Molecular characterization of methicillin-resistant
*Staphylococcus aureus* genotype ST764-SCCmec type II in Thailand

Sumalee Kondo1, Pimonwan Phokhaphan2, Sissades Tongsima2, Chumpol Ngamphiw2, Worawat Phonpirithawong2, Wuthiwat Ruangchai3, Areeya Disratthakit4, Pholawat Tingpej5, Surakameth Mahasirimongkol6, Aroonlug Lulitanond5, Anucha Apisarnthanarak1 & Prasit Palittapongarnpim3

Methicillin-resistant *Staphylococcus aureus* (MRSA) is a significant causative agent of hospital-acquired infections. We characterized MRSA isolated from August 2012 to July 2015 from Thammasat University Hospital. Genotypic characterization of MRSA SCCmec type II and III isolates were scrutinized by whole genome sequencing (WGS). The WGS data revealed that the MRSA SCCmec type II isolates belonged to ST764 previously reported mainly in Japan. All of tested isolates contained ACME Type II, SaPIn2, SaPIn3, seb, interrupted SA1320, and had a virulence gene profile similar to Japan MRSA ST764. Rigorous surveillance of MRSA strains is imperative in Thailand to arrest its potential spread.

Methicillin-resistant *Staphylococcus aureus* (MRSA) is a major cause of hospital-acquired infection (HA-MRSA) worldwide. There have been increasing reports of community-acquired MRSA (CA-MRSA)1 and livestock-associated MRSA (LA-MRSA)2. The resistance is generally conferred by the acquisition of a SCCmec element, which carries mecA, a gene encoding for PBP-2′, a penicillin-binding protein that is not inhibited by most beta-lactam antibiotics. The SCCmec element may also carry resistance genes to other classes of antibiotics such as aminoglycosides and glycopeptide antibiotics3. Many MRSA strains have become global multidrug resistant. For instance, MRSA strains isolated from a tertiary hospital in Malaysia showed co-resistance of tetracycline, ciprofloxacin and erythromycin4. MRSA-infected patients have a higher mortality rate than those infected with susceptible strains. Treatment of MRSA infection relies primarily on vancomycin. Although *S. aureus* isolates with complete resistance to vancomycin (VRSA) are still not common, milder vancomycin resistance phenotypes may cause problems in treatment. These include vancomycin-intermediate resistant *S. aureus* (VISA), heterogeneous vancomycin-resistant *S. aureus* (hVISA), and even the strains with reduced susceptibility to vancomycin at the MIC level of 1.5–2.0 µg/ml (vancomycin MIC creep)5–7. Strains with decreased susceptibility to vancomycin were reported to be a cause of treatment failure8. We selected the MRSA isolates by various methods. Despite the MRSA ST239 SCCmec type III genotype being commonly reported in Thailand, 41% of MRSA in Thammasat University Hospital (TUH)9 was SCCmec type II MRSA (unpublished data), which differed from other tertiary hospitals in Thailand10,11.

In this study, five isolates of HA-MRSA SCCmec type II and two isolates of SCCmec type III were whole-genome sequenced. All of the MRSA SCCmec type II isolates belonged to an unexpected MLST type, ST764 which had been reported mostly in Japan. The MRSA ST764 strains frequently contained ACME type II12. This strain was first discovered in a community setting in Japan in 2009 as a variant of ST5 genotype with SCCmec type II/spa2/seb2/PVL13. The ST764 genotype was later reported as a common genotype in many settings in Japan, including tertiary hospitals14,15, outpatient departments16, long-term care facilities17 and communities18, and this genotype has been increasing19.

---

1Faculty of Medicine, Thammasat University, Pathum Thani 12120, Thailand. 2National Biobank of Thailand (NBT), National Science and Technology Development Agency, Pathum Thani 12120, Thailand. 3Porntchai Matangkasombut Center for Microbial Genomics, Mahidol University, Bangkok 10400, Thailand. 4Medical Life Science Institute, Ministry of Public Health, Nonthaburi 11000, Thailand. 5Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kaen 40002, Thailand. *email: flower9great@yahoo.com; prasit.pal@mahidol.ac.th
This study reports, for the first time, WGS data of MRSA ST764 SCC\textit{mec} type II in Thailand, using Illumina NextSeq 500 platform. All isolates here were from severe hospital-acquired infection patients with high mortality. Although these MRSA strains were still susceptible to vancomycin, their MICs of vancomycin were mostly at the high end of the susceptible range. The discovery in Thailand of the ST764 genotype, a well-known Japan CA-MRSA, indicates the increase in the geographic range of the bacteria beyond East Asia. Since the ST764 genotype can be community-acquired, this emerging clone poses a significant threat not only to TUH but also to the other hospitals in the region. The presence of this MRSA strain warrants vigorous surveillance in hospitals and immediate investigation of community transmissions of the strains.

**Results**

Five isolates of HA-MRSA SCC\textit{mec} type II [SATU130 (SRA: SRR16971216), SATU131 (SRA: SRR16971215), SATU132 (SRA: SRR16971214), SATU134 (SRA: SRR16971213) and SATU135 (SRA: SRR16971212)] and two isolates of HA-MRSA SCC\textit{mec} type III [SATU133 (SRA: SRR17033804) and SATU135 (SRA: SRR17033804) were whole-genome sequenced. All isolates were from patients with invasive diseases, including pneumonia, septicemia, osteomyelitis and endocarditis as described in Table 1. The majority of the patients were over 60 years old. All five SCC\textit{mec} type II isolates belonged to the ST764 genotype while the SCC\textit{mec} type III isolates belonged to ST22 and ST239. None of the patients was HIV seropositive, but four patients with ST764 infection had diabetes with chronic kidney diseases. Only one patient with MRSA ST764 infection survived.

All isolates were resistant to cefpirome and clindamycin yet were susceptible to vancomycin, trimethoprim-sulfamethoxazole, linezolid, mupirocin, fusidic acid, TEC teicoplanin, Lzd linezolid.

### Table 1. Profiles of patients whose MRSA isolates were studied by WGS. WGS whole genome sequencing, ST sequence type, NA not available, DM diabetes mellitus, CKD chronic kidney disease.

| Strain code | ST  | Date of isolation | Age | Sex | Diagnosis                  | Sample type | Days to negative culture | Underlying diseases | Treatment outcomes |
|-------------|-----|------------------|-----|-----|----------------------------|-------------|-------------------------|--------------------|--------------------|
| SATU130     | ST764| 10/10/2013       | 84  | M   | Septicemia                 | Blood       | 8                       | Yes                | Death              |
| SATU131     | ST764| 5/11/2013        | 78  | M   | Pneumonia                  | Urine       | 8                       | No                 | No                 |
| SATU132     | ST764| 9/4/2013         | 61  | M   | Pneumonia                  | Sputum      | 19                      | Yes                | Success            |
| SATU133     | ST239| 1/2/2013         | 18  | M   | Osteomyelitis of left femur| Bone        | 17                      | No                 | No                 |
| SATU134     | ST764| 15/1/2013        | 80  | F   | Pneumonia                  | Sputum      | 13                      | Yes                | Death              |
| SATU135     | ST22 | 2/1/2015         | 78  | M   | Pneumonia                  | Sputum      | NA                      | No                 | Success            |
| SATU136     | ST764| 29/1/2014        | 80  | F   | Endocarditis and septicemia| Blood       | 8                       | Yes                | Death              |

### Table 2. Minimal inhibitory concentrations of antibiotics against the MRSA isolates. VAN vancomycin, CPR cefpirome, CLI clindamycin, TMP trimethoprim, SXT trimethoprim-sulfamethoxazole, RIF rifampicin, MUP mupirocin, FUS fusidic acid, TEC teicoplanin, Lzd linezolid.

| Strain code | VAN | CPR | CLI | TMP | SXT | RIF | MUP | FUS | TEC | LZD |
|-------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| SATU130     | 1.5 | >256| >256| 0.5 | 0.012| 0.012| 0.094| 0.125| 0.75| 2   |
| SATU131     | 2   | >256| >256| 0.38| 0.012| 0.004| 0.25 | 0.094| 1   | 1   |
| SATU132     | 0.75| >256| >256| 0.25| 0.008| 0.008| 0.094| 0.25 | 1   | 1.5 |
| SATU133     | 2   | >256| >256| >32 | >0.064| 0.094| 1   | 0.75 |     |     |
| SATU134     | 1.5 | >256| >256| 0.75| 0.016| 0.008| <0.064| 0.19 | 0.38| 1   |
| SATU135     | 0.38| >256| >256| 0.38| 0.032| 0.008| 0.19 | 0.19 | 0.5 | 1.5 |
| SATU136     | 2   | >256| >256| 0.38| 0.012| 0.008| 0.19 | 0.125| 0.5 | 1.5 |

Characterization of single nucleotide variants (SNVs). More than 90% of reads were uniquely mapped to \textit{S. aureus} N315 (ASM964v1, GenBank accession: BA000018.3), an ST5 isolate. The mean coverage of the sequence was 93% with 72X average depth. A total of 50,021 SNVs were identified based on alignment to the reference genome of \textit{S. aureus} N315. The number of SNVs identified among ST764 isolates ranged from 237 to 261. The average pairwise SNV distance among ST764 isolates was 49.8.

The maximum likelihood phylogenetic tree of five ST764 isolates, with N315 as the outgroup, was shown in Fig. 1. There were two pairs, SATU130–SATU131 and SATU132–SATU136, with the same SNV distances of 11 (Table 3), suggesting epidemiological linkages between each pair\(^1\). SATU130 and SATU131 were isolated from the same male surgical ward, 26 days apart, and therefore probably belonged to the same clones. Isolates SATU132
and SATU136 were from male and female medicine wards, 9 months apart. Moreover, isolate SATU136 did not contain seb, while SATU132 did. The epidemiological link between the two isolates was therefore uncertain.

All five ST764 isolates shared three deletions (g.ASM964v1:823521_823580del, g.ASM964v1:2068598_2068612del, g.ASM964v1:2793533_2793544del) and four insertions (g.ASM964v1:8342_8343ins[unknown], g.ASM964v1:941411_941412ins[unknown], g.ASM964v1:1104595_1104596ins[unknown], g.ASM964v1:1527689_1527690ins[unknown]), which were not present in the ST22 and ST239 isolates. There were some indels found in some but not all ST764 isolates. A deletion (g.ASM964v1:1608977_1609107del) was shared by SATU132 and SATU136 while a deletion (g.ASM964v1:383663_383683del) and an insertion (g.ASM964v1:676384_676385ins[unknown]) could be found only among isolates SATU130, SATU131, and SATU134, as shown in Fig. 1 and Supplementary Tables S2 and S3.

Characterization of SCCmec-ACME II’ composite island of ST764. We investigated the presence of various genes and genetic elements previously described in ACME-SCCmec composite island (ACME-SCCmec CI)12. All ST764 isolates here contained arcA–D, but not opp3; thus the ACME element was classified as ACME II’. They did not harbor any SNV in arcA–D as compared with N315, which was similar to those Japan ST764 isolates. It should be noted that C4913T mutation in arcB was found in an ST764 isolate in Japan22. ACME-SCCmec CI was previously classified into five types. The five ST764 isolates harbored cJR1 DNA segment but carried neither pUB110 nor ccrC. Thus, these isolates belonged to type A323. The ACME-SCCmec CI type A is the shortest form of the complex and probably puts less burden on the fitness cost for the bacteria.

The presence of genetic elements specific to ST764. In order to further evaluate similarity between the ST764 isolates in this study and the isolates reported from Japan, we identified other genetic elements previously described for ST764 in Japan apart from the ACME-SCCmec CI. The five ST764 isolates contained SaPIn2 and SaPIn3 pathogenic islands. The former exotoxin island, SaPIn2, contained set6 to set15. The latter enterotoxin island contained seg, sei, sem, sen, and seo. These isolates did not have SAPIn1, evidenced by the absence of tst, sea, sec, sel or sek. Our ST764 isolates did not contain sed, see, seq,
Discussion

HA-MRSA is a global problem, especially in large tertiary hospitals. Previous WGS studies in Thailand focused mainly on ST239 SCCmec type III, which was the most common MLST type4–6. It is, therefore, intriguing to find that SCCmec type II MRSA was common in TUH, unlike in other tertiary hospitals in Thailand10,11. Moreover, all belonged to ST764-MRSA-II and were from patients with severe hospital-acquired invasive diseases. The apparently high incidence remains to be confirmed.

ST764-MRSA-II, is a recently recognized genotype, reported as a common MRSA genotype in many settings in Japan and might be spreading16. Such hospital settings encompass outpatient departments and healthcare professionals17. This spreading may be attributed to the presence of ACME type II, which may provide the bacteria a better ability to colonize on skin and mucous membrane which could consequently lead to better transmission. It has recently been reported as a common genotype in Shanghai18 and Eastern China19 but not yet in Southern China20.

ST764 is a hybrid variant of the globally disseminated ST5 lineage of HA-MRSA, carrying a few virulence traits similar to community-acquired MRSA (CA-MRSA). ST764 has evolved via multiple steps, including the acquisition of ACME arcA and the staphylococcal enterotoxin B gene (seb) in SaPIn5421. Here we reported, for the first time, the WGS of the five MRSA SCCmec type II isolates identified from 2013 to 2015 and demonstrated that all ST764 isolates in this study carried ACME-SCCmec CI type A3 similar to some ST764 isolates in Japan23. The ST764 strains from Thailand and those from Japan also shared other genetic characteristics, such as the lack of pvl, tst, sec, sel and sep with the presence of SaPIn2, SaPIn3, seb, tet(M)22. This implies that the ST764-MRSA-II clone in Thailand was similar to and probably originated from the clone in Japan.

The genetic similarity of ST764-MRSA-II in Thailand to ST764-MRSA-II in Japan suggested a possibility that the ST764-MRSA-II strain was brought to Thailand by a healthy carrier, and later was transmitted into the hospital settings. There is also a possibility that the ST764-MRSA-II strain may continue to spread wider by colonization of healthy carriers.

Analysis of whole genome sequencing data suggested potential transmission between a pair of isolates. The isolates had a SNP-distance of 11 and were isolated from the same ward within a month. This finding of MRSA genotype ST764-SCCmec type II is the first HA occurrence reported in Southeast Asia. Two isolates of CA ST764-SCCmec type II from Thai pig farm workers were also reported23. The potential danger of spreading necessitates rigorous surveillance of emerging MRSA strains.

Methods

Bacterial strains. One-hundred and one clinical isolates of MRSA, confirmed by the detection of mecA, at TUH from August 2012 to July 201524, were characterized for SCCmec types, pulsed-field gel electrophoresis (PFGE) patterns, virulence genes patterns, and antibiotic resistance phenotypes. There were 41 samples belonging to SCCmec type II that revealed either PFGE pattern A or P. Five isolates with both A- and P-PFGE patterns were randomly selected for sequencing, together with two isolates belonging to SCCmec type III. Bacterial collection, demographic data, clinical and microbiological data retrieval from Microbiology laboratory, Thammasat University Hospital (TUH) were carried out in accordance with relevant guidelines and regulations after approval of the Human Research Ethics Committee of Thammasat University (approval no. MTU-EC-DS-1-193/63 and MTU-EC-DS-4-197/63) and TUH.

Microbiological and clinical data. Demographic data, clinical and microbiological data were retrieved from MRSA-infected patients' record files from the TUH database. Examples of patients' clinical data include lengths of stay, clinical outcomes and underlying diseases including diabetes mellitus (DM) and chronic kidney diseases (CKD). Successful treatment outcomes referred to events when the patients were cured or improved and MRSA was eradicated. Outcomes of persistent-MRSA, recurrent-MRSA or death due to MRSA were regarded as failure of treatments. The patients who were diagnosed with MRSA infection at an outpatient clinic or within 48 h of hospitalization without history of exposure to any healthcare facilities during the previous two months were defined as community-acquired MRSA infection (CA-MRSA). Hospital-acquired MRSA infection (HA-MRSA) referred to those patients who became infected after 48 h of hospitalization.

Antimicrobial susceptibility testing. Minimum inhibitory concentration (MIC) of vancomycin (VAN) was carried out by both broth microdilution method25,26 and E-test for determining the high end MICs. Other antibiotics including cefpirome (CPR), clindamycin (CLI), trimethoprim (TMP), trimethoprim-sulfamethoxazole (SXT), rifampicin (RIF), mupirocin (MUP), fusidic acid (FUS), teicoplanin (TEC) and linezolid (LZD) were tested by E-test (Liofilchem, Italy).
PCR detection of *S. aureus* mecA and pvl genes. Total DNA was purified from each overnight *S. aureus* culture using Genomic DNA Extraction Mini Kit (RBC Bioscience, New Taipei City, Taiwan). Amplification of mecA (5’-TCC AGA TTA CAA CTT CAC CAG G-3’ and 5’-CCA CTT CAC ATC TTG TAA CGG-3’) and pvl (5’-ATC ATG AGG TAA AAT GTC TGG AGA TGA TCC A-3’ and 5’-ACA TCA AST GTA TGG GAT AGC AAA AGC-3’) (Integrated DNA Technologies, Singapore) was carried out using the following conditions. The PCR mixture consisted of 50 μl of 10X PCR buffer, 50 mM MgCl2, 10 mM dNTPs, 100 μM specific primer pair and 1.25 U Taq polymerase (RBC Bioscience). Thermocycling was conducted in a MyCycler Thermal Cycler (Bio-Rad, Hercules, CA) as follows: 94°C for 2 min; followed by 30 cycles of 94°C for 30 s, 51°C (for mecA) or 56°C (for pvl) for 30 s, and 72°C for 1 min; with a final elongation step at 72°C for 5 min. Amplicons (162 bp and 433 bp of mecA and pvl, respectively) were analyzed by 1% agarose gel electrophoresis, stained with GelStar™ Nucleic Acid Gel Stain (Lonza Rockland, Rockland, ME) and visualized under UV light. *S. aureus* N315 and KNU-MS14 strains were used as mecA- and pvl-positive control, respectively.

**SCCmec typing.** SCCmec typing was carried out using a multiplex PCR assay. Specific primers for SCCmec types I, II, III, and IVc were included (Supplementary Table S4). Positive controls, SCCmec type I (613 bp); SCCmec type II (398 bp) and SCCmec type III (280 bp), were added.

**Whole-genome sequencing.** DNA purity was assessed using a spectrophotometer, NanoDrop™ 2000 (Thermo Fisher, USA) and DNA concentration was determined by measuring fluorescence in a fluorometer, Qubit™ 3 Fluorometer (Invitrogen™, USA). DNA integrity was visualized by running the isolated samples in agarose gel electrophoresis to assess extent of DNA shearing.

A DNA input of 1 ng was used to prepare a 150-base read library, using the Nextera XT DNA Library Preparation Kit (Illumina, USA). The library preparation protocol includes tagmentation of genomic DNA, library amplification, library clean up, library normalization and library pooling. Library denaturation and dilution were performed, and then diluted PhiX control V3 was spiked into the libraries to obtain 1% PhiX as an external control. The prepared library pool was loaded into the reagent cartridge and then sequenced by 150 bp paired-end NextSeq 500 sequencing (Illumina, USA). Sequences of the five isolates of HA-MRSA SCCmec type II were submitted to SRA. The accession numbers of each isolate were listed as follows. SATU130 (SRA: SRR16971216), SATU131 (SRA: SRR16971215), SATU132 (SRA: SRR16971214), SATU134 (SRA: SRR16971213) and SATU136 (SRA: SRR16971212) and two isolates of HA-MRSA SCCmec type III [SATU133 (SRA: SRR17033804) and SATU135 (SRA: SRR17033804)]

**Bioinformatic analyses.** The sequencing reads were quality filtered using Trimmomatic version 0.39 with the following parameters: LEADING: 3 TRAILING: 3 SLIDINGWINDOW: 4:20 MINLEN: 36. The quality-passed reads were aligned to the MRSA ST5 N315 reference genome (ASM964v1, GenBank accession: BA000018.3) using Burrows-Wheeler Aligner (BWA) version 0.7.17 with the BWA-MEM algorithm. Sequencing read duplication was handled by Picard Toolkit version 2.18.27 (available from http://broadinstitute.github.io/picard/). For each MRSA isolate, variant identification of single nucleotide variants (SNVs) and short insertion-deletions (Indels) was carried out using the HaplotypeCaller module of Genome Analysis Toolkit (GATK) version 4.1.4.19. Low quality variants that have read depth lower than 10 were discarded. Variant annotation was performed using SnpEff version 4.3.40. Structural variants (SVs) were identified by running the program Manta41 on an input BAM file (sorted and mapped reads) that has aligned to the reference sequence. Resulting SV events flagged with “IMPRECISE” were discarded. The remaining SV events were confirmed by visualizing them using integrative genomics viewer (IGV)42. De novo Genome assembly was done by SPAdes3.9 (available from http://cab.spbu.ru/software/spades/) (Supplementary Fig. S1).

**Multilocus sequencing typing (MLST).** MLST was done by identifying the presence of variants located in seven housekeeping genes, carbamate kinase (arcC), shikimate dehydrogenase (aroE), glyceraldehyde 3-phosphate dehydrogenase (gapd), guanylate kinase (gmk), phosphate acetyltransferase (pta), triosephosphate isomerase (tpi), and acetyl coenzyme A acetyltransferase (yqiL)43. Parts of the assembled contigs from each isolate were aligned with *Staphylococcus aureus* MLST primers (available from https://pubmlst.org/organisms/staphylococcus-aureus/primers). The allelic profiles of our isolates were matched with allele sequences present in the PubMLST database (available from https://pubmlst.org/)43 (Supplementary Fig. S1).

**Spa typing.** The spa sequence types were assigned using spaTyper1.0 (available from https://cge.cbs.dtu.dk/services/spatyper/)43.

**Phylogenetic tree construction.** To prepare an input sequence for building a phylogenetic tree, SNVs present on Mobile Genetic Element (MGE) were discarded and only homozygous and biallelic SNVs were concatenated into the SNP input sequence (in-house script). Multiple sequence alignment of five ST764 isolates and N315, representing an outgroup, were tested for recombination signals by using Recombination Detection Program v4.10 (RDP4)46 that includes seven well-known recombination checking routines, namely RDP, GENECONV, Chimera, MaxChi, Bootscan, Siscan, and 3Seq. If more than four utilities reported recombination regions, these regions would be ignored. These RDP-corrected sequences were used as the input for phylogenetic tree construction by the maximum likelihood (ML) method. PhyML version 3.0 was used to construct an ML-based phylogeny which can be visualized by the SeaView program47.48. That includes seven well-known recombination checking routines, namely RDP, GENECONV, Chimera, MaxChi, Bootscan, Siscan, and 3Seq. If
more than four utilities reported recombination regions, these regions would be removed. These RDP-corrected sequences were used as the input for phylogenetic tree construction by the maximum likelihood (ML) method. PhyML version 3.0 was used to construct an ML-based phylogeny which can be visualized by the SeaView program.

ACME/SCCmec type II. The structure of ACME/SCCmec type II was determined by identification of arcA, arcB, arcC, arcD, opp3 and ccrC using the sequence of MRSA USA300, (GenBank Accession: CP000255.1) as reference. Identification of PUB110 and cflR1 was done using N315 and NCTC10442 (SCCmec type I, GenBank Accession: UHC01000003.1) as reference, respectively.

Virulence gene identification. Genes present in three pathogenicity islands (SaPIn1, SaPIn2 and SaPIn3) were identified, which are SaPIn1: [stl, sea, sec, sel, sek], SaPIn2: [setf to set15] and SaPIn3: [seg, sei, sem, sen, seo]. The presence of other enterotoxins (seb (AB630021.1), sed, see, sep, seq), exfoliative toxins (eta, etb), leukocidin (lukE-lukD), panton-valentine leukocidin (pvl), hemolysins (hla, hld, hlb, hlgA, hlgB and hlgC), and adhesin genes (icaA, sdrC-E, ebpS, cflA, cflB, cna, bbp, femA, fnbA and fnbB), were also verified. The Sa1320 gene was checked for the presence of an insertion in the middle of the genome by blasting the assembled contigs with SA1320 sequence of N315.

The following sequences were used as reference: MRSA N315/SCCmec II/ST5/ACMEII (GenBank accession: BA000018.3)60, MRSA NN54/SCCmec II/ST764 (GenBank accession: BAFl00000000)51, MRSA PT2/SCCmec II/ST5/seeb (GenBank accession: AB630021.1)51, MRSA H0 5096_0412/SCCmec IV/ST22 (GenBank accession: HE681097)51, JKD6008/SCCmec III/ST239 (GenBank accession: CP002120), and TW20/SCCmec III/ST239 (GenBank accession: FN433596)51.

The in-house script for converting SNV data in VCF file format to SNV concatenated sequences in fasta file format was deposited at https://github.com/voravich/1_TB_platforn/blob/5ef4ed3dd93dd25df418a6fa632435663fc6f31/tb_convert_vcf_to_fasta.backbone.py.

References
1. Otter, J. A. & French, G. L. Community-associated methicillin-resistant Staphylococcus aureus strains as a cause of healthcare-associated infection. J. Hosp. Infect. 79, 189–193 (2011).
2. Petinaki, E. & Spiliopoulou, I. Methicillin-resistant Staphylococcus aureus colonization and infection risks from companion animals: Current perspectives. Vet. Med. (Auckl.) 6, 373–382 (2015).
3. Shindia, A. A., Ragab, F. A. & Nasrat, H. M. Emergence of high-level vancomycin-resistant Staphylococcus aureus isolates in Canada: Results from the Canadian Nosocomial Infection Surveillance Program, 1995–2006. Antimicrob. Agents Chemother. 54, 945–949 (2010).
4. Pitz, A. M. et al. Vancomycin susceptibility trends and prevalence of heterogeneous vancomycin-intermediate Staphylococcus aureus isolates in Canada: Results from the Canadian Nosocomial Infection Surveillance Program, 1995–2006. Antimicrob. Agents Chemother. 54, 945–949 (2010).
5. Pitt, A. M. et al. Vancomycin susceptibility trends and prevalence of heterogeneous vancomycin-intermediate Staphylococcus aureus isolates in Canada: Results from the Canadian Nosocomial Infection Surveillance Program, 1995–2006. Antimicrob. Agents Chemother. 54, 945–949 (2010).
6. Fitz, A. M. et al. Vancomycin susceptibility trends and prevalence of heterogeneous vancomycin-intermediate Staphylococcus aureus isolates in Canada: Results from the Canadian Nosocomial Infection Surveillance Program, 1995–2006. Antimicrob. Agents Chemother. 54, 945–949 (2010).
7. Shindia, A. A., Ragab, F. A. & Nasrat, H. M. Emergence of high-level vancomycin-resistant Staphylococcus aureus isolates in Canada: Results from the Canadian Nosocomial Infection Surveillance Program, 1995–2006. Antimicrob. Agents Chemother. 54, 945–949 (2010).
8. Howden, B. P., Davies, J. K., Johnson, P. D., Stinear, T. P. & Grayson, M. L. Reduced vancomycin susceptibility in Staphylococcus aureus, including vancomycin-intermediate and heterogeneous vancomycin-intermediate strains: Resistance mechanisms, laboratory detection, and clinical implications. Clin. Microbiol. Rev. 23, 99–130 (2010).
9. Phokhaphan, P., Tingpej, P., Apiaratanarak, A. & Kondo, S. Prevalence and antibiotic susceptibility of methicillin-resistant Staphylococcus aureus, collected at Thammasat University Hospital, Thailand, August 2012–July 2015. Southeast Asian J. Trop. Med. Public Health 48, 351–359 (2017).
10. Jaitphysathong, T. et al. MRSA carriage in a tertiary governmental hospital in Thailand: Emphasis on prevalence and molecular epidemiology. Eur. J. Clin. Microbiol. Infect. Dis. 29, 977–985 (2010).
11. Lilithanond, A. et al. Preliminary report of SCCmec-types and antimicrobial susceptibilities of methicillin-resistant Staphylococcus aureus isolates from a university hospital in Thailand. Southeast Asian J. Trop. Med. Public Health 41, 920–927 (2010).
12. Takano, T. et al. A new local variant (ST764) of the globally disseminated ST5 lineage of hospital-associated methicillin-resistant Staphylococcus aureus (MRSA) carrying the virulence determinants of community-associated MRSA. Antimicrob. Agents Chemother. 57, 1589–1595 (2013).
13. Ozaki, K. et al. Genotypes, intrafamilial transmission, and virulence potential of nasal methicillin-resistant Staphylococcus aureus in children from the community. J. Infect. Chemother. 15, 84–91 (2009).
14. Aung, M. S. et al. Clonal diversity and genetic characteristics of methicillin-resistant Staphylococcus aureus isolates from a Tertiary Care Hospital in Japan. Microb. Drug Resist. 23, 1164–1175 (2017).
15. Nakaminami, H. et al. Characterization of methicillin-resistant Staphylococcus aureus isolated from Tertiary Care Hospitals in Tokyo, Japan. J. Infect. Chemother. 20, 512–515 (2014).
16. Aung, M. S. et al. Molecular characterization of methicillin-resistant Staphylococcus aureus from outpatients in Northern Japan: Increasing tendency of ST5/ST764 MRSA-4a with arginine catabolic mobile element. Microb. Drug Resist. 23, 616–623 (2017).
17. Kawanura, K. et al. Spread of seb-positive methicillin-resistant Staphylococcus aureus SCCmec Type II-ST764 among elderly Japanese in non acute care settings. Microb. Drug Resist. 25, 915–924 (2019).
18. Otsuka, T. et al. Molecular Epidemiology of methicillin-resistant Staphylococcus aureus isolated from children in a community with low antimicrobial pressure in Japan. Jpn. J. Infect. Dis. 65, 483–488 (2012).
19. Goudarzi, M. et al. Molecular characteristics of methicillin-resistant Staphylococcus aureus (MRSA) strains isolated from patients with bacteremia based on MLST, SCCmec, spa, and agr locus types analysis. Microb. Pathog. 104, 328–335 (2017).
20. Monecke, S. et al. Molecular typing of ST239-MRSA-III from diverse geographic locations and the evolution of the SCCmec III element during its intercontinental spread. Front. Microbiol. 9, 1436 (2018).

21. Schurch, A. C., Arredondo-Alonso, S., Willems, R. J. L. & Goering, R. V. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene–gene-based approaches. Clin. Microbiol. Infect. 24, 350–354 (2018).

22. Kawaguchiya, M. et al. Genetic diversity of emerging Panton-Valentine leukocidine/arginine catabolic mobile element (ACME)-positive ST8 SCCmec-IVA methicillin-resistant Staphylococcus aureus (MRSA) strains and ACME-positive CC5 (ST5/ST764) MRSA strains in Northern Japan. J. Med. Microbiol. 62, 1852–1863 (2013).

23. Urushibara, N. et al. Novel structures and temporal changes of arginine catabolic mobile elements in methicillin-resistant Staphylococcus aureus genotypes ST5-MRSA-II and ST764-MRSA-II in Japan. Antimicrob. Agents Chemother. 60, 3119–3122 (2016).

24. Hall, M. D., et al. Improved characterization of MRSA transmission using within-host bacterial sequence diversity. Elife 8, e46402 https://doi.org/10.7554/eLife.46402 (2019).

25. Harris, S. R. et al. Evolution of MRSA during hospital transmission and intercontinental spread. Science 327, 469–474 (2010).

26. Tong, S. Y. et al. Genome sequencing defines phylogeny and spread of methicillin-resistant Staphylococcus aureus in a high transmission setting. Genome Res. 25, 111–118 (2015).

27. Tamasaki, F. et al. Prevalence and characteristics of methicillin-resistant Staphylococcus aureus colonization among healthcare professionals in a university hospital in Japan. J. Gen. Fam. Med. 20, 190–192 (2019).

28. Gu, F. et al. Antimicrobial resistance and molecular epidemiology of Staphylococcus aureus causing bloodstream infections at Ruijin Hospital in Shanghai from 2013 to 2018. Sci. Rep. 10(1), 6019 (2020).

29. Chen, W. K. et al. Prevalence and characterization of methicillin-resistant Staphylococcus aureus with mupirocin, fusidic acid and/or retapamulin resistance. BMC Microbiol. 20, 183 (2020).

30. Liu, F. et al. A comparative genomic analysis between methicillin-resistant Staphylococcus aureus strains of hospital acquired and community infections in Yunnan province of China. BMC Infect. Dis. 20, 137 (2020).

31. Sinlapasorn, S. et al. SCCmec IX in meticillin-resistant Staphylococcus aureus and meticillin-resistant coagulase-negative staphylococci from pigs and workers at pig farms in Khon Kaen, Thailand. J. Med. Microbiol. 64, 1087–1093 (2015).

32. CLSI. Performance Standards for Antimicrobial Susceptibility Testing, 30th edn. CLSI supplement M100. (Clinical and Laboratory Standards Institute, 2020).

33. CLSI. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard, 10th edn. CLSI document M07-A10. (Clinical and Laboratory Standards Institute, 2015).

34. Oliveira, D. C. & de Lencastre, H. Multiplex PCR strategy for rapid identification of structural types and variants of the mec element in methicillin-resistant Staphylococcus aureus. Antimicrob. Agents Chemother. 46, 2155–2161 (2002).

35. Lina, G. et al. Involvement of Panton-Valentine leukocidin-producing Staphylococcus aureus in primary skin infections and pneumonia. Clin. Infect. Dis. 29, 1128–1132 (1999).

36. Zhang, K., McClure, J. A., Elsayed, S., Louie, T. & Conly, J. M. Novel multiplex PCR assay for characterization and concomitant subtyping of staphylococcal cassette chromosome mec types I to V in meticillin-resistant Staphylococcus aureus. J. Clin. Microbiol. 43, 3026–3033 (2005).

37. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).

38. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26, 589–595 (2010).

39. Poplin, R., et al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv 201178 (2018).

40. Cingolani, P. et al. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. Front. Genet. 3, 35 (2012).

41. Chen, X. et al. Manta: Rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics 32, 1220–1222 (2016).

42. Robinson, J. T. et al. Integrative genomics viewer. Nat. Biotechnol. 29, 24–26 (2011).

43. Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J. & Spratt, B. G. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus. J. Clin. Microbiol. 38, 1008–1015 (2000).

44. Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res. 3, 124 (2018).

45. Bartels, M. D. et al. Comparing whole-genome sequencing with Sanger sequencing for spa typing of methicillin-resistant Staphylococcus aureus. J. Clin. Microbiol. 52, 4305–4308 (2014).

46. Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 1(1), vevo003 (2015).

47. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224 (2010).

48. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321 (2010).

49. Tamura, K. et al. MEGAS: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739 (2011).

50. Kuroda, M., et al. Whole genome sequencing of meticillin-resistant Staphylococcus aureus. Lancet 357, 1225–1240 (2001).

51. Iwao, Y. et al. Isolation and molecular characterization of meticillin-resistant Staphylococcus aureus from public transport. Microbiol. Immunol. 56, 76–82 (2012).

52. Koser, C. U. et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N. Engl. J. Med. 366, 2267–2275 (2012).

53. Holden, M. T. et al. Genome sequence of a recently emerged, highly transmissible, multi-antibiotic- and antiseptic-resistant variant of methicillin-resistant Staphylococcus aureus, sequence type 239 (TW). J. Bacteriol. 192, 888–892 (2010).

Acknowledgements

We owe thanks to Microbiology Laboratory, Thammasat University Hospital and Faculty of Medicine, Thammasat University, Fathum Thani, Thailand for strains collection and facility support, respectively. We also thank to Lalita Narachasima, Rujipat Wasiththakasem, Wanwisa Chareanchim, Nirinya Sudtachat and Maneerat Rak-sayot for technical support.

Author contributions

S.K. designed and directed the research project, P.Ph. performed experiments and clinical data analysis under supervision of S.K., A.A. and P.T., S.K., P.Ph., P.T., A.A. and P.Ps. interpreted results, C.N., W.P., W.R., P.Ph. and S.T. carried out bioinformatics analyses and prepared figures, A.D. and S.M. performed sequencing experiment.
A.L. provided bacterial strains for SCCmec typing, P.Pa., S.K. and P.Ph. wrote the main manuscript with input from all authors. All authors reviewed the manuscript.

Funding
This research was supported by the National Research Council of Thailand (94248/2557A10602093; 111854/2558A10602034), the Faculty of Medicine, Thammasat University, Thailand (2-07-2563), the National Science and Technology Development Agency (JRA-CO-2564-13515) and the Research Group in multidrug resistant bacteria and the antimicrobial herbal extracts, Faculty of Medicine, Thammasat University (RG2564).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-05898-1.

Correspondence and requests for materials should be addressed to S.K. or P.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022