPI_ISL_423040	2020-03-20
Thailand/Bangkok-0020/2020	EPI_ISL_423041	2020-03-21
Thailand/Bangkok-0021/2020	EPI_ISL_423042	2020-03-24
Thailand/Bangkok-0022/2020	EPI_ISL_423043	2020-03-18
Thailand/Bangkok-0025/2020	EPI_ISL_429164	2020-03-13
Thailand/Bangkok-0026/2020	EPI_ISL_429165	2020-03-13
Thailand/Bangkok-0028/2020	EPI_ISL_429166	2020-03-13
Thailand/Bangkok-0029/2020	EPI_ISL_429167	2020-03-13
Thailand/Bangkok-0030/2020	EPI_ISL_429168	2020-03-17
Thailand/Bangkok-0033/2020	EPI_ISL_429169	2020-03-17
Thailand/Bangkok-0034/2020	EPI_ISL_429170	2020-03-17
Thailand/Bangkok-0035/2020	EPI_ISL_429171	2020-03-18
Thailand/Bangkok-0036/2020	EPI_ISL_429172	2020-03-21
Thailand/Bangkok-0037/2020	EPI_ISL_429173	2020-03-21
Thailand/Bangkok-0038/2020	EPI_ISL_429174	2020-03-25
Thailand/Bangkok-0039/2020	EPI_ISL_429175	2020-03-28
Thailand/Bangkok-0040/2020	EPI_ISL_429176	2020-03-28
Thailand/Bangkok-0041/2020	EPI_ISL_429177	2020-03-28
Thailand/Bangkok-0042/2020	EPI_ISL_429178	2020-03-28
Thailand/Bangkok-0043/2020	EPI_ISL_429179	2020-03-28
Thailand/Bangkok-0044/2020	EPI_ISL_429180	2020-03-28
Thailand/Bangkok-0045/2020	EPI_ISL_429181	2020-03-28
Thailand/Bangkok-0046/2020	EPI_ISL_429182	2020-03-28
Thailand/Bangkok-0047/2020	EPI_ISL_429183	2020-03-28
Thailand/Bangkok-0048/2020	EPI_ISL_429184	2020-03-28

Supplementary table 2: List of representatives used for phylogenetics analyses

Name	GISAID ID	Collection date	Lineages
Australia/NSW09/2020	EPI_ISL_413595	2020-02-28	B.4
Australia/NSW13/2020	EPI_ISL_413599	2020-03-04	B.4
Australia/NSW14/2020	EPI_ISL_413600	2020-03-03	B.4
Australia/NSW146/2020	EPI_ISL_427682	2020-03-23	B.1.31
Australia/NSW147/2020	EPI_ISL_427764	2020-03-23	B.1.31
Australia/NSW149/2020	EPI_ISL_427775	2020-03-25	B.1.31
Australia/NT05/2020	EPI_ISL_419835	2020-02-24	B.5
Australia/VIC114/2020	EPI_ISL_419725	2020-03-18	B.1.5
Australia/VIC121/2020	EPI_ISL_419723	2020-03-18	B.2.2
Australia/VIC126/2020	EPI_ISL_419728	2020-03-20	A.2
Australia/VIC138/2020	EPI_ISL_419834	2020-02-23	A.5
Australia/VIC145/2020	EPI_ISL_419842	2020-03-17	B.2.5
Australia/VIC164/2020	EPI_ISL_419861	2020-03-18	B.1.5.3
Australia/VIC176/2020	EPI_ISL_419873	2020-03-18	B.2.4
Australia/VIC188/2020	EPI_ISL_419884	2020-03-19	B.6
Australia/VIC196/2020	EPI_ISL_419892	2020-03-19	B.1.13
Australia/VIC248/2020	EPI_ISL_419943	2020-03-21	B.1.31
Australia/VIC26/2020	EPI_ISL_419745	2020-03-10	B.3
Australia/VIC30/2020	EPI_ISL_419749	2020-03-10	B.4
Australia/VIC342/2020	EPI_ISL_426651	2020-03-20	B.2.4
Australia/VIC344/2020	EPI_ISL_426653	2020-03-20	B.2.4
Australia/VIC370/2020	EPI_ISL_426678	2020-03-22	B.1.25
Australia/VIC378/2020	EPI_ISL_426685	2020-03-22	B.2.4
Australia/VIC38/2020	EPI_ISL_419756	2020-03-11	B.1.6
Australia/VIC382/2020	EPI_ISL_426689	2020-03-23	B.1.25
Australia/VIC40/2020	EPI_ISL_419758	2020-03-11	A.1.2
Australia/VIC41/2020	EPI_ISL_419760	2020-03-11	A.1.2
Australia/VIC42/2020	EPI_ISL_419761	2020-03-11	B.6
Australia/VIC475/2020	EPI_ISL_426907	2020-03-28	B.1.23
Australia/VIC476/2020	EPI_ISL_426908	2020-03-28	B.1.23
Australia/VIC515/2020	EPI_ISL_426785	2020-03-23	B.1.25
Australia/VIC547/2020	EPI_ISL_426810	2020-03-25	B.4
Australia/VIC551/2020	EPI_ISL_426813	2020-03-25	B.1.25
Australia/VIC554/2020	EPI_ISL_426815	2020-03-25	B.1.25
Australia/VIC555/2020	EPI_ISL_426816	2020-03-25	B.1.3
Australia/VIC565/2020	EPI_ISL_426825	2020-03-25	A.2
Australia/VIC571/2020	EPI_ISL_426831	2020-03-25	B.1.2
Australia/VIC584/2020	EPI_ISL_426841	2020-03-25	A.2
Australia/VIC59/2020	EPI_ISL_419771	2020-03-13	B.6
Australia/VIC591/2020	EPI_ISL_426848	2020-03-22	B.3
Australia/VIC598/2020	EPI_ISL_426855	2020-03-25	A.2
Australia/VIC60/2020	EPI_ISL_419772	2020-03-13	A.5
Australia/VIC605/2020	EPI_ISL_426860	2020-03-27	B.1.31
Australia/VIC613/2020	EPI_ISL_426867	2020-03-26	B.1.10
Australia/VIC623/2020	EPI_ISL_426927	2020-03-28	A.5

Australia/VIC628/2020	EPI_ISL_426932	2020-03-29	B.1.23
Australia/VIC646/2020	EPI_ISL_426942	2020-03-24	A.3
Australia/VIC78/2020	EPI_ISL_419790	2020-03-14	B
Australia/VIC785/2020	EPI_ISL_427060	2020-03-30	B.1.23
Australia/VIC789/2020	EPI_ISL_427064	2020-03-31	B.1.22
Australia/VIC79/2020	EPI_ISL_419791	2020-03-14	B.2.1
Australia/VIC80/2020	EPI_ISL_419792	2020-03-14	B.2.1
Australia/VIC81/2020	EPI_ISL_419793	2020-03-14	B.2.1
Australia/VIC818/2020	EPI_ISL_427090	2020-04-03	B.1.22
Australia/VIC819/2020	EPI_ISL_427091	2020-04-03	B.6
Australia/VIC82/2020	EPI_ISL_419794	2020-03-14	B.2.1
Australia/VIC83/2020	EPI_ISL_419799	2020-03-14	B.2
Australia/VIC86/2020	EPI_ISL_419796	2020-03-15	A.3
Australia/VIC87/2020	EPI_ISL_419797	2020-03-15	B.2.1
Australia/VIC91/2020	EPI_ISL_419804	2020-03-15	B.1.5
Australia/VIC912/2020	EPI_ISL_426881	2020-03-24	B.1.23
Australia/WA07/2020	EPI_ISL_420537	2020-03-13	B.2.5
Beijing/233/2020	EPI_ISL_413520	2020-01-28	A
Beijing/235/2020	EPI_ISL_413521	2020-01-28	A
Belgium/BM-03012/2020	EPI_ISL_415154	2020-03-01	B.1.16
Belgium/LY-030575/2020	EPI_ISL_420331	2020-03-05	B.1.16
Belgium/MNL-0324130/2020	EPI_ISL_420387	2020-03-24	B.2.3
Belgium/SN-03031/2020	EPI_ISL_416469	2020-03-03	B.1.16
Belgium/SR-0319112/2020	EPI_ISL_420369	2020-03-19	B.2.3
Belgium/ULG-10022/2020	EPI_ISL_421197	2020-03-31	B.2.3
Belgium/ULG-10047/2020	EPI_ISL_421214	2020-03-30	B.1.6
Belgium/ULG-10062/2020	EPI_ISL_424664	2020-04-05	B.1.6
Belgium/ULG-10085/2020	EPI_ISL_424645	2020-04-04	B.1.9
Belgium/ULG-10122/2020	EPI_ISL_427370	2020-04-07	B.1.6
Belgium/ULG-3662/2020	EPI_ISL_417008	2020-03-07	B.2.3
Belgium/ULG-6670/2020	EPI_ISL_417018	2020-03-14	B.1.9
Belgium/ULG-7729/2020	EPI_ISL_418630	2020-03-17	B.1.6
Belgium/ULG-8578/2020	EPI_ISL_418638	2020-03-19	B.1.12
Belgium/ULG-8617/2020	EPI_ISL_418639	2020-03-19	B.1.9
Belgium/ULG-9694/2020	EPI_ISL_418654	2020-03-22	B.1.9
Belgium/VAG-03013/2020	EPI_ISL_415155	2020-03-01	B.1.16
Belgium/VL-0318104/2020	EPI_ISL_420360	2020-03-18	B.1.16
Canada/ON_PHL0178/2020	EPI_ISL_418342	2020-03-11	A.1.2
Canada/ON_PHL1083/2020	EPI_ISL_418340	2020-03-09	A.1.2
Canada/ON_PHL3459/2020	EPI_ISL_418376	2020-03-12	B.6
Canada/ON_PHL7590/2020	EPI_ISL_418373	2020-03-14	B.1.10
DRC/523/2020	EPI_ISL_420849	2020-03-28	B.1.1
England/200991076/2020	EPI_ISL_414524	2020-03-01	B.1.20
England/20104004902/2020	EPI_ISL_417226	2020-03-02	B.1.7
England/20104008802/2020	EPI_ISL_417231	2020-03-03	B.1.7
England/20106004803/2020	EPI_ISL_417240	2020-03-03	B.1.7
England/20108007002/2020	EPI_ISL_417255	2020-03-04	B.1.7

England/20109050406/2020	EPI_ISL_417265	2020-03-08	B.1.1
England/20109054806/2020	EPI_ISL_417280	2020-03-07	B.2.2
England/20109093706/2020	EPI_ISL_417286	2020-03-05	B.8
England/20109094006/2020	EPI_ISL_417289	2020-03-06	B.2.2
England/20130034204/2020	EPI_ISL_423605	2020-03-22	B.1.11
England/20132068204/2020	EPI_ISL_423796	2020-03-23	B.1.30
England/20132068304/2020	EPI_ISL_423797	2020-03-23	B.1.30
England/20134081904/2020	EPI_ISL_423181	2020-03-20	B.1.20
England/201380055/2020	EPI_ISL_421810	2020-03-25	B.1.5.3
England/201380056/2020	EPI_ISL_421811	2020-03-25	B.1.5.3
England/20138083204/2020	EPI_ISL_421839	2020-03-27	B.1.20
England/20139024804/2020	EPI_ISL_421894	2020-03-27	B.1.13
England/20139053804/2020	EPI_ISL_420709	2020-03-28	B.1.13
England/20139058104/2020	EPI_ISL_420723	2020-03-28	B.1.1
England/20139058304/2020	EPI_ISL_420724	2020-03-27	B.1.13
England/20139059004/2020	EPI_ISL_420727	2020-03-28	B.1.13
England/20140039304/2020	EPI_ISL_420755	2020-03-27	B.1.5
England/20140071104/2020	EPI_ISL_423411	2020-03-28	B.1.30
England/20142019504/2020	EPI_ISL_423443	2020-03-30	B.1.30
England/20142060404/2020	EPI_ISL_420766	2020-03-27	B.2.5
England/20142061804/2020	EPI_ISL_420772	2020-03-30	B.1.11
England/20142062904/2020	EPI_ISL_420775	2020-03-30	B.1.20
England/20144074604/2020	EPI_ISL_423564	2020-04-01	B.1.30
England/20144076304/2020	EPI_ISL_421964	2020-04-01	B.1.11
England/20146004804/2020	EPI_ISL_423573	2020-04-01	B.1.5.3
England/CAMB-738A6/2020	EPI_ISL_425249	2020-03-31	B.1.5.3
England/CAMB-7477B/2020	EPI_ISL_425333	2020-03-29	B.1.20
England/NOTT-10E256/2020	EPI_ISL_425517	2020-03-20	B.8
England/NOTT-10E335/2020	EPI_ISL_425530	2020-03-21	B.3
England/NOTT-10E3BD/2020	EPI_ISL_425537	2020-03-21	B.2.2
England/NOTT-10E46F/2020	EPI_ISL_425548	2020-03-22	B.1.11
England/NOTT-10E5E4/2020	EPI_ISL_425569	2020-03-23	B.8
England/NOTT-10E6C3/2020	EPI_ISL_425581	2020-03-24	B.1.11
England/NOTT-10E7A2/2020	EPI_ISL_425593	2020-03-25	B.1.7
England/NOTT-10EAA9/2020	EPI_ISL_425635	2020-03-30	B.3
England/NOTT-10EAD6/2020	EPI_ISL_425638	2020-03-30	B.8
England/SHEF-C0646/2020	EPI_ISL_420278	2020-03-21	B.1.1
Finland/14M82/2020	EPI_ISL_418411	2020-03-14	B
France/ARA11036/2020	EPI_ISL_418432	2020-03-18	B.2
Germany/NRW-33/2020	EPI_ISL_419550	2020-03-16	B.1.8
Germany/NRW-34/2020	EPI_ISL_419551	2020-03-16	B.1.8
HongKong/HKPU1_2101/2020	EPI_ISL_417176	2020-01-21	A
HongKong/HKPU29_0102/2020	EPI_ISL_417187	2020-02-08	B.7
HongKong/HKPU30_2901/2020	EPI_ISL_417188	2020-02-08	B.7
HongKong/HKPU34_3001/2020	EPI_ISL_417197	2020-02-09	B.7
HongKong/HKPU70-1302/2020	EPI_ISL_419243	2020-02-22	B.7
HongKong/VM20002907/2020	EPI_ISL_414517	2020-02-25	B.7

Iceland/117/2020	EPI_ISL_417542	2020-03-17	B.1.5.1
Iceland/118/2020	EPI_ISL_417543	2020-03-17	B.1.5.1
Iceland/238/2020	EPI_ISL_417566	2020-03-17	B.1.14
Iceland/249/2020	EPI_ISL_417577	2020-03-17	A.1.1
Iceland/254/2020	EPI_ISL_417582	2020-03-17	B.1.8
Iceland/287/2020	EPI_ISL_417615	2020-03-18	A.1.1
Iceland/29/2020	EPI_ISL_417618	2020-03-16	A.1.1
Iceland/293/2020	EPI_ISL_417622	2020-03-18	B.1.8
Iceland/303/2020	EPI_ISL_417632	2020-03-18	B.1.14
Iceland/334/2020	EPI_ISL_417663	2020-03-14	B.1.10
Iceland/339/2020	EPI_ISL_417668	2020-03-15	B.2.2
Iceland/356/2020	EPI_ISL_424380	2020-03-19	B.1.5.1
Iceland/358/2020	EPI_ISL_424382	2020-03-19	B.2
Iceland/360/2020	EPI_ISL_424384	2020-03-19	B.1.14
Iceland/368/2020	EPI_ISL_424392	2020-03-19	B.1.5.1
Iceland/369/2020	EPI_ISL_424393	2020-03-19	B.1.5.1
Iceland/412/2020	EPI_ISL_424436	2020-03-19	A.1.1
Iceland/438/2020	EPI_ISL_424462	2020-03-20	A.1.1
Iceland/444/2020	EPI_ISL_424468	2020-03-20	B.1.14
Iceland/452/2020	EPI_ISL_424476	2020-03-19	B.1.14
Japan/DP0654/2020	EPI_ISL_416613	2020-02-17	B.5
Japan/TK-20-31-3/2020	EPI_ISL_413459	2020-02-20	B.5
Korea/BA-ACH_2604/2020	EPI_ISL_420799	2020-02-27	B.2
Luxembourg/LNS4134806/2020	EPI_ISL_419586	2020-03-14	B.1.12
Luxembourg/LNS5731562/2020	EPI_ISL_417529	2020-03-18	B.1.19
Luxembourg/LNS9086704/2020	EPI_ISL_419601	2020-03-15	B.1.12
Luxembourg/LNS9324837/2020	EPI_ISL_419602	2020-03-12	B.1.12
Luxembourg/LNS9982497/2020	EPI_ISL_419607	2020-03-14	B.1.12
Malaysia/MKAK-CL-2020-5096/2020	EPI_ISL_416885	2020-02-20	A
Netherlands/Friesland_6/2020	EPI_ISL_422571	2020-03-28	B.1.5.2
Netherlands/NA_126/2020	EPI_ISL_422670	2020-03-17	B.1.22
Netherlands/NA_190/2020	EPI_ISL_422725	2020-03-23	B.2.3
Netherlands/NA_28/2020	EPI_ISL_415485	2020-03-12	B.1.22
Netherlands/NA_296/2020	EPI_ISL_422605	2020-04-01	B.1.8
Netherlands/NA_301/2020	EPI_ISL_422610	2020-04-01	B.1.22
Netherlands/ZuidHolland_75/2020	EPI_ISL_422943	2020-03-26	B.1.5.2
Netherlands/ZuidHolland_82/2020	EPI_ISL_422950	2020-03-27	B.1.5.2
Netherlands/ZuidHolland_88/2020	EPI_ISL_422955	2020-03-27	B.1.5.2
Netherlands/ZuidHolland_90/2020	EPI_ISL_422957	2020-03-29	B.1.5.2
NewZealand/CoV002/2020	EPI_ISL_417212	2020-03-11	B.2.4
Scotland/CVR139/2020	EPI_ISL_425682	2020-03-16	B.2.5
Scotland/EDB222/2020	EPI_ISL_425948	2020-03-26	A.5
Shenzhen/SZTH-002/2020	EPI_ISL_406593	2020-01-13	A
Slovenia/808/2020	EPI_ISL_420541	2020-03-05	B.2.5
Spain/PaisVasco201602/2020	EPI_ISL_419709	2020-03-04	A.2
Turkey/GLAB-CoV012/2020	EPI_ISL_428346	2020-04-17	B.1.9
Uruguay/UY-5/2020	EPI_ISL_426480	2020-03-17	A.5

USA/AZ-TG273294/2020	EPI_ISL_426561	2020-03-31	B.1.29
USA/CruiseA-12/2020	EPI_ISL_413617	2020-02-20	B.5
USA/CruiseA-4/2020	EPI_ISL_413609	2020-02-21	B.5
USA/CT-UW-4238/2020	EPI_ISL_427162	2020-03-30	B.1.26
USA/CT-UW-4346/2020	EPI_ISL_426053	2020-03-30	B.1.26
USA/FL_6318/2020	EPI_ISL_420785	2020-03-02	B.1
USA/GA_1299/2020	EPI_ISL_420787	2020-03-03	B.1
USA/GA_1320/2020	EPI_ISL_420786	2020-03-03	A.3
USA/GA_1445/2020	EPI_ISL_420788	2020-03-04	A.3
USA/IL_1293/2020	EPI_ISL_420790	2020-03-01	B
USA/IL_1375/2020	EPI_ISL_420789	2020-03-01	B
USA/MN26-MDH26/2020	EPI_ISL_417194	2020-03-12	A.1
USA/NH_0004/2020	EPI_ISL_420791	2020-02-29	B.1
USA/NH_0008/2020	EPI_ISL_420792	2020-03-02	B.1
USA/NY_2929/2020	EPI_ISL_420793	2020-03-02	B.1
USA/NY-NYUMC124/2020	EPI_ISL_424931	2020-04-01	B.1.28
USA/NY-NYUMC18/2020	EPI_ISL_418203	2020-03-17	B.1.28
USA/NY-NYUMC39/2020	EPI_ISL_419700	2020-03-18	B.1.3
USA/NY-NYUMC42/2020	EPI_ISL_419703	2020-03-18	B.1.26
USA/NY-NYUMC44/2020	EPI_ISL_419705	2020-03-18	B.1.28
USA/NY-NYUMC60/2020	EPI_ISL_420575	2020-03-18	B.1.26
USA/NY-NYUMC62/2020	EPI_ISL_420577	2020-03-18	B.1.3
USA/NY-NYUMC66/2020	EPI_ISL_420581	2020-03-18	B.1.3
USA/NY-NYUMC72/2020	EPI_ISL_420588	2020-03-18	B.1.28
USA/NY-NYUMC74/2020	EPI_ISL_420590	2020-03-18	B.1.3
USA/NY-PV09151/2020	EPI_ISL_422519	2020-03-22	B.1.28
USA/NY-PV09164/2020	EPI_ISL_422528	2020-03-21	B.1.29
USA/OR_5430/2020	EPI_ISL_420794	2020-03-01	B.2
USA/OR-UW-1885/2020	EPI_ISL_424307	2020-03-23	B.1.2
USA/RI_0556/2020	EPI_ISL_420795	2020-03-01	B.1.5
USA/TX_2039/2020	EPI_ISL_420796	2020-02-29	B
USA/VA-DCLS-0021/2020	EPI_ISL_419713	2020-03-11	A
USA/WA-S87/2020	EPI_ISL_417140	2020-03-01	A.1
USA/WA-S89/2020	EPI_ISL_417142	2020-02-29	A.1
USA/WA-S90/2020	EPI_ISL_417143	2020-02-29	A.1
USA/WA-S96/2020	EPI_ISL_417149	2020-02-28	A.1
USA/WA-UW104/2020	EPI_ISL_416642	2020-03-11	B.1.27
USA/WA-UW-1470/2020	EPI_ISL_423028	2020-03-18	B.1.2
USA/WA-UW-1561/2020	EPI_ISL_423011	2020-03-19	B.1.27
USA/WA-UW-1706/2020	EPI_ISL_424227	2020-03-21	B.1.21
USA/WA-UW-1707/2020	EPI_ISL_424225	2020-03-21	B.1.2
USA/WA-UW-1724/2020	EPI_ISL_424229	2020-03-21	B.1.21
USA/WA-UW-1733/2020	EPI_ISL_424235	2020-03-21	B.1.21
USA/WA-UW-1741/2020	EPI_ISL_424238	2020-03-21	B.1.21
USA/WA-UW-1772/2020	EPI_ISL_424233	2020-03-21	B.1.5
USA/WA-UW-1797/2020	EPI_ISL_424259	2020-03-22	B.1.21
USA/WA-UW-1930/2020	EPI_ISL_424304	2020-03-24	B.1.26

USA/WA-UW-1947/2020	EPI_ISL_424327	2020-03-23	B.1.2
USA/WA-UW-2131/2020	EPI_ISL_426092	2020-03-25	B.1.27
USA/WI-06/2020	EPI_ISL_417200	2020-03-21	A.4
USA/WI-07/2020	EPI_ISL_417201	2020-03-21	A.4
USA/WI-09/2020	EPI_ISL_417203	2020-03-21	A.1.2
USA/WI-102/2020	EPI_ISL_425159	2020-03-30	B.1.27
USA/WI-145/2020	EPI_ISL_427452	2020-04-10	B.1.19
USA/WI-146/2020	EPI_ISL_427453	2020-04-06	A.4
USA/WI-147/2020	EPI_ISL_427454	2020-04-10	B.1.19
USA/WI-184/2020	EPI_ISL_428280	2020-03-22	A.3
USA/WI-22/2020	EPI_ISL_417514	2020-03-13	A.4
USA/WI-53/2020	EPI_ISL_421311	2020-03-18	B.1.27
USA/WI-61/2020	EPI_ISL_421319	2020-03-23	A.4
USA/WI-64/2020	EPI_ISL_421322	2020-03-24	B.1.19
USA/WI-65/2020	EPI_ISL_421323	2020-03-22	B.1.19
USA/WI-GMF-00707/2020	EPI_ISL_422461	2020-04-06	B.1.29
USA/WI-GMF-00744/2020	EPI_ISL_422462	2020-04-03	B.1.29
USA/WI-GMF-00921/2020	EPI_ISL_426160	2020-04-03	B.1.29
Wales/PHW32/2020	EPI_ISL_415920	2020-03-12	B.3
Wales/PHWC-25ABC/2020	EPI_ISL_422053	2020-03-24	B.8
Wales/PHWC-260AA/2020	EPI_ISL_422073	2020-03-24	B.1.24
Wales/PHWC-2617A/2020	EPI_ISL_422094	2020-03-25	B.1.1
Wales/PHWC-2666C/2020	EPI_ISL_422046	2020-03-27	B.1.10
Wales/PHWC-26796/2020	EPI_ISL_422016	2020-03-28	B.1.24
Wales/PHWC-26D3A/2020	EPI_ISL_422076	2020-03-30	B.1.24
Wales/PHWC-26DC1/2020	EPI_ISL_422088	2020-03-30	B.1.24
Wales/PHWC-26E46/2020	EPI_ISL_422075	2020-03-29	B.1.10
Wales/PHWC-26E64/2020	EPI_ISL_422143	2020-03-29	B.1.24
Wuhan-Hu-1/2019	EPI_ISL_402125	2019-12-31	B

Supplementary table 3: genomes from the Thai samples deposited in GISAID by other groups.

Name	GISAID ID	Collection date
Thailand/NIH-15/2020	EPI_ISL_434692	2020-01-05
Thailand/NIH-162/2020	EPI_ISL_434694	2020-01-22
Thailand/NIH-1889/2020	EPI_ISL_434697	2020-03-03
Thailand/NIH-190/2020	EPI_ISL_434695	2020-01-22
Thailand/NIH-2294/2020	EPI_ISL_434698	2020-03-13
Thailand/NIH-2296/2020	EPI_ISL_434699	2020-03-13
Thailand/NIH-2420/2020	EPI_ISL_434700	2020-03-13
Thailand/NIH-2472/2020	EPI_ISL_434701	2020-03-14
Thailand/NIH-2488/2020	EPI_ISL_434702	2020-03-14
Thailand/NIH-2489/2020	EPI_ISL_434703	2020-03-14
Thailand/NIH-2492/2020	EPI_ISL_430841	2020-03-17
Thailand/NIH-2492-2/2020	EPI_ISL_434704	2020-03-17
Thailand/NIH-2720/2020	EPI_ISL_434705	2020-03-16
Thailand/NIH-2721/2020	EPI_ISL_434706	2020-03-16
Thailand/NIH-2980/2020	EPI_ISL_434707	2020-03-18
Thailand/NIH-2982/2020	EPI_ISL_430842	2020-03-18
Thailand/NIH-2982-2/2020	EPI_ISL_434708	2020-03-18
Thailand/NIH-363/2020	EPI_ISL_434696	2020-01-24
Thailand/NIH-59/2020	EPI_ISL_434693	2020-01-16
Thailand/TNIC-1889/2020	EPI_ISL_430837	2020-03-03