Tropical Medicine and Health

RESEARCH

Open Access

Genomic analysis of the genetic background underlying *Streptococcus pneumoniae* beta-lactam nonsusceptibility in central Vietnam: increased beta-lactam nonsusceptibility and dynamics of the *pbp2x* gene

Hiroshi Fujii[1], Mitsuhiko Sato[2,11], Hien Anh Thi Nguyen[3], Huong Thi Thu Vu[3], Satoshi Kakiuchi[4,12], Bhim Gopal Dhoubhadel[5], Shota Nakamura[6], Daisuke Motooka[6], Yoshitoshi Ogura[2,13], Satoshi Nakano[7], Christopher M. Parry[5,14], Konosuke Morimoto[1,8], Lay-Myint Yoshida[5,9], Dang Duc Anh[3], Tetsuya Hayashi[2], Tetsuya Iida[6,10] and Koya Ariyoshi[4,5*]

Abstract

Background We previously reported alarmingly high carriage rates of *Streptococcus pneumoniae* (SP) serotype 19F and serogroup 6 isolates, which were not susceptible to multiple beta-lactams among children under five years of age in Vietnam. Multilocus sequence typing analysis revealed the predominance of two major lineages, ST320 and ST13223, among serotype 19F and serogroup 6 isolates, respectively. Investigating the association between nonsusceptible genotypes and clinical outcomes could help optimize patient care or lead to the development of new diagnostic tests.

Methods We performed WGS on SP isolates randomly selected from the two major lineages and their related strains. FASTQ quality control and de novo assembly were performed using CLC Genomics Workbench ver. 7.5.1. Draft genome sequences were annotated using DFAST (DDBJ Fast Annotation and Submission Tool), which revealed the serogroups/serotypes and the sequences of the three major penicillin-binding protein genes and the sequence types. Draft sequences were aligned using MUMmer ver. 3.23, and putative recombination events and phylogenetic relationships excluding recombination regions were identified using Gubbins ver. 2.4.1. Finally, the association between a detected nonsusceptible genotype and the duration of hospital stay was evaluated in patients with acute respiratory infection.

Results WGS analysis (serotype 19F/ST320, n = 22; serogroup 6/ST13223, n = 13; and isolates closely related to ST13223, n = 4) revealed substantial differences in genomic diversity and antimicrobial susceptibility between serogroup 6/ST13223 and serotype 19F/ST320 isolates, particularly the recombination-prone nature of serogroup 6/

*Correspondence:
Koya Ariyoshi
kari@nagasaki-u.ac.jp
Full list of author information is available at the end of the article

ST13223. Among the 23 recombination events observed in serogroup 6/ST13223, only those spanning the *pbp2x* region (15.5 kb and 6.4 kb) were associated with high MICs for multiple beta-lactams. A subset of ST13223 isolates and all ST320 isolates carried the identical *pbp2x* allele 16, which was significantly associated with a lack of susceptibility to the combination of penicillin, cefotaxime, and meropenem ($p < 0.0001$; odds ratio 11.5; 95% confidence interval [CI] 3.35–39.3). No significant association was demonstrated between the presence of this *pbp2x* allele and prolonged hospitalization ($p = 0.6123$).

Conclusions We revealed that the widespread nonsusceptibility to multiple beta-lactams among SP isolates circulating in central Vietnam was primarily driven by the dynamics of the *pbp2x* gene. However, the nonsusceptible *pbp2x* allele had little effect on clinical outcome.

Keywords *Streptococcus pneumoniae*, Whole-genome sequencing, Molecular epidemiology, Pneumococcal conjugate vaccine, Antimicrobial nonsusceptibility, Penicillin-binding protein

Background

Streptococcus pneumoniae (SP) is a major pathogen that causes multiple infectious diseases, including pneumonia, bacteraemia, meningitis, and otitis media, thus resulting in significant health burdens in both high- and low- to middle-income countries [1, 2]. Moreover, the increasing antimicrobial nonsusceptibility of SP is a pressing global concern; since the 1990 s, the prevalence of pneumococcal strains that are not susceptible to beta-lactam antimicrobials has increased rapidly worldwide [3, 4]. Several epidemiological studies have also shown that antimicrobial nonsusceptibility was quite common among SP isolates in Vietnamese children before the introduction of the pneumococcal conjugate vaccine (PCV) and that it was primarily driven by the emergence of multidrug-resistant, globally circulating clones (Spain^{23F}−1 and Taiwan^{19F}−14) [5–8]. Serotypes/serogroups 19, 23, 14, and 6 were predominant in Vietnam before the introduction of the PCV [5, 7].

It was suggested that beta-lactam nonsusceptibility among SP isolates was associated with specific serotypes/serogroups. In our previous community-based study conducted in Nha Trang, located in central Vietnam, we reported alarmingly high carriage rates of SP serotype 19 F and serogroup 6, which were not susceptible to multiple beta-lactams, including broad-spectrum antimicrobials. These strains were found both in children with acute respiratory infections (ARIs) and in healthy children [9]. Our study revealed that 18.0%, 25.8%, and 75.6% of SP isolates were not susceptible to penicillin (PEN) (nonmeningitis breakpoint: ≥ 4 μg/ml), cefotaxime (CTX) (≥ 2 μg/ml), and meropenem (MEM) (≥ 0.5 μg/ml), respectively. This nonsusceptibility to multiple beta-lactams was significantly associated with serotype 19 F and daycare attendance but not the presence of an ARI, age, or prior antimicrobial use. Fortunately, the majority (90.0%) of beta-lactam-nonsusceptible isolates were PCV13 vaccine serotypes, suggesting that vaccine introduction would significantly increase drug susceptibility. Our preliminary analysis using multilocus sequence typing (MLST) revealed the predominance of two major lineages, ST320 and ST13223, among serotype 19 F and serogroup 6 isolates, respectively. However, the detailed mechanisms underlying beta-lactam nonsusceptibility in these SP lineages remain unclear.

Koch et al. recently examined the relationships among resistance genotypes, phenotypes, and clinical outcomes in hospitalized Israeli patients with Gram-negative infections [10]. Investigating the association between nonsusceptible genotypes and clinical outcomes may optimize patient care and facilitate the development of new rapid diagnostic tests in clinical settings. Genotype determination is faster than traditional phenotypic susceptibility tests, making it a promising scientific approach. In the previous study, the association between the nonsusceptible phenotype and clinical outcomes was analyzed [9]. Detection of pneumococci nonsusceptible to multiple beta-lactams among patients with ARI did not significantly prolong hospital stay or influence the choice of antimicrobials administered during hospitalization. However, uncertainty related to the agar dilution method [11] may have contributed to these negative findings.

Frequent genetic recombination in SP can lead to rapid clonal selection through the acquisition of DNA fragments that confer survival advantageous characteristics, such as antimicrobial nonsusceptibility [12]. Although MLST is a reproducible and valid molecular typing method, it analyzes only small genomic segments and lacks sufficient resolution for closely related bacterial populations, providing limited insights into genetic evolution beyond its targeted loci. Whole-genome sequencing (WGS) has emerged as a cost-effective and scalable approach for characterizing genetic relationships among closely related isolates [13–15]. However, molecular epidemiological data on pneumococcal strains circulating in Vietnamese communities using WGS remain scarce. Although some studies have applied molecular techniques to characterize SP in Vietnam [5, 7, 16],

comprehensive genomic analysis, particularly of penicillin-binding protein genes, which are the primary determinants of beta-lactam nonsusceptibility [17], is still required.

To develop long-term, effective interventions against antimicrobial nonsusceptibility, it is essential to analyze the detailed mechanisms of beta-lactam nonsusceptibility in SP strains prevalent within the community. The present study aimed to investigate the genetic backgrounds of SP isolates exhibiting nonsusceptibility to multiple beta-lactam antimicrobials in central Vietnam. Whole-genome sequencing was employed to confirm the genomic determinants of antimicrobial nonsusceptibility and to characterize the major lineages identified in previous research. Additionally, the association between the nonsusceptible genotype and clinical outcomes was explored.

The findings indicated that a subset of serogroup 6/ST13223 isolates and all serotype 19F/ST320 isolates carried the identical *pbp2x* allele 16, which was significantly associated with nonsusceptibility to the combination of beta-lactams (PEN, CTX, and MEM). The *pbp2x* allele 16 in certain serogroup 6/ST13223 strains was variably acquired from serotype 19F/ST320 strains through horizontal gene transfer. Although no significant association between the specific *pbp2x* allele and clinical outcome was demonstrated in this study, continuous monitoring of the emergence of nonsusceptibility genes in SP isolates with nonvaccine serotypes remains essential, even after the introduction of the pneumococcal conjugate vaccine.

Methods

Overview of SP surveillance in Nha Trang, central Vietnam

The study design for the original surveillance has been described elsewhere [9, 18, 19]. Briefly, a community-based study on SP colonization was conducted in Nha Trang city between 2008 and 2009. Nasopharyngeal swab samples were obtained from two groups of children under 5 years of age: 331 healthy children randomly selected from the community in July 2008 and 552 sick children with ARIs admitted to the paediatric ward of Khanh Hoa General Hospital between April 7, 2008, and March 30, 2009. The samples were subjected to on-site bacterial culture, followed by antimicrobial susceptibility testing, conventional serotyping, and molecular analyses in a research laboratory. SP isolates were obtained from 95 (28.7%) of the healthy children and 202 (36.6%) of the children with ARIs [9]. An initial agar-dilution screening test revealed alarmingly high nasal carriage rates for SP strains that were nonsusceptible to multiple antimicrobial agents, particularly broad-spectrum beta-lactams, chief among them serotype 19 F and serogroup 6 [9]. Of the 40 SP strains nonsusceptible to multiple beta-lactams, 20 strains (50.0%) belonged to serogroup 19 F, followed by serogroup 6 (n = 14: 35.0%). Further investigation of sequence type (ST) distribution using MLST revealed two major lineages: serotype 19F/ST320 and serogroup 6/ST13223. ST320 is a double-locus variant (DLV) of ST236: Taiwan^{19F}−14 (Pneumococcal Molecular Epidemiology Network [PMEN] 14 clone). ST13223 is a newly identified ST, a DLV of ST473 and a single-locus variant (SLV) of ST647, that was previously reported in Ho Chi Minh City (Additional File 1; Supplementary Table 1).

The impact of pneumococci nonsusceptible to multiple beta-lactams on clinical indices was evaluated. The detection of pneumococci nonsusceptible to multiple beta-lactams among ARI cases did not result in a significant prolongation of hospital stay or affect the type of antimicrobials administered during hospitalization. The data set used for analysis and the details of the analysis are provided in Additional File 2 (Supplementary Table 2).

Antimicrobial susceptibility testing

Minimum inhibitory concentrations (MICs) were re-evaluated using the microdilution method according to Clinical and Laboratory Standards Institute (CLSI) guidelines [20] to validate the results obtained from the agar dilution method in the prior study.

Whole-genome sequencing (WGS)

A preliminary analysis suggested that most SP strains nonsusceptible to multiple beta-lactams belonged to serotype 19 F and serogroups 6, and that nonsusceptible strains belonged to two major STs (serotype 19F/ST320 and serogroup 6/ST13223). WGS was principally performed for these SP lineages. Genomic DNA was extracted using a PowerSoil DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, CA) and fragmented using Covaris S220 (Covaris, Inc., Woburn, MA). DNA libraries were prepared following the method described by Sasagawa et al. [21] using TruSeq adapters (Illumina, San Diego, CA) and a KAPA library preparation kit (KAPA Biosystems, Wilmington, MA). Denatured and diluted DNA samples were sequenced on an Illumina MiSeq platform as multiplexed libraries.

Bioinformatic analysis

FASTQ quality control and de novo assembly were performed using CLC Genomics Workbench ver. 7.5.1 (CLC bio, Aarhus, Denmark). Contigs with low coverage (< 10x) and short lengths (< 300 bp) were excluded from subsequent analyses. Draft genome sequences were annotated using DFAST (DDBJ Fast Annotation and Submission Tool: https://dfast.ddbj.nig.ac.jp) [22]. On the basis of the draft sequences, the serogroups/serotypes and sequences of the three major penicillin-binding protein (*pbp*) genes

were determined. Serogroups/serotypes were determined using the Basic Local Alignment Search Tool (BLAST) from the National Center for Biotechnology Information (NCBI) [23] to search for the sequences of the *cps* locus. For the determination of sequence types, the reference sequences of seven MLST loci were first obtained from the PubMLST website (https://pubmlst.org/organisms/streptococcus-pneumoniae). Allele sequences of a sample were determined by mapping short reads to the reference sequences. MLST profiles were determined by combining alleles using the PubMLST website. *pbp* gene alleles were assigned a unique identifier (e.g., ×*1*, *a1*, and *b1*). In addition to determining the nucleotide sequences of the *pbp* genes, the amino acid sequences of the transpeptidase domain (allele number) of the *pbp* genes, estimated susceptibility patterns to beta-lactam antimicrobials, and global pneumococcal sequencing cluster (GPSC) were analyzed with the genomic surveillance platform Pathogenwatch (https://pathogen.watch). Newly identified *pbp* gene alleles were assigned a unique identifier (e.g., vn_a1, vn_b1, and vn_x1).

MUMmer ver. 3.23 [24] was used for draft sequence alignment with default settings, generating a FASTA-format alignment for sequences sharing 95.0% homology with the reference genome sequence. Scaffolds shorter than 1000 bps, SNPs near the ends of the scaffolds, and indels were excluded manually because of their unreliability. Putative recombination events and phylogenetic relationships excluding recombination regions were identified using Gubbins ver. 2.4.1 with default settings [25]. The genomes of GPSC13 (ST473, a double-locus variant of ST13223) (accession no: ERS1021669) [26] and Taiwan^{19F}−14 (accession no: CP000921) [27] were used as references for the serogroup 6/CC13223 and serotype 19F/ST320 strains, respectively. A maximum likelihood (ML) tree was constructed with 1000 bootstrap replicates using filtered polymorphic site alignments and MEGA ver. 7.0.26 [28]. Phylogenetic trees were constructed using Figtree version 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). The mean pairwise SNPs were also determined using filtered polymorphic site alignment.

Identification of horizontally acquired regions encompassing the pbp2x gene in the serogroup 6/ST13223 genomes

To identify horizontally acquired regions encompassing the *pbp2x* gene in the genomes of serogroup 6/ST13223 isolates, scaffolds containing the *pbp2x* region were selected from each of the serotype 19F/ST320 and serogroup 6/ST13223 genome sequences, and the sequences 20 kb upstream and 5 kb downstream from the *pbp2x* gene were aligned with the appropriate references. Afterwards, the number of sequence variants (SNPs, insertions, and deletions) was counted within a sliding window of 1000 bp with a step size of 100 bp to define the borders of the regions horizontally acquired via recombination.

Determination of pbp2x alleles in nasopharyngeal DNA samples and analysis of their association with a clinical index

The 539–668 bp region of the *pbp2x* gene was amplified via PCR from nasopharyngeal DNA samples and sequenced to explore the association between specific *pbp2x* alleles and a clinical index (hospital stay duration) in ARI cases; carriage of SP, defined as *lyt-A*-positive nasopharyngeal DNA, was identified in 186 ARI cases. The primer pairs used for PCR and sequencing are listed in Supplementary Table 3 (Additional File 3). DNA samples from patients with ARIs and cocolonization of multiple serotypes—detected in approximately 18% of patients using a nanofluidic real-time PCR assay [18]—were excluded from the analysis, because cocolonization of multiple pneumococcal serotypes might affect the sequencing process of the *pbp* gene. The association between a specific *pbp2x* allele and hospital stay duration was examined in the cases where a *pbp2x* allele was confirmed. MICs determined for PEN, CTX, and MEM were based on the results of the previous study [9]. The data set used for analysis and the details of the analysis are provided in Additional File 2 (Supplementary Table 2).

Statistical analysis

The numbers of SP isolates and SP isolates nonsusceptible to beta-lactams were expressed as both real numbers and percentages, categorized by *pbp2x* alleles. The duration of hospital stay was expressed as the median and interquartile range in days. Susceptibility to beta-lactams was classified as a categorical variable. The chi-square test was used when all expected cell counts were ≥ 5; otherwise Fisher's exact test was applied to compare categorical variables. The Wilcoxon rank-sum test was used to compare the length of hospital stay between *pbp2x* allele groups. P-values lower than 0.05 were considered significant. Analyses were performed using JMP ver. 9.0.2 (JMP Statistical Discovery, LLC, Cary, NC).

Results

Whole-genome sequencing of SP isolates randomly selected from the two major lineages and their related strains

To investigate genetic evolution in the SP genome outside MLST-targeted loci (the three major *pbp* genes in

particular), we performed WGS on 39 SP isolates randomly selected from the two major lineages and their related strains: 22 isolates of serotype 19F/ST320, 13 isolates of serogroup 6/ST13223, and four isolates closely related to ST13223. This last group included A1248 (ST14128: SLV of ST13223), A998 (ST2644: DLV of ST13223), H2146 (ST9650: SLV of ST13223), and A1298 (SLV of ST2644). Although A1298 is a triple-locus (*gki*, *spi*, and *ddl*) variant of ST13223, it was included in the lineage analysis; the *ddl* locus, which is near the *pbp2b* gene on the SP genome, is highly variable because of a "hitchhiking" effect involving *ddl* and *pbp2b* [29, 30]. These four isolates were included to provide insights into the genome characteristics of the ST13223 lineage, as detailed genomic data on ST320 and the clonal complex (CC320, including ST236 and ST283) have been well documented [31]. However, such data were lacking for ST13223.

The results of WGS ultimately included data such as inferred serotypes, nucleotide sequences of *pbp* genes (allele numbers), amino acid sequences of the transpeptidase domains of *pbp* genes (allele numbers), estimated beta-lactam susceptibility patterns to analyzed using Pathogenwatch, and STs (Table 1). MIC values against PEN, CTX, and MEM are also shown in Table 1. Across the three major *pbp* genes, we identified 3, 5, and 4 distinct alleles for the *pbp1a*, *pbp2b* and *pbp2x* transpeptidase domains, respectively. All serotype 19F/ST320 isolates shared an identical allele combination for the three major *pbp* genes (13:11:16), and the majority (20 out of 22) were nonsusceptible to both CTX and MEM. This allele combination (13:11:16) was previously reported for serotype 19A/ST320 [32]. In contrast, serogroup 6/ST13223 isolates exhibited three distinct alleles for the *pbp2x* transpeptidase domain (16, 47, and vn_x1), whereas the alleles of *pbp1a* and *pbp2b* (vn_a1 and 12, respectively) were consistent. Notably, the *pbp2x* allele 16, which was shared by all serotype 19F/ST320 isolates, was also shared by eight serogroup 6/ST13223 isolates. Allele 47, found in strain A1548, matched that of Taiwan^{19F}−14 (accession no: CP000921). Importantly, ST13223 isolates with either the 16 or 47 allele exhibited PEN and CTX MICs of 1 μg/ml or higher, whereas isolates with the vn_x1 allele did not; moreover, the MEM MICs did not clearly differ between these groups. *pbp2b* allele 12 in serogroup 6/ST13223 was identical to that of Spain^{23F}−1 (accession no: FM211187) [33]. Among the four non-ST13223 serogroup 6 isolates, three (A1248, A998 and A1298) with the *pbp1a* allele vn_a1 and *pbp2x* allele vn_x1 displayed PEN and CTX MIC values of 0.5 μg/ml or lower. However, their *pbp2b* alleles (vn_b1 or 16) differed from those of ST13223. Although the remaining non-ST13223 (H2146) isolate yielded the same PEN and CTX MIC values, it had a unique transpeptidase domain profile (vn_a2, vn_b2, vn_x2). Moreover, serotype 6E was inferred for H2146, whereas serotype 6 A was inferred for the other non-ST13223 isolates, consistent with the serotype observed in all ST13223 isolates. These features suggest that H2146 was distantly related to the other serogroup 6 isolates.

Novel alleles of the *pbp* genes identified in this study are provided in Additional File 4. The details of the sequence data, including the accession numbers, are provided in Supplementary Table 4 (Additional File 5). Amino acid sequences (allele numbers) of the transpeptidase domains of the PBPs from Taiwan^{19F}−14 and Spain^{23F}−1 are provided in Supplementary Table 5 (Additional File 6).

In summary, the *pbp2x* allele 16, which was shared by all serotype 19F/ST320 isolates and a subset of serogroup 6/ST13223 isolates, and SP isolates with this allele showed increased nonsusceptibility to multiple beta-lactams, regardless of serotype.

Whole-genome phylogeny of serogroup 6/CC13223 isolates

To characterize the genomic features of serogroup 6/CC13223 strains, an ML phylogenetic tree of 17 serogroup 6/CC13223 isolates was constructed on the basis of their genome sequences (Fig. 1a) (Additional File 7, Supplementary Fig. 1a and b), with GPSC 13 (ST473, a DLV of ST13223) [26] as a reference. The ST13223 isolates formed a distinct clade. In this clade, eight isolates with the *pbp2x* allele 16 formed a monophyletic subclade. Among the remaining three non-ST13223 isolates, H2146, which, as mentioned above, demonstrated a unique profile, formed an independent branch, supporting the presumption that it was distantly related to the other isolates. Two isolates (A1298 and A998) with the *pbp2b* allele 16, which is unique among the CC13223 isolates, clustered together and formed an independent branch, suggesting their relatively distant relatedness to the other CC13223 isolates.

Analysis using the Gubbins program revealed 87 recombination events among the serogroup 6/CC13223 isolates, with four recombination regions specifically observed in the genome region spanning *pbp2x*. Owing to the high density of base substitutions in GPSC13, attempts to align the sequences in the *pbp1a* region failed. However, except H2146, all the CC13223 strains possessed the *a1* allele of *pbp1a*, suggesting that no significant recombination events associated with antimicrobial nonsusceptibility occurred in this region among the ST13223 strains. In the genome region spanning *pbp2b*, four recombination events were detected, but all resulted in alterations in both *pbp2b* and *ddl*. Additional analysis

Table 1 Inferred serotypes, MICs, *pbp* alleles of whole lengths, *pbp* alleles of the transpeptidase domain, estimated susceptibility patterns to beta-lactams and sequence types of 39 SP isolates sequenced in this study

Strain ID	Inferred serotype	PEN	CTX	MEM	*pbp* alleles of whole lengths			Transpeptidase domain			Estimated susceptibility pattern to beta-lactams	Sequence type	GPSC
					pbp1a	*pbp2b*	*pbp2x*	*pbp1a*	*pbp2b*	*pbp2x*			
A1256	6A	2	2	0.5	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
A1108	6A	2	2	0.5	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
A1288	6A	2	2	0.5	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
A1212	6A	2	2	0.5	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
A1476	6A	2	2	0.5	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
A1300	6A	2	2	0.5	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
A1139	6A	1	1	0.25	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
H2045	6A	2	2	0.5	*a1*	*b1*	*×1*	vn_a1	12	16	RII	ST13223	1001
A1361	6A	0.5	0.25	0.5	*a1*	*b1*	*×2*	vn_a1	12	vn_x1	ISS	ST13223	1001
A1506	6A	0.5	0.25	0.5	*a1*	*b1*	*×2*	vn_a1	12	vn_x1	ISS	ST13223	1001
A1348	6A	0.5	0.25	0.5	*a1*	*b1*	*×2*	vn_a1	12	vn_x1	ISS	ST13223	1001
A1548	6A	2	1	0.5	*a1*	*b1*	*×3*	vn_a1	12	47	RII	ST13223	1001
A1553	6A	–	–	–	*a1*	*b1*	*×2*	vn_a1	12	vn_x1	ISS	ST13223	1001
A1248	6A	0.5	0.25	0.25	*a1*	*b2*	*×2*	vn_a1	vn_b1	vn_x1	ISS	ST14128	1001
H2146	6E	0.5	0.25	0.25	*a2*	*b4*	*×4*	vn_a2	vn_b2	vn_x2	ISS	ST9650	1001
A1298	6A	0.25	0.25	0.25	*a1*	*b3*	*×2*	vn_a1	16	vn_x1	ISS	SLV of ST2644	1001
A998	6A	0.5	0.25	0.5	*a1*	*b3*	*×2*	vn_a1	16	vn_x1	ISS	ST2644	1001
A1131	19F	2	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1144	19F	2	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1129	19F	2	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1543	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
H2155	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1081	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
H2222	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1014	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1069	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1068	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1211	19F	2	1	0.5	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1093	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1267	19F	2	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1207	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1422	19F	2	1	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1086	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1159	19F	2	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1161	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1191	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1210	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1266	19F	4	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1
A1569	19F	2	2	1	*a3*	*b5*	*×1*	13	11	16	HRR	ST320	1

MICs minimum inhibitory concentrations, *SLV* single-locus variant, *SP Streptococcus pneumoniae*, *PEN* penicillin, *CTX* cefotaxime, *MEM* meropenem, and *PBP*: penicillin-binding protein

The transpeptidase domain amino acid sequences of *pbp* genes and estimated susceptibility patterns to beta-lactam antimicrobials were analyzed using the genomic surveillance platform Pathogenwatch (https://pathogen.watch)

Newly identified alleles of *pbp* genes were assigned arbitrary identifiers

The order of the beta-lactam resistance phenotypes is as follows: (1) PEN, (2) CTX, and (3) MEM

Susceptibility (S), intermediate resistance (I), resistance (R), and high-level resistance (H) were defined according to the criteria established by Metcalf et al. [32]

The three-number combination (13:11:16) was observed in serotype 19 A ST320

Table 1 (continued)

(See figure on next page.)

Fig. 1 Substantial Differences in Genomic Diversity and Antimicrobial Susceptibility between Serogroup 6/CC13223 **a** and Serotype 19F/ST320 Isolates **b**: The analysis of Putative Recombination Regions and Maximum-Likelihood Phylogenetic Trees. Strain IDs are followed by the allele combinations of three major *pbp* genes (*pbp1a*, *pbp2b*, *pbp2x*), listed in parentheses in that order. An ML phylogenetic tree generated using vertically inherited mutations outside regions of horizontal sequence transfer is displayed on the left. For CC13223 isolates (Fig. 1a), the assembled sequence of serotype 6 A global pneumococcal sequence cluster (GPSC) 13 (accession number: GCA 900693075.1) was used as the reference. For the ST320 isolates (Fig. 1b), the genome of the Taiwan^{19F}−14 strain (accession number: CP000921) served as the reference. The middle section illustrates putative recombination events. Isolates of ST13223 and ST320 from Vietnamese children are outlined in light green frames in Figs. 1a and 1b, respectively. The nucleotide positions of significant *pbp* genes are marked on the reference genome at the top. Coloured blocks indicate recombination events; red blocks represent recombination events shared among multiple isolates through common descent, and blue blocks indicate unique recombination events found in individual isolates. The MICs for PEN, CTX, and MEM are displayed on the right. Publicly available whole-genome data for serotype 19F/ST320 strains, including strain SN28306 isolated in Germany in 2006 (accession number: ERR052048) and strain 416185 isolated in the USA in 2001 (accession number: ERR069843**),** were also included in the analysis, as shown in Fig. 1b. In Fig. 1a, the branch length to the outgroup (GPSC13) is substantially long and has been intentionally shortened for clarity. Phylogenetic trees with bootstrap estimates of reliability are presented in Supplementary Fig. 1a, b (Additional File 7). The analysis showed substantial differences in genomic diversity and antimicrobial susceptibility between serogroup 6/ST13223 and serotype 19F/ST320 isolates, particularly the recombination-prone nature of serogroup 6/ST13223. Furthermore, despite frequent recombination events throughout the genome, only the ST13223 subclade with *pbp2x* allele 16, in which two different recombination events in the *pbp2x* region were detected, displayed high PEN and CTX MIC phenotypes (Fig. 1a)

revealed that the 13.5-kb sequence containing *ddl* and *pbp2b*, shared by all the serogroup 6/ST13223 isolates, was also shared by the PMEN clone Spain^{23F}−1 (Additional File 8, Supplementary Fig. 2).

Despite frequent recombination events throughout the genome, only the ST13223 subclade with the *pbp2x* allele 16, which showed two distinct recombination events in the *pbp2x* region, displayed high PEN and CTX MIC phenotypes (Fig. 1a).

Whole-genome phylogeny of serotype 19F/ST320

All serotype 19F/ST320 isolates shared an identical allele combination for the three major *pbp* genes (13:11:16), in contrast to serogroup 6/ST13223 isolates, which exhibited three distinct alleles for the *pbp2x* transpeptidase domain (16, 47, and vn_x1). To characterise the genomic features of serotype19F/ST320 strains, an ML phylogenetic tree was also constructed for serotype 19F/ST320 isolates using the same methodology alongside publicly available genome data for two additional serotype 19F/ST320 isolates: strain SN28306, reported in Germany in 2006, and strain 416185, reported in the USA in 2001 (Accession Numbers: ERR052048 and ERR069843) [31, 34] (Fig. 1b). The genome of Taiwan^{19F}−14 [27] was used as a reference in this analysis.

The 22 serotype 19F/ST320 isolates sequenced in this study formed a monophyletic clade from those of strains SN28306 and 416185. Analysis with Gubbins revealed a recombination profile distinct from that of the serogroup 6/ST13223 strains. This analysis revealed 38 recombination regions; however, most (33 out of 38) were specific to Taiwan^{19F}−14. In addition, no recombination region associated with alterations in *pbp* genes was observed. However, the sequence diversity of the ST320 strains isolated in Nha Trang was comparable to that of the serogroup 6/ST13223 isolates: the mean pairwise SNP distances were 32.6 and 20.7 for the ST320 and serogroup 6/ST13223 isolates, respectively.

In summary, WGS analysis revealed substantial differences in genomic diversity and antimicrobial susceptibility between serogroup 6/ST13223 and serotype 19F/ST320 isolates: the recombination-prone nature was specific to serogroup 6/ST13223.

Horizontally acquired pbp2x-encompassing regions in the serogroup6/ST13223 genomes and identification of sequences possibly derived from serotype19F/ST320 strains

To analyze the sequences surrounding the *pbp2x* gene among the ST13223 isolates, where notable recombination-related variation was observed, we compared these sequences using A1553 as a reference. This isolate was selected because it was the first to separate from other ST13223 isolates, and its *pbp2x* surrounding sequences had no sequence gaps. For this comparison, the sequences of A1256, A1212, A1361, and A1548 were selected from each of the four groups where different recombination events were suggested to have occurred in this region (Fig. 1a), and the numbers of sequence variants (SVs), including SNPs, insertions, and deletions, were counted in a 1000 bp sliding window (with a 100 bp step size) across the 25 kb region from 20 kb upstream to 5 kb downstream of the *pbp2x* gene. Notably, the

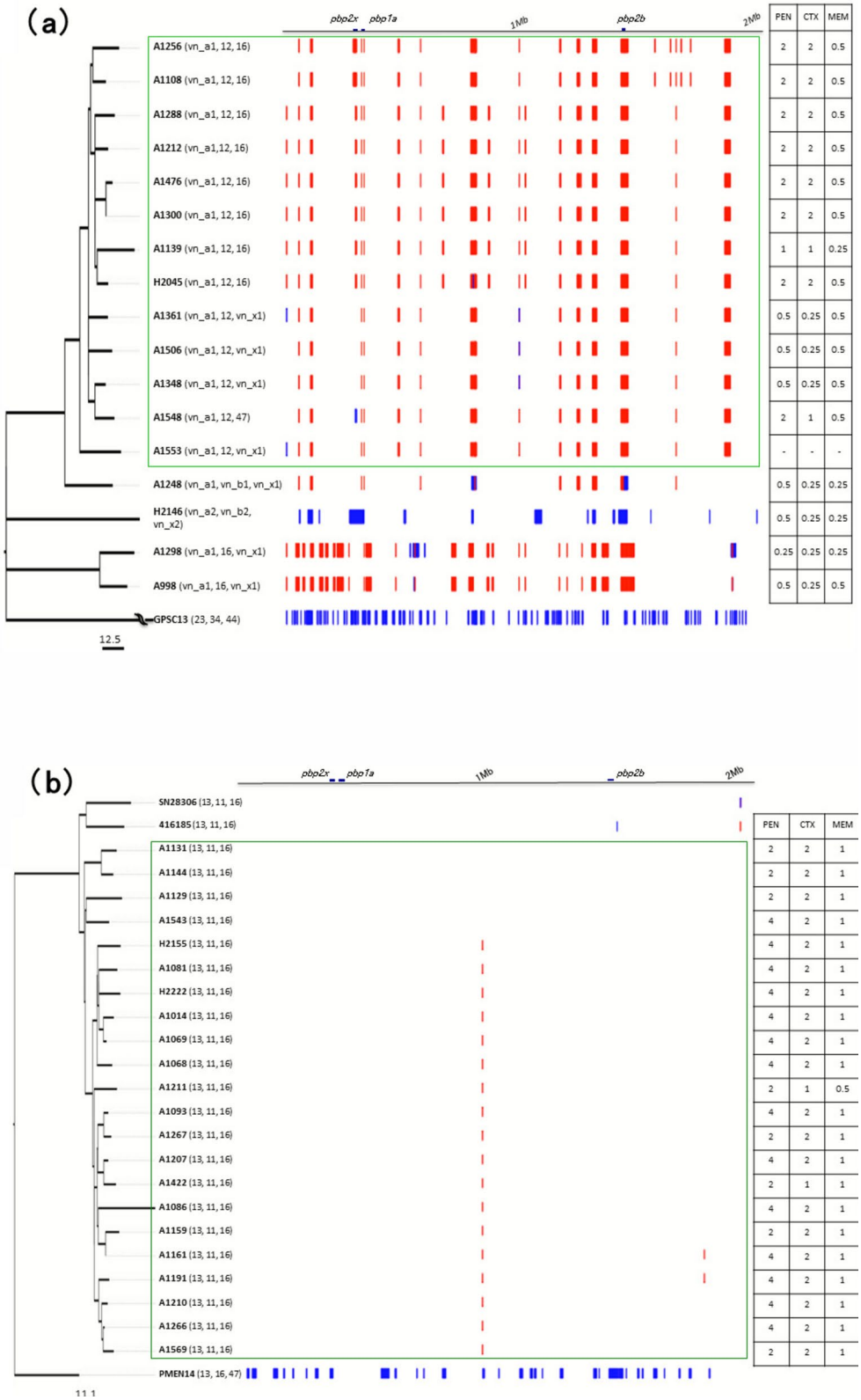

Fig. 1 (See legend on previous page.)

sequences in each group were identical except for a few SNPs in the A1212 group [Supplementary Table 6 (Additional File 9)].

As shown in Fig. 2a, no SNPs were noted between the sequences of A1361 and the reference, which shared the same *pbp2x* allele (vn_x1). In the sequences of A1256, A1212, and A1548, which harboured different *pbp2x* alleles (16, 16, 47), sequences that were apparently introduced by recombination were detected. Consistent with the results shown in Fig. 1a, the recombinogenic regions differed among the three isolates. Notably, the upstream regions of A1256 and A1212, both of which harboured the *pbp2x* allele 16, were apparently of different origins. Although their *pbp2x* genes and downstream regions were nearly identical (while containing multiple SVs), these findings suggest that additional recombination events occurred in the A1256/A1108 lineage or that different recombination events occurred in the two lineages. As the eight ST13223 isolates belonging to the A1256 and A1212 groups shared the *pbp2x* allele 16 with the ST320 isolates, we performed a similar comparison of their *pbp2x*-surrounding sequences. In this analysis, in addition to the A1256 and A1212 sequences, those of A1548 and A1553 were also compared with the sequence of A1131 as a representative of strain ST320 (note that the sequences of 22 ST320 isolates were identical with the exception of a few SNPs (Supplementary Table 7 (Additional File 10)). As shown in Fig. 2b, although the sequence of A1553 (*pbp2x* allele vn_x1) clearly differed from that of ST320, A1256 and A1212 shared almost identical *pbp2x*-spanning sequences with ST320. Interestingly, however, the shared regions of A1256 and A1212 differed; specifically, the A1256 region was much longer (approximately 15.5 kb) than that of A1212 (6.4 kb), supporting the idea that distinct recombination events occurred in the two lineages. Unexpectedly, A1548 (*pbp2x* allele 47) also shared a region with ST320. However, the shared sequence was limited to a 2.4 kb region that included only the 3'-terminal region of *pbp2x*. Moreover, the 5.5 kb sequence of A1548, including the remaining region of *pbp2x* and its upstream region, differed from that of A1553 (*pbp2x* allele vn_x1), which appeared to be the ancestral sequence of the ST13223 lineage. Considering this feature and that A1548 emerged from the ST13223 lineage with the *pbp2x* allele vn_x1 (Fig. 1a), A1548 likely acquired this sequence via a recombination event that differed from those experienced by A1256 and A1212.

As schematically shown in Fig. 3, the results of this group of analyses indicated that *pbp2x*-surrounding sequences were frequently and variably acquired by ST13223 strains from ST320 or their close relatives via recombination. Consequently, it was hypothesized that the *pbp2x* allele16, identified in a subset of serogroup 6/ST13223 strains, was variably obtained by serotype 19F/ST320 strains.

Analysis of the associations of specific pbp2x alleles with drug susceptibility and a clinical index

Is there a significant association of specific *pbp2x* alleles with drug susceptibility and clinical outcome? To investigate the association, we identified the *pbp2x* alleles in nasopharyngeal DNA samples from 134 (88.2%) of the 152 ARI patients tested via PCR and sequencing. The *pbp2x* allele 16 was amplified only from samples associated with serotype 19F/ST320 (n = 23), single locus variant of ST320 (n = 1), and serogroup 6/ST13223 (n = 10). The analysis revealed one additional major allele (allele 18), which was found in 27 patients, predominantly those with serotype 23 F (n = 19), followed by serotype 11 A (n = 7) and nontypeable SP (n = 1). We performed WGS on two of these isolates (serotypes 11 A and 23 F; accession numbers DRR725989 and DRR725990, respectively) and confirmed that their *pbp2x* alleles were compatible

(See figure on next page.)

Fig. 2 Analysis of sequence variants (SVs) in the sequences surrounding the *pbp2x* gene. Scaffolds containing the *pbp2x* region were selected from each of the serotype 19F/ST320 and serogroup 6/ST13223 genome sequences, and the sequences 20 kb upstream and 5 kb downstream from the *pbp2x* gene were aligned with the references (**a**: strain A1553 (ST13223), **b**: strain A1131 (ST320)). The number of SVs was counted in a sliding window of 1000 bp with a step size of 100 bp to define the borders of the regions horizontally acquired via recombination. The number of SVs was considered 100 when the number was more than 100 or if the alignment failed. As shown in Fig. 2a, no SNPs were noted between the sequences of A1361 and the reference, which shared the same *pbp2x* allele (vn_x1). In the sequences of A1256, A1212, and A1548, which harboured different *pbp2x* alleles (16, 16, 47), sequences that were apparently introduced by recombination were detected. Consistent with the results shown in Fig. 1a, the recombinogenic regions differed among the three isolates. Notably, the upstream regions of A1256 and A1212, both of which harboured the *pbp2x* allele 16, were apparently of different origins. Although their *pbp2x* genes and downstream regions were nearly identical (while containing multiple SVs), these findings suggest that additional recombination events occurred in the A1256/A1108 lineage or that different recombination events occurred in the two lineages. As shown in Fig. 2b, although the sequence of A1553 (*pbp2x* allele vn_x1) clearly differed from that of ST320, A1256 and A1212 shared almost identical *pbp2x*-spanning sequences with ST320. Interestingly, however, the shared regions of A1256 and A1212 were different; specifically, the region of A1256 was much longer (approximately 15.5 kb) than that of A1212 (6.4 kb), supporting the idea that different recombination events occurred in the two lineages. Unexpectedly, A1548 (*pbp2x* allele 47) also shared a region with ST320

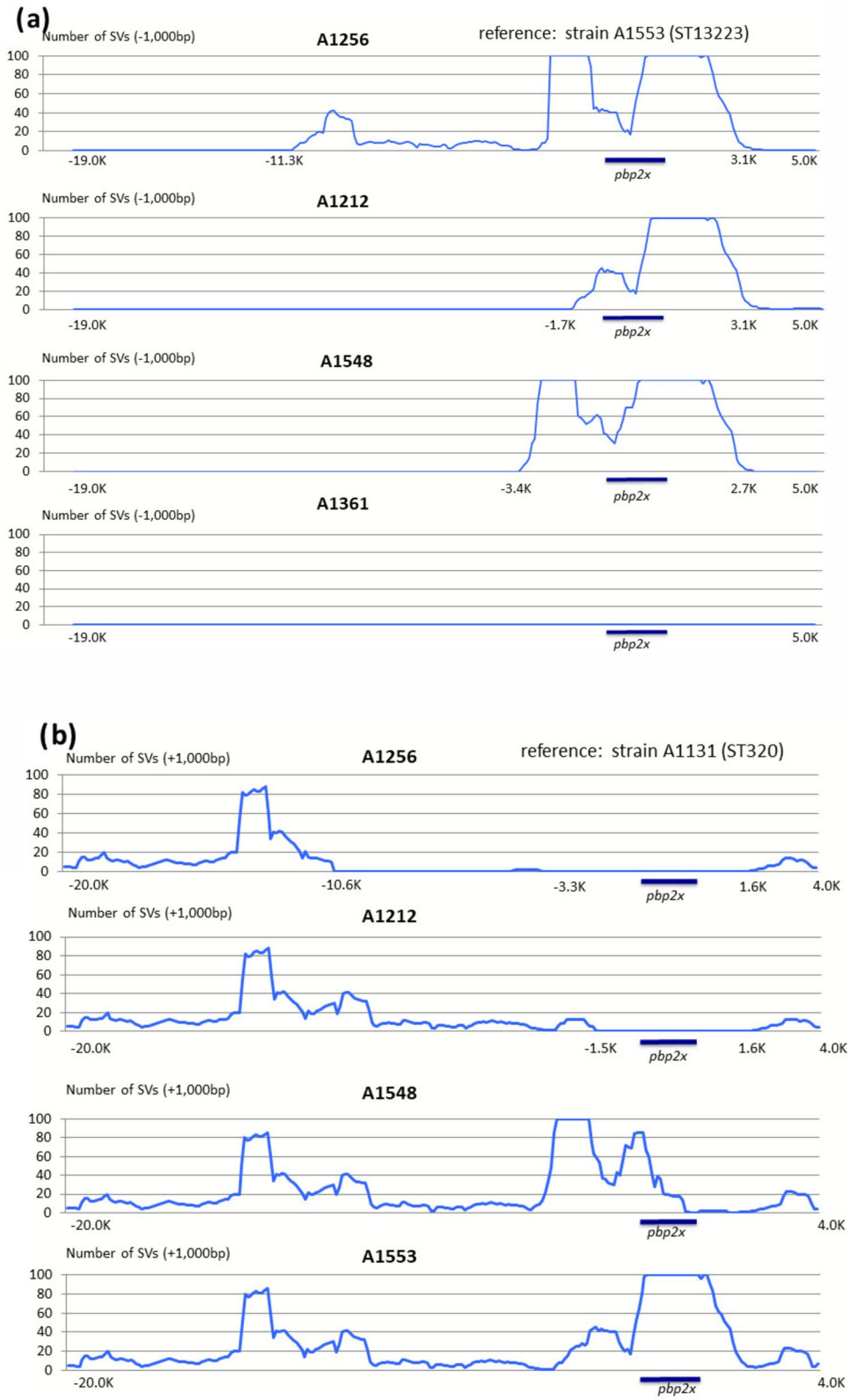

Fig. 2 (See legend on previous page.)

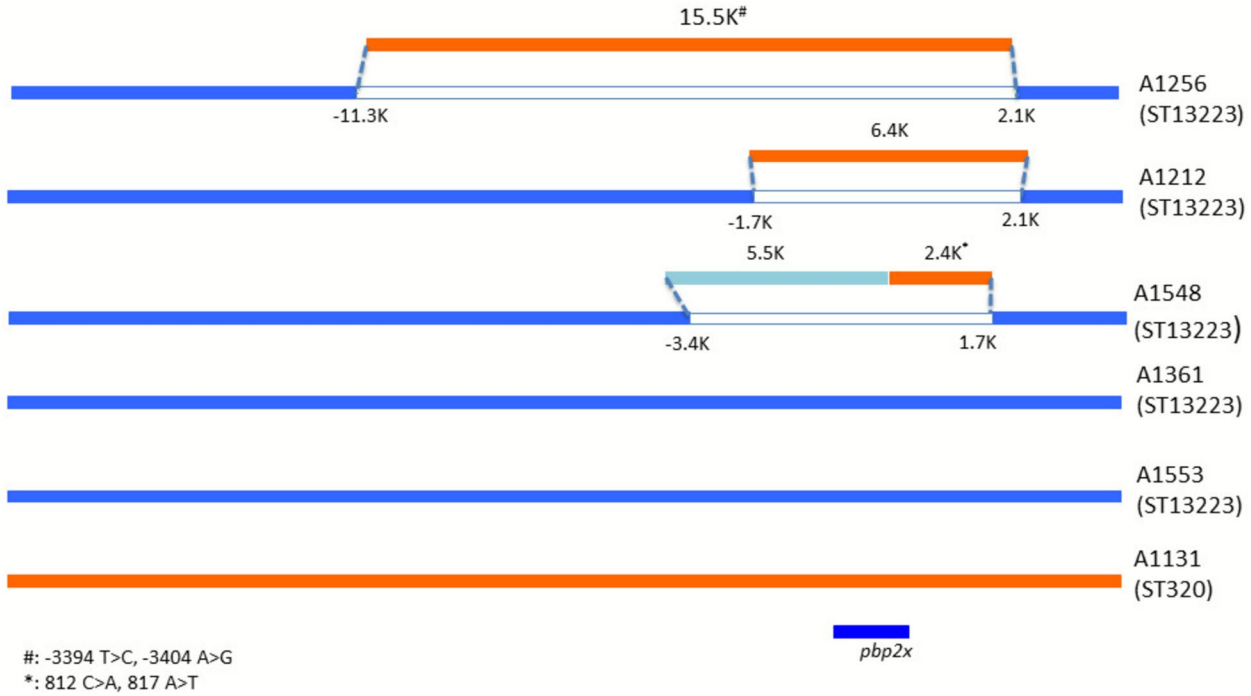

Fig. 3 Schematic of a putative recombination between the ST13223 and ST320 strains. A1256 and A1212 shared almost identical *pbp2x*-spanning sequences with ST320. However, the regions shared by A1256 and A1212 differed. Specifically, the region of A1256 was much longer (approximately 15.5 kb) than that of A1212 (6.4 kb), supporting the idea that different recombination events occurred in the two lineages. A1548 (*pbp2x* allele 47) also shared a region with ST320, but the shared sequence was limited to a 2.4 kb region that included only the 3′-terminal region of *pbp2x*. Moreover, the 5.5 kb sequence of A1548, including the remaining region of *pbp2x* and its upstream region, differed from that of A1553. Considering this feature and that A1548 emerged from the ST13223 lineage with the *pbp2x* allele vn_x1, A1548 acquired this sequence via a recombination event different from those in A1256 and A1212

Table 2 MIC^{50} and hospital stay duration of *S. pneumoniae* isolates categorized by *pbp2x* alleles

***pbp2x* allele**	**n (%)**	**PEN**	**CTX**	**MEM**	**Nonsusceptible to PEN&CTX, n (%)**	**Nonsusceptible to CTX&MEM, n (%)**	**Nonsusceptible to PEN&CTX&MEM, n (%)**	**Duration of hospital stay***
16	34 (25.4)	2	2	1	11 (32.4)	18 (52.9)	11 (32.4)	5 (3–6)
18	27 (20.1)	2	1	1	0 (0)	0 (0)	0 (0)	4 (4–6)
47	19 (14.2)	2	0.5	0.5	1 (5.3)	4 (21.1)	1 (5.3)	4 (3–6)
vn_x1	14 (10.4)	0.5	0.25	0.5	1 (7.1)	4 (28.6)	1 (7.1)	4 (3–6.25)
vn_x2	5 (3.7)	1	0.25	0.25	0 (0)	0 (0)	0 (0)	5 (2.5–10)
Others#	35 (26.1)	0.5	0.5	0.25	2 (5.7)	4 (11.4)	2 (5.7)	5 (3–6)

MIC minimum inhibitory concentration, *PBP* penicillin-binding protein, *PEN* penicillin, *CTX* cefotaxime, and *MEM* meropenem

* Median (IQR) in days

\# "Others" includes 24 different alleles, each occurring at a low frequency ranging from 1 to 5. These alleles were included in this analysis

with those of Spain^{23F}−1 (PMEN 1). Although the origin of this allele was unclear, SP isolates with serotypes 11 A and 23 F in Vietnam shared the identical pbp2x allele with the PMEN clone.

The MIC^{50} values for the SP isolates classified by *pbp2x* alleles are summarized in Table 2. Allele 16 accounted for 25.4% of the cases, followed by alleles 18 (20.1%), 47 (14.2%), and vn_x1 (10.4%). The analysis using Fisher's exact test revealed that allele 16 was significantly associated with phenotypic nonsusceptibility to multiple beta-lactams (PEN, CTX and MEM) ($p < 0.0001$; odds ratio [95% CI] 11.5 [3.35–39.3]). Notably, allele 16 was detected only in the vaccine-type strains (that is, serotype 19 F and serogroup 6). The analysis using the Wilcoxon rank-sum test revealed no significant association

between the presence of this allele and prolonged hospitalization (p = 0.6123).

In summary, allele 16 was significantly associated with phenotypic nonsusceptibility to multiple beta-lactams and was detected only in the vaccine-type strains. However, no significant association between specific *pbp2x* alleles and clinical outcome was demonstrated.

Discussion

The correlation between antimicrobial-nonsusceptible genotypes and clinical outcomes was investigated using data from a previously established clinical epidemiological field study, in which cases were consecutively registered at a single hospital, and nasopharyngeal DNA samples were collected without bias [9]. WGS was utilized to confirm the genomic determinants of antimicrobial nonsusceptibility and to characterize the major lineages highlighted in the prior study: serotype 19F/ST320 and a new ST, ST13223, belonging to serogroup 6. These lineages accounted for the high prevalence of SP isolates that were nonsusceptible to multiple beta-lactams in the community. WGS analysis revealed substantial genomic diversity among serogroup 6/ST13223 isolates and elucidated the genetic background underlying their antimicrobial susceptibility. In contrast to serotype 19F/ST320 isolates, serogroup 6/ST13223 isolates were characterized by frequent recombination events; in particular, those spanning *pbp2x* were associated with high MICs for multiple beta-lactams.

The introduction of the PCV was shown to reduce vaccine serotype nonsusceptible SP carriage and disease [35, 36], whereas subsequent increases in less susceptible nonvaccine types had been observed [37]. Particularly, serotype 19A/ST320 has been identified as a rapidly emerging multidrug-resistant clone responsible for invasive pneumococcal disease cases following the introduction of the 7-valent pneumococcal conjugate vaccine (PCV-7) in the USA [37], Canada [38], Spain [39], South Korea, and numerous other Asian countries [40]. In this study, ST320 was the major lineage among SP isolates with serotype 19F, and this strain possessed the *pbp* allele combination (13:16:16), which was previously reported for serotype 19A/ST320 [32]. Further genomic analyses would clarify how SP isolates with serotype 19F in Vietnam acquired this genotype.

Obviously, the clonal spread of serotype 19F/ST320 strains was one of the causes of high carriage rates of SP nonsusceptibility to multiple beta-lactams in Vietnamese children. The increased beta-lactam nonsusceptibility was validated using the microdilution method. In addition to serotype 19A/ST320, the multidrug-resistant serotype 19F/ST320 has recently become prevalent in several Asian countries [41–43]. However, ST320 had not been reported in Vietnam prior to this study; however, its ancestral strain, ST236, was known to circulate in the region [5, 7]. In the preliminary analysis, we identified five ST236 isolates and eight SLV isolates of ST236 (Additional File-1; Supplementary Table-1), none of which exhibited nonsusceptibility to both CTX and MEM in our initial screening using the agar-dilution method. This finding suggests that ST320 is advantageous under the selective pressure of beta-lactams. Analysis of the whole-genome phylogeny of the serotype 19F/ST320 lineage revealed notable genetic diversity within this lineage, suggesting that ST320 may have been introduced into the central Vietnamese community on multiple occasions or emerged locally from existing lineages. The 22 serotype 19F/ST320 isolates from central Vietnam formed a distinct clade from SN28306 and 416185. To determine the timing and origin of this lineage in the region, further analyses using larger genomic datasets from various periods and locations are necessary.

Unlike serotype 19F/ST320 isolates, several serogroup 6/ST13223 isolates did not satisfy nonsusceptible MIC breakpoints for either CTX or MEM. Further WGS analysis and genomic comparison between susceptible and nonsusceptible isolates revealed the existence of serogroup 6/ST13223 sublineages that shared identical *pbp* alleles (*pbp2b* and *pbp2x*) with different PMEN-related strains. We hypothesized that multiple recombination events occurred near the local community, conferring a selective advantage to these SP strains. This hypothesis aligns with our observations of frequent nasopharyngeal cocolonization of multiple SP serotypes, including serotype 19F and serogroup 6, in this community [18]. Additional contributing factors include the high prevalence of PMEN strains, extensive antimicrobial consumption [9], and longer colonization durations of serotype 6A [44]. The recombination-prone nature of serogroup 6/ST13223 and reduced eradication rates of non-penicillin–susceptible SP following the administration of inadequate doses of beta-lactams [45] may have further promoted frequent recombination events in the *pbp* spanning region. Further investigation would be needed to determine whether rational use of beta-lactams can restore drug susceptibility.

More than half (59.7%) of the analysed SP isolates (n = 134) carried *pbp2x* alleles associated with beta-lactam-nonsusceptible PMEN strains: allele 16 (25.4%, ST320), allele 18 (20.1%, ST81 (Spain^{23F}–1 clone) or ST166 (SLV of ST156 (Spain^{9V}–3 clone)), and allele 47 (14.1%, ST236 (Taiwan^{19F}–14 clone). Most isolates that did not exhibit susceptibility to multiple beta-lactams belonged to PCV serotypes, suggesting that introducing a PCV could reduce the incidence of

antibiotic-nonsusceptible disease caused by vaccine-type serotypes [35] and restore drug susceptibility. However, subsequent increases in less susceptible nonvaccine types had also been observed [37]. Furthermore, nonencapsulated SP lineages, which were not covered by the current vaccines, were shown to have a major role in genetic exchange, including the *pbp* genes [10]. Continuous monitoring of the prevalence of nonsusceptibility genes and the emergence of nonsusceptibility in nonvaccine serotypes through genomic recombination is therefore essential, and rigorous antimicrobial regulation should be prioritized.

This study examined the association between the nonsusceptible genotype and clinical outcome, as this investigation could contribute to the development of new diagnostic tests in clinical settings. The *pbp2x* allele associated with the nonsusceptible genotype demonstrated minimal impact on clinical outcome. While treatment failure due to penicillin nonsusceptibility phenotype has been reported in patients with pneumococcal meningitis [46], most studies have not established a relationship between elevated beta-lactam MICs and adverse clinical outcomes, such as mortality or hospital stay duration, in pneumococcal pneumonia [47–49]. An increase in pneumococcal mastoiditis cases caused by multidrug-resistant serotype 19A SP, including some ST320 isolates, was reported in one institution [50]. However, the primary driver of this complication—whether antimicrobial nonsusceptibility, serotype, or other genotypic factors—remains unclear. Although serotype 19F/ST320 is prevalent in other Asian countries [41–43], its impact on clinical outcomes has not been investigated. Delayed bacterial eradication because of less potent antimicrobial agents with high MICs could facilitate the dissemination of serotype 19F/ST320. To identify genotypes that significantly influence clinical outcomes, studies with larger sample sizes and more comprehensive clinical data, such as fever duration, may be required. Alternatively, focusing on more severe disease presentations, such as invasive pneumococcal disease, may be necessary. This strategy could involve the utilization of nationwide cohorts and the collection of whole-genome data for all identified strains. Ongoing reductions in whole-genome sequencing costs and the development of analytical pipelines are expected to support these efforts.

This study has several limitations. First, the study used relatively old stored samples collected in 2008 and 2009, which are unlikely to reflect the current state of SP strain nonsusceptibility and may have limited implications for ongoing public health measures. However, we have continued to conduct pneumococcal surveillance at the study site. In a hospital-based surveillance programme from 2015 to 2016, during which ST320 and ST13223 remained the predominant STs [51]. Therefore, the findings in this study are worth publishing, as they provide valuable baseline data prior to the introduction of PCVs, offering a foundation for further molecular epidemiological studies. Second, selection bias may be present, as our MLST results were primarily obtained from patients with ARIs. Nevertheless, we believe that two major lineages (serotype 19F/ST320 and serogroup 6/ST13223) predominated among both children with ARIs and healthy children. Our previous studies revealed high levels of non-beta-lactam-susceptibility among SP isolates, regardless of the health status of the child [9]. Third, antimicrobial susceptibility testing of a limited number of SP isolates suggested that the initial screening using the agar-dilution method may have overestimated non-beta-lactam-susceptibility, particularly in serogroup 6 isolates. Finally, the analyzed strains represent only a subset of strains of interest, and the majority of circulating diversity was excluded, which may obscure the origins of recombination. However, analysis of nasopharyngeal DNA samples showed that the *pbp2x* allele 16 was amplified only from samples associated with serotype 19F/ST320, a single locus variant of ST320, and serogroup 6/ST13223. It is therefore plausible that some serogroup 6/ST13223 strains acquired a *pbp2x* allele from serotype 19F/ST320, although introduction from another lineage outside Nha Trang cannot be excluded.

In conclusion, the findings indicate that the dynamics of the *pbp2x* allele (allele 16) are the primary drivers of nonsusceptibility to multiple beta-lactams among SP isolates circulating in central Vietnam. However, no significant association between the *pbp2x* allele and clinical outcome was observed. Continuous monitoring of the emergence of nonsusceptibility genes in SP isolates with nonvaccine serotypes remains essential, even following the introduction of the pneumococcal conjugate vaccine.

Abbreviations

ARI	Acute respiratory infection
BLAST	Basic Local Alignment Search Tool
CI	Confidence interval
CLSI	Clinical and Laboratory Standards Institute
CTX	Cefotaxime
DFAST	DDBJ Fast Annotation and Submission Tool
DLV	Double-locus variant
GPSC	Global Pneumococcal Sequencing Cluster
MEM	Meropenem
MIC	Minimum inhibitory concentration
ML	Maximum likelihood
MLST	Multilocous sequence typing
NCBI	National Center for Biotechnology Information
PCR	Polymerase chain reaction
PCV	Pneumococcal conjugate vaccine
PEN	Penicillin
PBP	Penicillin-binding protein
SLV	Single-locus variant
SNP	Single nucleotide polymorphism
SP	*Streptococcus pneumoniae*
ST	Sequence type

SV Sequence variant
WGS Whole-genome sequencing

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s41182-025-00889-0.

Additional file 1: Table S1: Sequence Type Distribution of *S. pneumoniae* with common serotypes

Additional file 2: Table S2: The data set used for analysis and the details of the analysis

Additional file 3: Table S3: Primer Pairs Designed for Partial Sequencing of the *pbp2x* Gene

Additional file 4: Novel alleles of the *pbp* genes identified in this study

Additional file 5: Table-S4: The Details of Sequence Data

Additional file 6: Table-S5: Serotypes, GPSC, Transpeptidase domain amino acid sequences of PBP genes, Estimated susceptibility pattern to beta-lactam antimicrobials, and Sequence Types of the two PMEN clones

Additional file 7: Figure S1a and b: An ML Phylogenetic Tree with an estimation of the reliability by bootstrapping

Additional file 8: Figure S2: Illustration of the shared 13.5kb nucleotide sequence containing both *ddl* and *pbp2b* between serogroup 6/ST13223 and the PMEN clone Spain^{23F}-1

Additional file 9: Table S6: The details of the sequences of the genome region surrounding *pbp2x*

Additional file 10: Table S7: The details of the sequences of the genome region surrounding *pbp2x*

Acknowledgements
We thank the staff members of Khanh Hoa General Hospital and Khanh Hoa Health Service for their help with the study. We also thank Kyoko Uchibori and Stephanie Jane Airs for their technical assistance with the susceptibility testing. We are grateful to Dr. Takayuki Wada for his advice on the bioinformatics analyses. We thank Ryan McCorkell for revising the draft of the manuscript.

Author contributions
LMY, DDA, and KA designed the study and coordinated the field work. HATN and HTTV were responsible for data and sample collection, laboratory testing, and data analysis. HF performed susceptibility testing. BGD provided the data on serotyping. HF and SK performed WGS with the technical support of S. Nakamura, DM, and TI. HF and MS performed bioinformatic analyses. HF performed the statistical analysis. HF drafted the manuscript with YO, S. Nakano, CMP, KM, TH, and KA. All authors read and approved the final manuscript.

Funding
This research is supported by the Japan Initiative for Global Research Network on Infectious Diseases (J-GRID) from the Ministry of Education, Culture, Sport, Science & Technology in Japan and the Japan Agency for Medical Research and Development (AMED). It is also supported by the Joint Usage/Research Center on Tropical Diseases, Institute of Tropical Medicine, Nagasaki University (2014-Ippan-10, 2015-Ippan-25).

Data availability
The data set used for analysis and the details of the analysis are provided in Additional File 2. The details of the sequence data, including the accession numbers, are provided in Additional File 5.

Declarations

Ethics approval and consent to participate
The research protocol was approved by the Nagasaki University Ethical Review Board, the National Institute of Hygiene and Epidemiology scientific review committee in Hanoi, and the Khanh Hoa Provincial Health Service ethical review board. Written informed consent was obtained from the parents or guardians prior to sample collection and the interviews.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Infection Research, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8102, Japan. [2]Department of Bacteriology, Faculty of Medical Sciences, Kyushu University, Fukuoka 812-8582, Japan. [3]Department of Bacteriology, National Institute of Hygiene and Epidemiology, Hanoi 100000, Vietnam. [4]Department of Clinical Medicine, Institute of Tropical Medicine, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan. [5]School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki 852-8523, Japan. [6]Department of Infection Metagenomics, Bioinformatics Center, Research Institute for Microbial Diseases, Osaka University, Suita 565-0871, Japan. [7]Antimicrobial Resistance Research Center, Japan Institute for Health Security, National Institute of Infectious Diseases, Tokyo 162-0052, Japan. [8]Department of Respiratory Infections, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan. [9]Department of Pediatric Infectious Diseases, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan. [10]Department of Bacterial Infections, Research Center for Infectious Disease Control, Research Institute for Microbial Diseases, Osaka University, Suita 565-0871, Japan. [11]Department of Frontier Research and Development, Kazusa DNA Research Institute, Kisarazu 292-0818, Japan. [12]Infection Control and Education Center, Nagasaki University Hospital, Nagasaki 852-8501, Japan. [13]Division of Microbiology, Department of Infectious Medicine, Kurume University School of Medicine, Kurume 830-0011, Japan. [14]Centre for Tropical Medicine and Global Health, University of Oxford, Oxford OX3 7LG, UK.

Received: 1 October 2025 Accepted: 22 December 2025
Published online: 10 February 2026

References

1. Walker CLF, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, et al. Global burden of childhood pneumonia and diarrhoea. Lancet. 2013;381:1405–16.
2. O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, et al. Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. Lancet. 2009;374:893–902.
3. Jacobs MR, Felmingham D, Appelbaum PC, Grüneberg RN. The alexander project 1998–2000: susceptibility of pathogens isolated from community-acquired respiratory tract infection to commonly used antimicrobial agents. J Antimicrob Chemother. 2003;52:229–46.
4. Song JH, Lee NY, Ichiyama S, Yoshida R, Hirakata Y, Fu W, et al. Spread of drug-resistant *Streptococcus pneumoniae* in Asian countries: Asian network for surveillance of resistant pathogens (ANSORP) study. Clin Infect Dis. 1999;28:1206–11.
5. Schultsz C, Vien le M, Campbell JI, Chau NVV, Diep TS, Hoang NVM, et al. Changes in the nasal carriage of drug-resistant *Streptococcus pneumoniae* in urban and rural Vietnamese schoolchildren. Trans R Soc Trop Med Hyg. 2007;101:484–92. https://doi.org/10.1016/j.trstmh.2006.08.010.
6. Hoa NQ, Trung NV, Larsson M, Eriksson B, Phuc HD, Chuc NT, et al. Decreased *Streptococcus pneumoniae* susceptibility to oral antibiotics among children in rural Vietnam: a community study. BMC Infect Dis. 2010;10:85. https://doi.org/10.1186/1471-2334-10-85.
7. Parry CM, Duong NM, Zhou J, Mai NTH, Diep TS, Thinh LQ, et al. Emergence in Vietnam of *Streptococcus pneumoniae* resistant to multiple antimicrobial agents as a result of dissemination of the multiresistant Spain^{23F}-1 clone. Antimicrob Agents Chemother. 2002;46:3512–7.
8. Kim SH, Song JH, Chung DR, Thamlikitkul V, Yang Y, Wang H, et al. Changing trends in antimicrobial resistance and serotypes of Streptococcus pneumoniae isolates Asian countries: an Asian in network for surveillance of resistant pathogens (ANSORP) study. Antimicrob Agents Chemother. 2012;56:1418–26.

9. Nguyen HAT, Fujii H, Vu HTT, Parry CM, Dang AD, Ariyoshi K, et al. An alarmingly high nasal carriage rate of *Streptococcus pneumoniae* serotype 19F non-susceptible to multiple beta-lactam antimicrobials among Vietnamese children. BMC Infect Dis. 2019;19:241.
10. Koch RE, Barth J, Clark AE, Desai D, Kim J, Pybus CA, et al. Antibiotic resistance genotype, phenotype, and clinical outcomes in patients with Gram-negative infections at Rabin medical center in Israel. Microbiol Spectr. 2024;13:e00383-e424.
11. Zhang SX, Rawte P, Brown S, Lo S, Siebert H, Porter SP, et al. Evaluation of CLSI agar dilution method and Trek sensititre broth microdilution panel for determining antimicrobial susceptibility of *Streptococcus pneumoniae*. J Clin Microbiol. 2011;49:704–6.
12. Chewapreecha C, Harris SR, Croucher NJ, Turner C, Marttinen P, Cheng L, et al. Dense genomic sampling identifies highways of pneumococcal recombination. Nat Genet. 2014;46:305–9.
13. Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, et al. Evolution of MRSA during hospital transmission and intercontinental spread. Science. 2010;327:469–74.
14. Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, et al. Rapid pneumococcal evolution in response to clinical interventions. Science. 2011;331:430–4.
15. Croucher NJ, Hanage WP, Harris SR, McGee L, van der Linden M, de Lencastre H, et al. Variable recombination dynamics during the emergence, transmission and "disarming" of a multidrug-resistant pneumococcal clone. BMC Biol. 2014;12:49.
16. Song JH, Chang HH, Suh JY, Ko KS, Jung SI, Oh WS, et al. Macrolide resistance and genotypic characterization of *Streptococcus pneumoniae* in Asian countries: a study of the Asian network for surveillance of resistant pathogens (ANSORP). J Antimicrob Chemother. 2004;53:457–63.
17. Dowson CG, Coffey TJ, Spratt BG. Origin and molecular epidemiology of penicillin-binding-protein-mediated resistance to beta-lactam antibiotics. Trends Microbiol. 1994;2:361–6.
18. Dhoubhadel BG, Yasunami M, Nguyen HA, Suzuki M, Vu TH, Nguyen ATT, et al. Bacterial load of pneumococcal serotypes correlates with their prevalence and multiple serotypes is associated with acute respiratory infections among children less than 5 years of age. PLoS One. 2014;9:e110777.
19. Vu HT, Yoshida LM, Suzuki M, Nguyen HA, Nguyen CD, Nguyen AT, et al. Association between nasopharyngeal load of *Streptococcus pneumoniae,* viral coinfection, and radiologically confirmed pneumonia in Vietnamese children. Pediatr Infect Dis J. 2011;30:11–8.
20. Clinical and Laboratory Standards Institute. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically. 10th ed. Wayne, PA: CLSI; 2015.
21. Sasagawa Y, Nikaido I, Hayashi T, Danno H, Uno KD, Imai T, et al. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol. 2013;14:R31.
22. Tanizawa Y, Fujisawa T, Kaminuma E, Nakamura Y, Arita M. DFAST and DAGA: web-based integrated genome annotation tools and resources. Bioscience of Microbiota, Food and Health. 2016;35:173–84.
23. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
24. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biol. 2004;5:R12.
25. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using gubbins. Nucleic Acids Res. 2015;43:e15.
26. Gladstone RA, Lo SW, Goater R, Yeats C, Taylor B, Hadfield J, et al. Visualizing variation within global pneumococcal sequence clusters (GPSCs) and country population snapshots to contextualize pneumococcal isolates. Microb Genom. 2020;6:e000357.
27. Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli SV, et al. Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. Genome Biol. 2010;11:R107.
28. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–4.
29. Enright MC, Spratt BG. Extensive variation in the ddl gene of penicillin-resistant *Streptococcus pneumoniae* results from a hitchhiking effect driven by the penicillin-binding protein 2b gene. Mol Biol Evol. 1999;16:1687–95.
30. Chi F, Nolte O, Bergmann C, Ip M, Hakenbeck R. Crossing the barrier: evolution and spread of a major class of mosaic pbp2x in *Streptococcus pneumoniae*, *S. mitis* and *S. oralis*. Int J Med Microbiol. 2007;297:503–12.
31. Croucher NJ, Chewapreecha C, Hanage WP, Harris SR, McGee L, van der Linden M, et al. Evidence for soft selective sweeps in the evolution of pneumococcal multidrug resistance and vaccine escape. Genome Biol Evol. 2014;6:1589–602.
32. Metcalf BJ, Gertz RE Jr, Gladstone RA, Walker H, Sherwood LK, Jackson D, et al. Strain features and distributions in pneumococci from children with invasive disease before and after 13-valent conjugate vaccine implementation in the USA. Clin Microbiol Infect. 2016;22(60):e9-29.
33. Croucher NJ, Walker D, Romero P, Lennard N, Paterson GK, Bason NC, et al. Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone *Streptococcus pneumoniae* Spain23F ST81. J Bacteriol. 2009;191:1480–9.
34. Croucher NJ, Finkelstein JA, Pelton SI, Mitchell PK, Lee GM, Parkhill J, et al. Population genomics of post-vaccine changes in pneumococcal epidemiology. Nat Genet. 2013;45:656–63.
35. Kyaw MH, Lynfield R, Schaffner W, Craig AS, Hadler J, Reingold A, et al. Active bacterial core surveillance of the emerging infections program network. Effect of introduction of the pneumococcal conjugate vaccine on drug-resistant *Streptococcus pneumoniae*. N Engl J Med. 2006;354:1455–63. https://doi.org/10.1056/NEJMoa051642.
36. Huang SS, Platt R, Rifas-Shiman SL, Pelton SI, Goldmann D, Finkelstein JA. Post-PCV7 changes in colonizing pneumococcal serotypes in 16 Massachusetts communities, 2001 and 2004. Pediatrics. 2005;116:e408–13. https://doi.org/10.1542/peds.2004-2338.
37. Moore MR, Gertz RE Jr, Woodbury RL, Barkocy-Gallagher GA, Schaffner W, Lexau C, et al. Population snapshot of emergent *Streptococcus pneumoniae* serotype 19A in the United States, 2005. J Infect Dis. 2008;197:1016–27.
38. Pillai DR, Shahinas D, Buzina A, Pollock RA, Lau R, Khairnar K, et al. Genome-wide dissection of globally emergent multi-drug resistant serotype 19A *Streptococcus pneumoniae*. BMC Genomics. 2009;10:642.
39. Ardanuy C, Rolo D, Fenoll A, Tarrago D, Calatayud L, Liñares J. Emergence of a multidrug-resistant clone (ST320) among invasive serotype 19A pneumococci in Spain. J Antimicrob Chemother. 2009;64:507–10.
40. Shin J, Baek JY, Kim SH, Song JH, Ko KS. Predominance of ST320 among *Streptococcus pneumoniae* serotype 19A isolates from 10 Asian countries. J Antimicrob Chemother. 2011;66:1001–4.
41. Zhao C, Zhang F, Chu Y, Liu Y, Cao B, Chen M, et al. Phenotypic and genotypic characteristic of invasive pneumococcal isolates from both children and adult patients from a multicenter surveillance in China 2005-2011. PLoS ONE. 2013;8:e82361.
42. Safari D, Kuo LC, Huang YT, Liao CH, Sheng WH, Hsueh PR. Increase in the rate of azithromycin-resistant *Streptococcus pneumoniae* isolates carrying the erm(B) and mef(A) genes in Taiwan, 2006–2010. BMC Infect Dis. 2014;14:704.
43. Ip M, Ang I, Liyanapathirana V, Ma H, Lai R. Genetic analyses of penicillin binding protein determinants in multidrug-resistant *Streptococcus pneumoniae* serogroup 19 CC320/271 clone with high-level resistance to third-generation cephalosporins. Antimicrob Agents Chemother. 2015;59:4040–5.
44. Hill PC, Cheung YB, Akisanya A, Sankareh K, Lahai G, Greenwood BM, et al. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. Clin Infect Dis. 2008;46:807–14.
45. Dagan R, Leibovitz E. Bacterial eradication in the treatment of otitis media. Lancet Infect Dis. 2002;2:593–604.
46. John CC. Treatment failure with use of a third-generation cephalosporin for penicillin-resistant pneumococcal meningitis: case report and review. Clin Infect Dis. 1994;18:188–93.
47. Tan TQ, Mason EO Jr, Barson WJ, Wald ER, Schutze GE, Bradley JS, et al. Clinical characteristics and outcome of children with pneumonia attributable to penicillin-susceptible and penicillin-nonsusceptible *Streptococcus pneumoniae*. Pediatrics. 1998;102:1369–75.
48. Ewig S, Ruiz M, Torres A, Marco F, Martinez JA, Sanchez M, et al. Pneumonia acquired in the community through drug-resistant *Streptococcus pneumoniae*. Am J Respir Crit Care Med. 1999;159:1835–42.
49. Deeks SL, Palacio R, Ruvinsky R, Kertesz DA, Hortal M, Rossi A, et al. Risk factors and course of illness among children with invasive penicillin-resistant *Streptococcus pneumoniae*. Pediatrics. 1999;103:409–13.

50. Ongkasuwan J, Valdez TA, Hulten KG, Mason EO Jr, Kaplan SL. Pneumococcal mastoiditis in children and the emergence of multidrug-resistant serotype 19A isolates. Pediatrics. 2008;122:34–9.
51. Wambugu P, Shah MM, Nguyen HA, Le KA, Le HH, Vo HM, et al. Molecular epidemiology of *Streptococcus pneumoniae* detected in hospitalized pediatric acute respiratory infection cases in central Vietnam. Pathogens. 2023;12:943.

Publisher's Note