scientific reports

Check for updates

OPEN

Evaluating long-read metagenomics for bloodstream infection diagnostics: a pilot study from a Thai Tertiary Hospital

Thunchanok Yaikhan[1], Thidathip Wongsurawat[2,3,4], Piroon Jenjaroenpan[2,3], Iyarit Thaipisuttikul[5], Methee Chayakulkeeree[6], Chanwit Tribhuddarat[5], Perapon Nitayanon[5], Maleewan Theerawatthananon Peizner[5], Supathep Tansirichaiya[5], Witchuda Kamolvit[5]✉ & Komwit Surachat[1]✉

Bloodstream infections (BSIs) are life-threatening and require rapid, accurate pathogen characterization to guide antimicrobial therapy. Conventional culture-based diagnostics offer limited insight into the genetic basis of antimicrobial resistance (AMR) and virulence. In this study, we applied Oxford Nanopore Technology (ONT) metagenomic sequencing directly to 40 positive blood culture bottles collected at Siriraj Hospital, Thailand (2022 and 2025). Long-read data enabled species identification, AMR marker detection, virulence profiling, and plasmid replicon analysis. Diverse Gram-negative and Gram-positive pathogens were identified, including ESBL-producing *Escherichia coli*, carbapenem-resistant *Klebsiella pneumoniae*, *Enterococcus* spp., and *Staphylococcus* spp. Comprehensive genomic profiling revealed complex resistance mechanisms, multiple virulence factors related to adhesion, biofilm formation, and toxin production, and diverse plasmid types associated with horizontal gene transfer (HGT). This study demonstrates the value of ONT-based metagenomics as a faster workflow that is blood culture-dependent but subculture-independent, enabling species identification and AMR gene detection within 6–8 h, compared with 5–7 days for conventional methods, while supporting integrated genomic characterization for diagnostics, infection control, and regional AMR surveillance.

BSIs remain a major cause of morbidity and mortality in hospitalized patients. Characterization of their AMR profiles, is essential for guiding appropriate antimicrobial therapy and improving clinical outcomes[1]. Conventional laboratory workflows that rely on automated systems are limited in capturing the full genomic content underlying resistance, virulence, and mobile genetic elements (MGEs). Mixed infections, uncommon species, and novel resistance markers are often missed by standard phenotypic assays[2].

High-throughput sequencing has transformed clinical microbiology by providing genome-level insights into pathogen biology. Among sequencing technologies, ONT offers distinct advantages for clinical applications through long reads generated in real time, facilitating near-complete assembly of bacterial genomes and plasmids[3].

Metagenomic sequencing directly from positive blood culture bottles has emerged as a powerful approach for comprehensive genomic profiling of bloodstream pathogens. Beyond faster species identification, this method allows simultaneous detection of resistance genes, virulence factors, and plasmid replicons within a single workflow[4]. In this study, we applied long-read metagenomic sequencing using ONT to positive blood cultures collected at Siriraj Hospital, Thailand. Our objectives were to characterize species diversity, AMR markers,

[1]Department of Biomedical Sciences and Biomedical Engineering, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand. [2]Siriraj Long-Read Lab (Si-LoL), Division of Medical Bioinformatics, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [3]Oxford Nanopore Centre of Excellence, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [4]Division of Medical Bioinformatics, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [5]Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [6]Division of Infectious Diseases and Tropical Medicine, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. ✉email: witchuda.kam@mahidol.ac.th; komwit.s@psu.ac.th

virulence-associated genes, and plasmid content among bloodstream pathogens, thereby demonstrating the applicability of ONT-based metagenomics for integrated genomic surveillance and enhanced diagnostics in a tertiary-care clinical setting. Moreover, the ability of ONT sequencing to rapidly generate genomic data within hours of blood culture positivity highlights its potential to shorten time to pathogen identification and antimicrobial resistance inference compared with conventional diagnostic workflows, which typically require several days for definitive identification and antimicrobial susceptibility testing.

Materials and methods

Sample collection

Forty positive blood culture bottles were collected during two independent periods (2022 and 2025) from the Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand. Blood samples were collected using standard aseptic techniques and inoculated into aerobic and anaerobic blood culture bottles according to routine hospital protocols. Bottles were incubated and continuously monitored using automated blood culture systems (BACTEC™, BD Diagnostics, and BacT/ALERT®, bioMérieux), with time to positivity varying by bacterial species and inoculum size. Metagenomic sequencing was performed only after bottles flagged positive and was evaluated in parallel with the routine laboratory workflow applied to the same positive cultures.

Sampling across two years allowed inclusion of temporally distinct isolates, capturing possible shifts in pathogen composition, resistance patterns, and strain diversity over time. This approach also served to validate the reproducibility and long-term applicability of the ONT-based metagenomic workflow. To avoid duplication, bottles originating from the same patient were excluded, and all samples were anonymized prior to processing. From each positive bottle, Gram staining was performed, followed by species identification and antimicrobial susceptibility testing (AST) using the VITEK® 2 system and disk diffusion, following CLSI M100—35th Edition guidelines[5]. In polymicrobial blood culture samples, antimicrobial susceptibility testing was performed independently for each recovered isolate, and susceptibility categories were interpreted on an isolate-by-isolate basis rather than at the sample level. All methods were performed in accordance with relevant institutional guidelines, national regulations, and biosafety standards.

DNA extraction for Oxford Nanopore sequencing

Ten milliliters of culture fluid were collected directly from positive bottles. Human cells were removed by centrifugation (1600 rpm, 5 min), followed by bacterial pelleting at 8000 rpm for 15 min. Genomic DNA was extracted using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, Irvine, CA, USA). DNA concentration was measured using a Qubit™ 4 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), while purity was assessed with a NanoDrop™ One spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and integrity was verified by agarose gel electrophoresis.

Oxford Nanopore DNA sequencing

Samples from 2022 and 2025 were sequenced on R9.4.1 and R10.4.1 flow cells, respectively. Libraries were prepared using ONT's Rapid Barcoding Kit (SQK-RBK004) and Native Barcoding Kit V14 (SQK-NBD114.96), with six and twelve samples multiplexed per run. Sequencing was performed on a MinION device using MinKNOW software (versions current at the time of sequencing in 2022 and 2025) in high-accuracy mode for up to 72 h, and base calling was done with Guppy v6.5.7 ($Q \geq 8$). Reads shorter than 500 bp were excluded.

For turnaround time assessment, T0 was defined as the time point at which the blood culture bottle flagged positive in the automated blood culture system. All sequencing-derived time points reported in this study are referenced to T0 and reflect the post–blood culture diagnostic phase. Although sequencing runs were continued for up to 72 h, preliminary taxonomic assignment and species identification with AMR gene detection were achievable within the first hours of sequencing.

Bioinformatics and data analysis

Reads were basecalled and quality filtered, and all reads generated per sample, including those from rapid and ligation-based library preparations were merged prior to de novo assembly using metaFlye v3.0 with default parameters[6]. Genomes were reconstructed primarily at the contig level, with assembly completeness varying by sequencing depth and sample complexity. Assembly quality was assessed using QUAST v5.2[7]. Genome assemblies varied in completeness, ranging from fragmented contigs to fully circularized chromosomes. Contigs with insufficient coverage or ambiguous taxonomic assignment were excluded from downstream analyses (Supplementary Table 1). Taxonomic classification was conducted using the Kraken2 and GTDB-Tk pipelines, and the resulting contigs were further validated by BLASTn analysis. Human-derived sequences were assessed and excluded through this combined approach: contigs were screened using Kraken2 (with the standard RefSeq bacterial database) for taxonomic classification and GTDB-Tk (Supplementary Table 2), followed by BLASTn searches against the NCBI database. Contigs were additionally manually inspected in Geneious software to confirm bacterial origin and to remove any sequences mapping to the human genome prior to downstream analyses[8–10]. AMR and virulence genes were identified using ABRicate V.1.0.1 with the CARD and VFDB databases ($\geq 80\%$ coverage and identity)[11]. Plasmid replicons were detected using PlasmidFinder[12]. Data visualization was performed using the ChiPlot online platform (https://www.chiplot.online/, accessed on 2 October 2025) and in-house Python scripts.

Results

Routine identification using VITEK® 2 and antimicrobial susceptibility profiles

A total of 45 bacterial isolates were recovered from 40 positive blood culture samples by routine laboratory workflows, indicating polymicrobial infections in three cases (Si10, Si20, and Si32), whereas metagenomic sequencing was performed directly on blood culture material without subculturing. These samples contained combinations such as *Enterococcus faecium* with *Stenotrophomonas maltophilia* (Si20) and *Escherichia coli*, *K. pneumoniae*, and *Aeromonas veronii* (Si32). Routine identification revealed a diverse pathogen spectrum dominated by Enterobacterales, particularly *E. coli* and *K. pneumoniae*, together accounting for nearly 40% of isolates. Polymicrobial growth was observed in three samples, highlighting the complexity of bloodstream infections encountered in routine diagnostics. Less frequent species included *Klebsiella aerogenes*, *Aeromonas veronii* biovar sobria, *Serratia odorifera*, *Campylobacter* spp., *Bacillus cereus*, *Staphylococcus lugdunensis*, and coagulase-negative staphylococci (CoNS). Additionally, three isolates could not be confidently identified to species level, including two Gram-negative rods, non-fermentative (Si10_1 and Si40) and one Gram-negative rod (Si10_2).

Phenotypically, multidrug resistance (MDR) was most frequent among *K. pneumoniae*, *E. coli*, and *A. baumannii*. β-lactam resistance was widespread across the Enterobacterales. Several *E. coli* and *K. pneumoniae* isolates exhibited resistance to third-generation cephalosporins, consistent with extended-spectrum β-lactamase (ESBL) production. One *K. pneumoniae* isolate (Si34) showed resistance to carbapenems, indicating the presence of a possible carbapenemase. Among non-fermenting Gram-negative bacilli, *A. baumannii* Si09 displayed an MDR phenotype, whereas *Pseudomonas aeruginosa* and *S. maltophilia* remained susceptible to all tested antibiotics. Gram-positive isolates demonstrated variable resistance patterns: some *Enterococcus* isolates were β-lactam–resistant but all remained susceptible to vancomycin, whereas MSSA remained broadly susceptible except for a single tetracycline-resistant isolate (Si36) (Fig. 1).

Turnaround time and sequencing workflow performance

Hands-on time for DNA extraction and library preparation was approximately 0.5–1 h, while sequencing and real-time analysis were performed instrument-side. Preliminary taxonomic assignments were obtained within approximately 1 h of sequencing initiation, species-level classification within 2–4 h and AMR gene detection within 6–8 h. Sequencing runs were extended up to 72 h to enhance genome coverage and assembly continuity, enable detection of low-abundance organisms, and improve reconstruction of plasmid-associated sequences; however, extended sequencing primarily enhanced genomic completeness and contextual analyses rather than altering initial species identification or major AMR predictions.

Species identification and taxonomic classification using metagenomic sequencing

ONT-based metagenomic sequencing was performed on all 40 positive blood culture samples. While routine culture yielded 45 phenotypic isolates, genomic analysis recovered 43 distinct bacterial genomes. Two isolates from polymicrobial samples were genetically indistinguishable from co-isolated strains and were therefore merged during assembly. ONT sequencing enabled species identification and AMR prediction within 6–8 h, despite continuing runs up to 72 h for maximal assembly completeness—representing a substantial reduction compared with the 5–7 days required for culture-based workflows.

Taxonomic assignment using Kraken2, GTDB-Tk, and BLASTn identified 18 bacterial taxa. Metagenomic classification confirmed Enterobacterales as the dominant group, while additionally resolving taxa not detected or misclassified by routine workflows. Rare or unexpected species included *Ralstonia mannitolilytica*, *A. veronii*, *Campylobacter jejuni*, and *Bacillus cereus*. Notably, ONT-based genomic analysis uniquely identified *R. mannitolilytica* and *Sphingomonas* sp., which were not detected by conventional methods, and reclassified two CoNS as *Staphylococcus aureus* and *S. epidermidis*. In contrast, one isolate reported as *E. faecium* by VITEK® could not be resolved by sequencing due to insufficient genomic evidence (Fig. 2).

Antimicrobial resistance (AMR) and virulence associated gene profiles

Enterobacterales

Across the 19 Enterobacterales genomes, AMR and virulence profiles were heterogeneous but showed clear species-level trends (Fig. 3). Efflux- and transporter-associated resistance mechanisms predominated overall, while β-lactamases and aminoglycoside-modifying enzymes were widely distributed. *E. coli* genomes were enriched for macrolide resistance markers and efflux-related genes, whereas *K. pneumoniae* exhibited broader multidrug resistance, particularly involving β-lactams, aminoglycosides, and efflux systems. In contrast, *P. mirabilis* and *S. marcescens* displayed comparatively limited resistance repertoires.

Virulence-associated genes were more abundant in *E. coli* and *K. pneumoniae*, largely reflecting functions related to adhesion/biofilm formation, secretion systems, and iron acquisition. By comparison, *P. mirabilis* and *S. marcescens* encoded few virulence factors, primarily linked to motility and surface polysaccharide biosynthesis (Supplementary Table 3).

Non-fermenting Gram-negative bacilli

Among non-fermenters, *A. baumannii* Si09 displayed the most extensive AMR profile, encompassing multiple β-lactamases, aminoglycoside resistance markers, and additional resistance classes, consistent with a multidrug-resistant phenotype (Fig. 4). The three *P. aeruginosa* genomes showed highly similar resistance architectures dominated by multidrug efflux systems, regulatory elements, and intrinsic β-lactamases.

Virulence gene analysis revealed a marked contrast between species. *A. baumannii* encoded a broad array of factors associated with biofilm formation, iron acquisition, and secretion systems, whereas *P. aeruginosa* carried

Fig. 1. Antimicrobial susceptibility profiles of bacteria recovered from positive blood culture samples. Susceptibility categories (susceptible, intermediate, resistant) were determined using routine phenotypic AST (VITEK® 2 and/or disk diffusion) following CLSI M100—35th Edition guidelines. AST results of polymicrobial blood culture samples are shown separately for each phenotypically isolated organism.

an even larger virulence repertoire, including quorum sensing, motility, and multiple secretion pathways, reflecting its well-characterized pathogenic versatility.

ONT-based analysis additionally detected *R. mannitolilytica* (Si40) and *Sphingomonas* sp. (Si20). Given their reported association with environmental or reagent contamination, these findings should be interpreted cautiously. Correspondingly, these taxa harbored few or no AMR or virulence-associated genes, supporting conservative interpretation of their clinical relevance (Supplementary Table 4).

Gram-positive cocci

Gram-positive genomes exhibited generally lower AMR and virulence burdens than Gram-negative organisms, with notable species-level differences (Fig. 5). *E. faecalis* Si08 encoded resistance genes primarily targeting

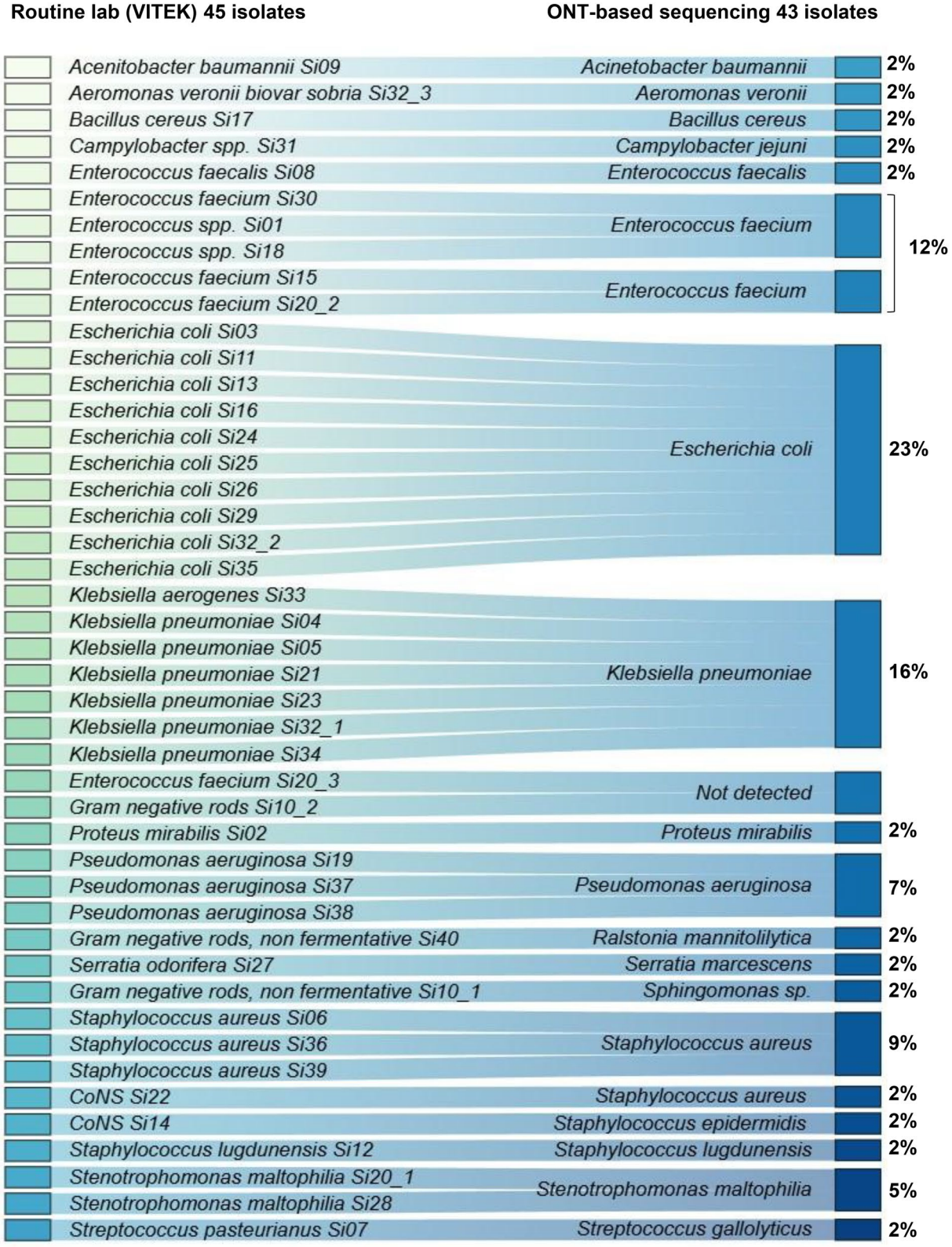

Fig. 2. Comparison of bacterial species identification between the routine clinical laboratory (VITEK) and Oxford Nanopore technology (ONT)-based sequencing. Each line represents the correspondence between the bacterial species identified by conventional VITEK biochemical profiling (left) and the species identified by WGS-based analysis (right). The width of each line reflects the relative proportion of isolates within the dataset. Percentages on the right indicate the relative frequency of each species detected by ONT-based sequencing.

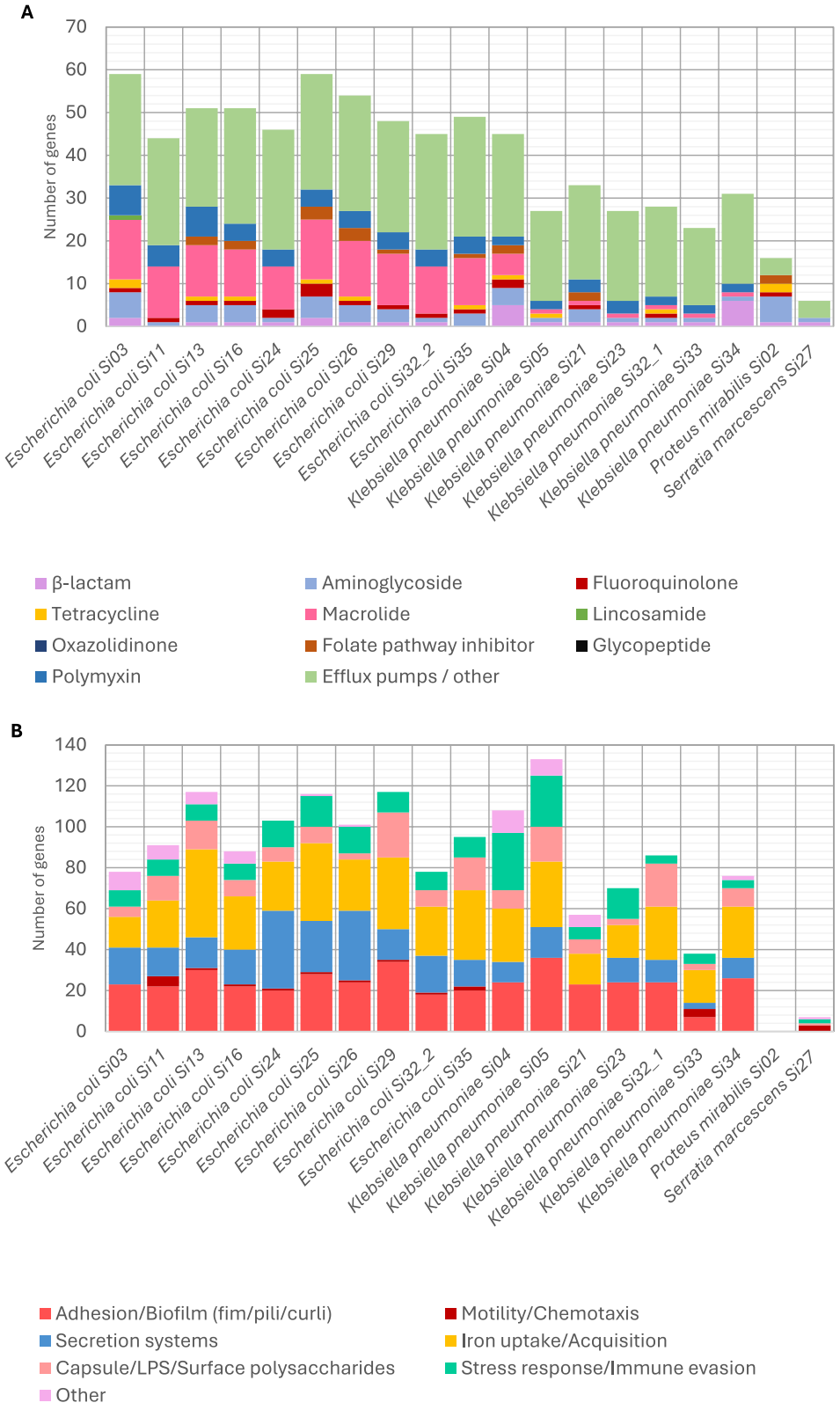

Fig. 3. The graph reveals the identification of antimicrobial resistance (**A**) and virulence genes (**B**) in Enterobacterales isolates of this study. Gene counts reflect the presence of annotated genes identified by in silico analysis and do not indicate gene expression or phenotypic activity.

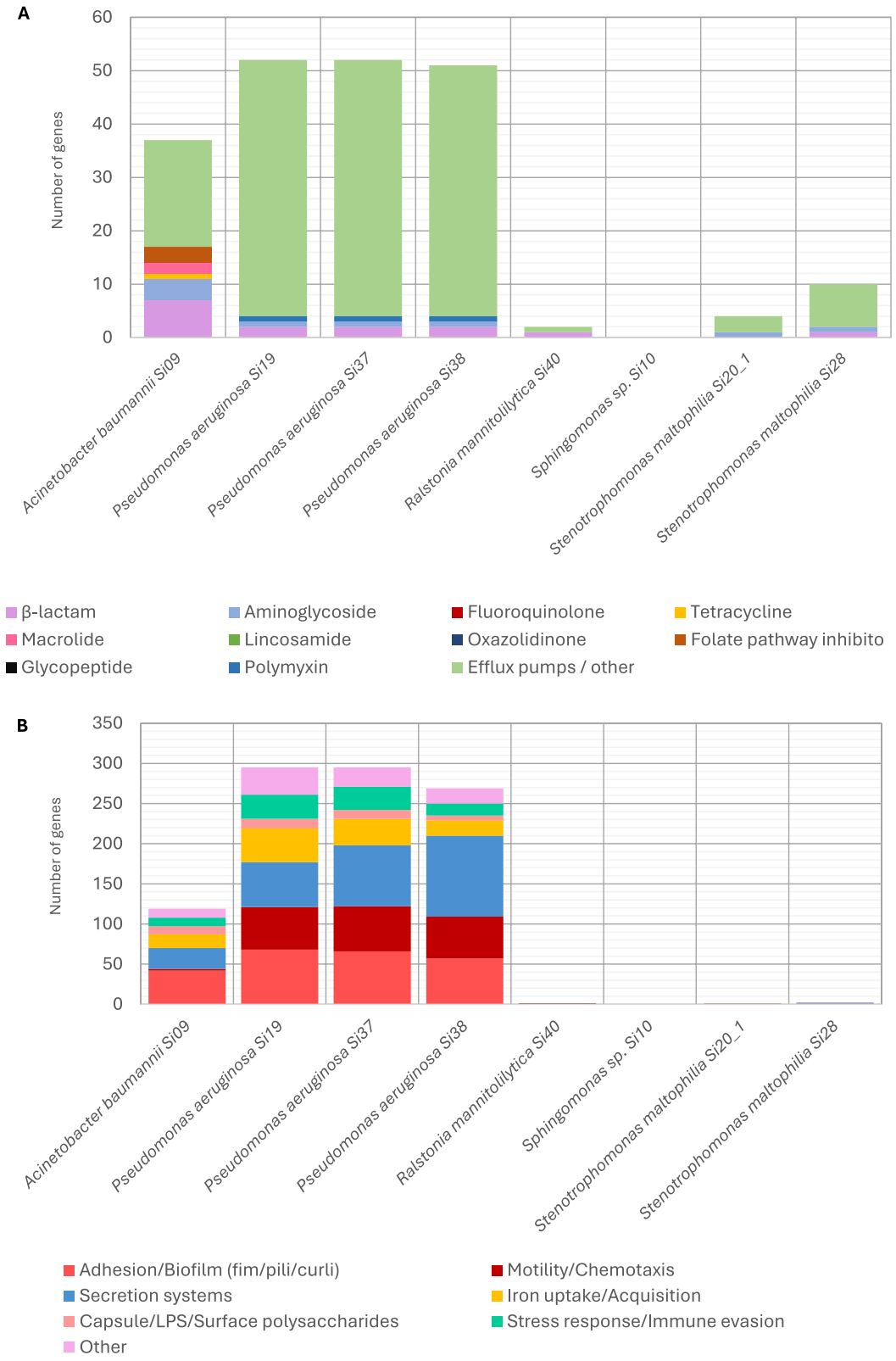

Fig. 4. The graph reveals the identification of antimicrobial resistance (**A**) and virulence genes (**B**) in non-fermenting Gram-negative isolates of this study. Gene counts reflect the presence of annotated genes identified by in silico analysis and do not indicate gene expression or phenotypic activity.

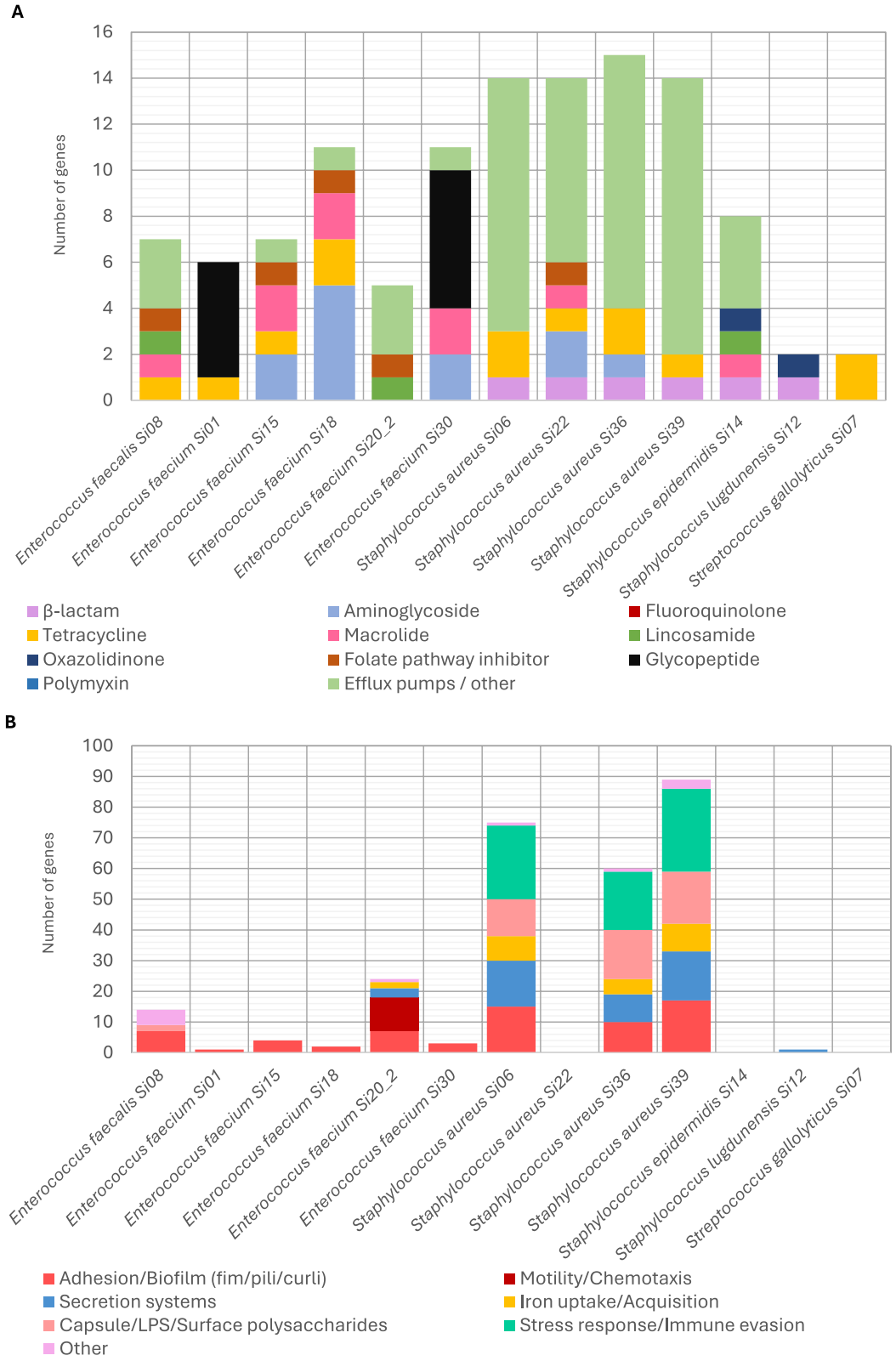

Fig. 5. The graph reveals the identification of antimicrobial resistance (**A**) and virulence genes (**B**) in Gram-positive cocci isolates of this study. Gene counts reflect the presence of annotated genes identified by in silico analysis and do not indicate gene expression or phenotypic activity.

tetracyclines and macrolides and possessed virulence factors related to adhesion and biofilm formation. *E. faecium* genomes showed variable AMR content, while virulence gene repertoires were limited in most isolates, except for Si20_2, which encoded a broader range of functions spanning adhesion, secretion, and stress adaptation.

S. aureus genomes carried consistent resistance markers, mainly targeting β-lactams and tetracyclines, but exhibited substantial variability in virulence gene content. In contrast, coagulase-negative staphylococci and *Streptococcus gallolyticus* encoded few resistance or virulence-associated genes (Supplementary Table 5).

Unusual Gram-negative bacilli

Uncommon bloodstream-associated taxa exhibited relatively restricted AMR and virulence profiles. *A. veronii* Si32_3 carried only a single tetracycline resistance marker alongside adhesion-related virulence genes. *B. cereus* Si17 encoded a modest number of resistance genes and a diverse but limited virulence repertoire. In contrast, *C. jejuni* Si31 harbored multiple efflux-associated resistance markers and a broad virulence gene set dominated by motility, surface structures, and secretion-related functions, consistent with its known invasive potential (Fig. 6; Supplementary Table 6).

Plasmid identification in metagenome-derived bacterial genomes

Plasmids identified in Gram-negative isolates

Seventeen Gram-negative bacterial genomes reconstructed from positive blood culture samples were screened for plasmid replicons using PlasmidFinder. In *E. coli*, IncF-type plasmids predominated, particularly IncFIB-AP001918, detected in multiple strains with 100% sequence identity. Other IncF variants, including IncFIA, IncFIB-H89, IncFII-Coo, and IncFIC-FII, highlighted the diversity and widespread distribution of IncF plasmids. *K. pneumoniae* isolates displayed greater heterogeneity, harboring Col-type, Inc-type, and unclassified plasmids (pENTAS02, pESA2, repB-KlebVIR), indicating both common and species-specific replicons. *P. mirabilis* Si02 carried a unique Col3M plasmid, whereas *S. marcescens* Si27 harbored IncFII-ECLA and IncR replicons, reflecting interspecies variation and potential HGT. Thirty-four distinct plasmid types were identified across Gram-negative blood culture–derived genomes. Certain replicons appeared species-specific (IncFIB-AP001918 in *E. coli*, IncFII-KP91 in *K. pneumoniae*), while others, such as Col-pHAD28, ColRNAI, and IncR, were shared across multiple species, highlighting complex plasmid dynamics among Enterobacterales detected in bloodstream infections (Fig. 7).

Plasmids identified in Gram-positive isolates

Twelve Gram-positive blood culture–derived genomes harbored diverse Rep-family plasmids. Among *Enterococcus* (n = 5), *E. faecium* strains Si15, Si18, and Si30 carried multiple replicons, including Rep14b-pRI1, Rep18a-p200B, Rep18b-pE1p13, and RepUS15-pNB2354, with Si15 showing the highest diversity (eight replicons). *S. aureus* isolates (Si06, Si22, Si36, Si39) displayed varied plasmid content; Si22 contained four replicons (Rep10-pNE131, Rep20-SAP105B, RepUS23-SAP099B, RepUS46-SAP099B), Si39 three, and Si06/Si36 two each. *S. epidermidis* Si14 carried three replicons, *S. lugdunensis* Si12 two, and *B. cereus* Si17 one (RepUS47-pLFE1), reflecting both inter- and intra-species plasmid variation in Gram-positive bacteria detected in positive blood culture samples (Fig. 8).

Discussion

This study demonstrates the applicability of ONT-based sequencing for comprehensive metagenomic characterization of bloodstream pathogens directly from positive blood culture samples in a tertiary-care hospital in Thailand. Samples collected during two periods (2022 and 2025) captured temporal variation in pathogen diversity and antimicrobial resistance profiles within the same institution, enhancing dataset representativeness and demonstrating the reproducibility and long-term applicability of ONT-based metagenomic sequencing for hospital genomic surveillance. Although no statistically meaningful shifts in pathogen composition or resistance profiles were observed between periods, the inclusion of temporally distinct samples highlights workflow robustness and supports future longitudinal surveillance applications.

A total of 45 isolates were recovered from 40 positive blood culture samples through routine laboratory workflows, reflecting the presence of polymicrobial infections in three cases (Si10, Si20, and Si32). Enterobacterales dominated the collection, with *E. coli* and *K. pneumoniae* together accounting for nearly 40% of isolates, consistent with regional and global BSI epidemiology[13–15]. Most *E. coli* and *K. pneumoniae* isolates exhibited resistance to third-generation cephalosporins and β-lactam/β-lactamase inhibitor combinations, consistent with the presence of ESBL production. One *K. pneumoniae* isolate (Si34) was resistant to carbapenems, suggesting carbapenemase activity, while *P. mirabilis* Si02 and *S. odorifera* Si27 displayed resistance to aminoglycosides and β-lactams. Non-fermenting Gram-negative bacilli—including *P. aeruginosa*, *A. baumannii*, and *S. maltophilia*—represented a substantial fraction of isolates and are well known for their intrinsic and acquired MDR mechanisms that complicate clinical management. Among Gram-positive organisms, *E. faecium* exhibited β-lactam resistance but remained susceptible to vancomycin and linezolid, while MSSA and CoNS remained largely sensitive to tested agents. Several isolates, including Si10 and Si40, lacked AST data, likely due to contamination or growth interference. Less frequent species such as *A. veronii* and *Campylobacter* spp. emphasize the microbial diversity of BSIs and the value of molecular methods for accurate detection. However, it should be noted that this study did not include dedicated negative controls, such as extraction or library blanks, which represents a limitation when interpreting low-abundance or unexpected taxa detected by metagenomic sequencing. Environmental or water-associated bacteria, including *Ralstonia* spp. and *Sphingomonas* spp., have been reported as potential reagent- or laboratory-associated contaminants in sequencing workflows. Accordingly, although *R. mannitolilytica* Si40 and *Sphingomonas* sp. Si20 were detected in positive blood culture samples and exhibited distinct genomic features,

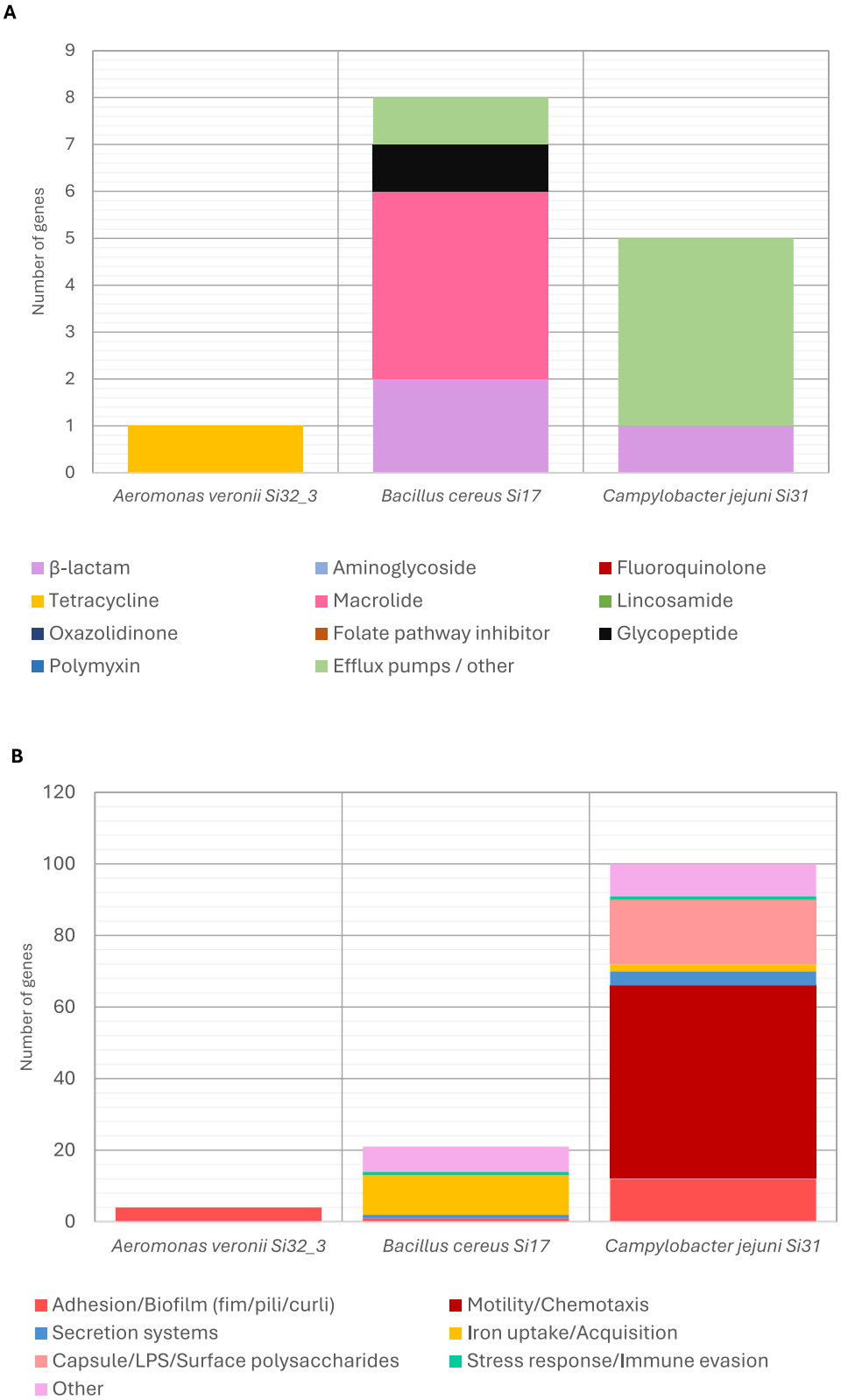

Fig. 6. The graph reveals the identification of antimicrobial resistance (**A**) and virulence genes (**B**) in unusual Gram-negative bacilli isolates of this study. Gene counts reflect the presence of annotated genes identified by in silico analysis and do not indicate gene expression or phenotypic activity.

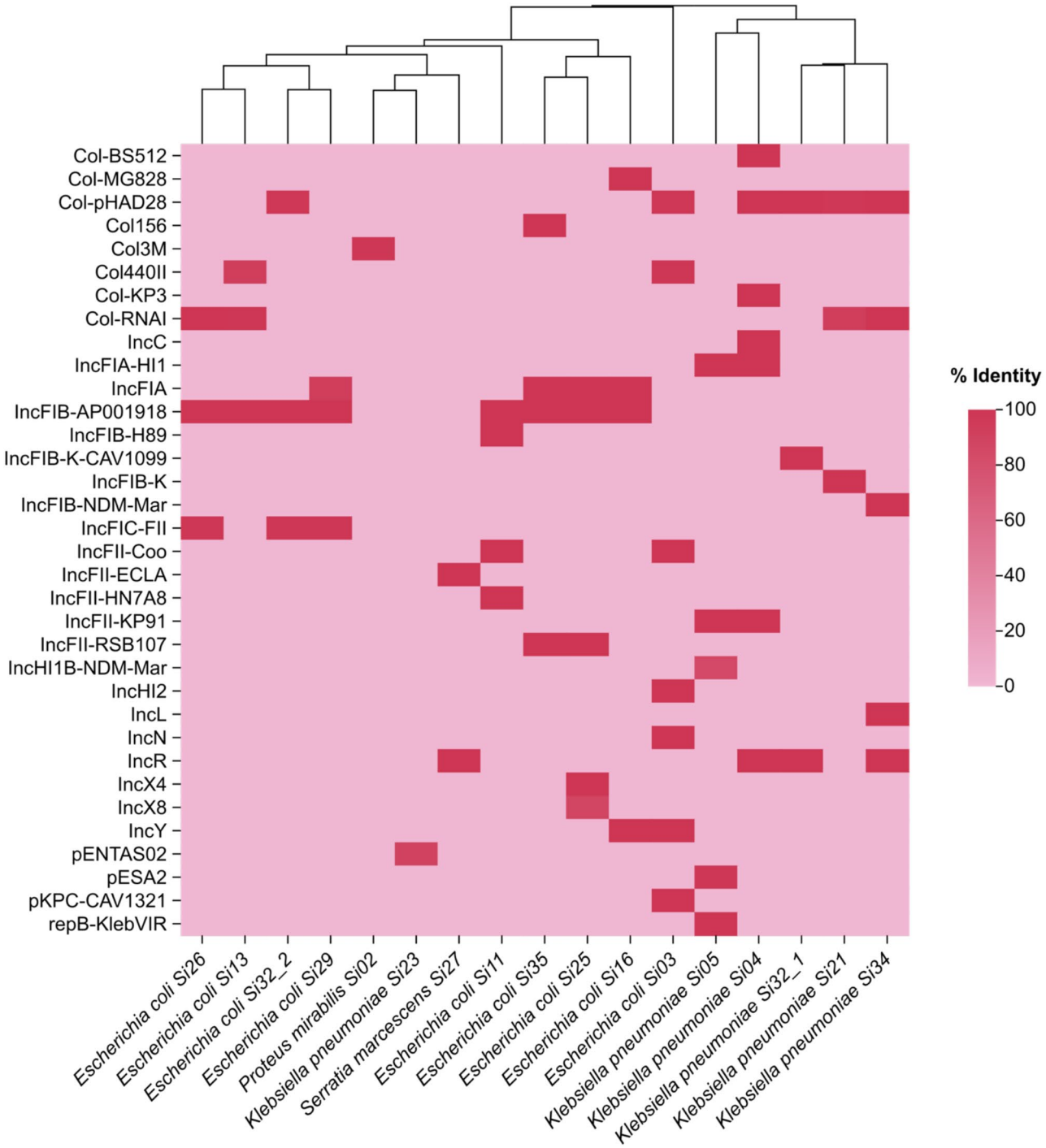

Fig. 7. Heatmap showing the distribution of plasmid replicon types identified using PlasmidFinder among Gram-negative bacterial isolates collected in this study. Each column represents a bacterial isolate, and each row represents a detected plasmid replicon type. The color intensity indicates the presence (red) or absence (light pink) of the corresponding replicon within each isolate.

their clinical relevance cannot be conclusively established based on metagenomic data alone and should be interpreted with caution[16,17]. The low abundance of AMR and virulence markers in these taxa further supports cautious interpretation. Future studies incorporating systematic negative controls and quantitative thresholds will be essential to distinguish true bloodstream pathogens from background or contaminant taxa.

Metagenomic sequencing was subsequently performed on the same 40 samples without prior subculturing, yielding 43 bacterial genomes. Taxonomic classification using Kraken2, GTDB-Tk, and BLASTn revealed some discrepancies with conventional identification, mainly due to methodological and database differences. For instance, sample Si10, identified as containing two non-fermenting Gram-negative rods by routine workflows,

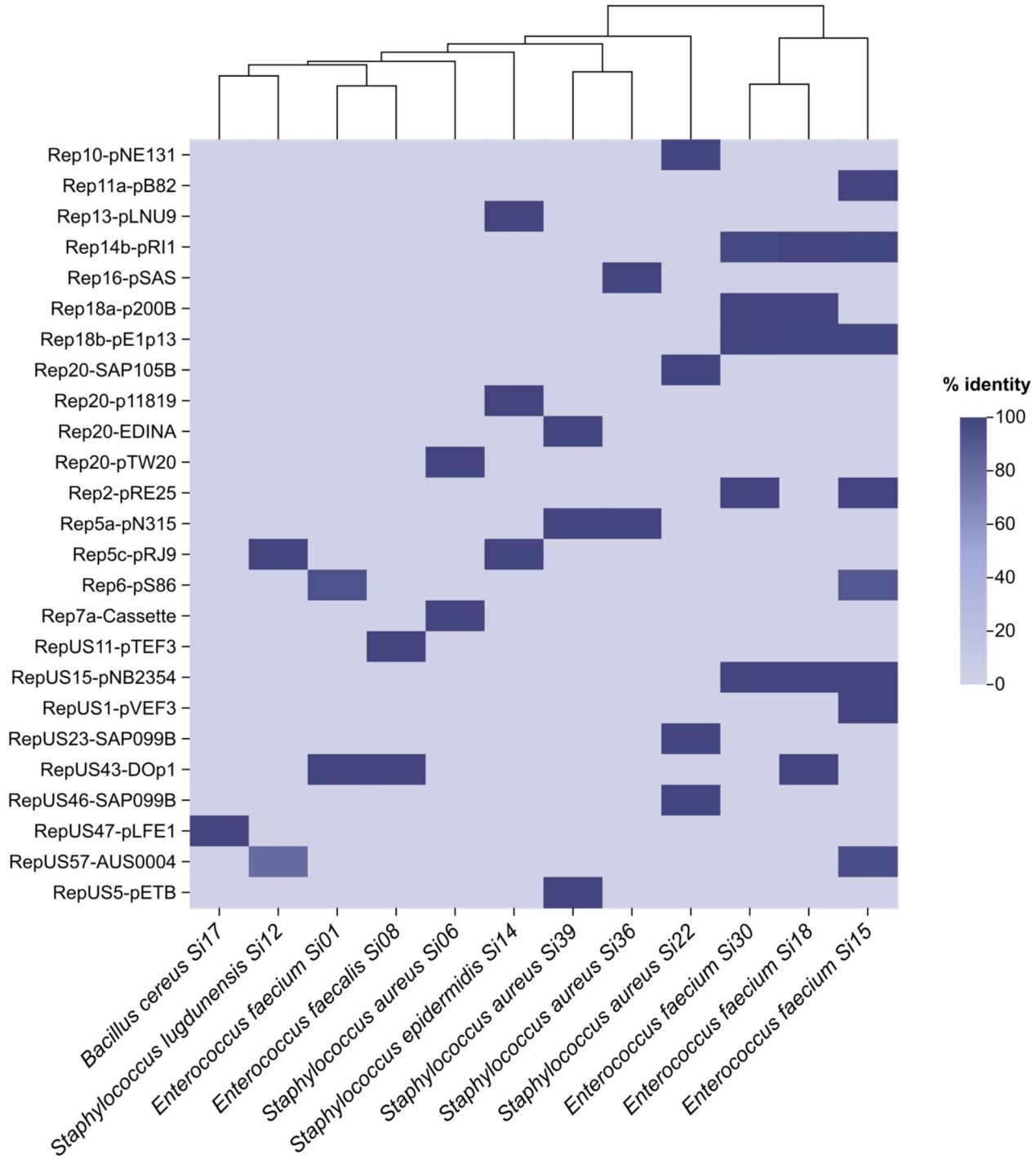

Fig. 8. Heatmap showing the distribution of plasmid replicon types identified using PlasmidFinder among Gram-positive bacterial isolates collected in this study. Each column represents a bacterial isolate, and each row represents a detected plasmid replicon type. The color intensity indicates the presence (dark violet) or absence (light violet) of the corresponding replicon within each isolate.

was assigned to a single *Sphingomonas* sp. genome by ONT metagenomics. Similarly, sample Si20, which was reported by routine workflows as containing *S. maltophilia* and two *E. faecium* isolates with differing AST profiles, was resolved by ONT-based sequencing as one *S. maltophilia* genome and one *E. faecium* genome. These discrepancies illustrate a recognized limitation of metagenomic sequencing—for example, the difficulty in resolving multiple closely related species in mixed samples due to uneven read representation, shared genomic sequences and read-mapping ambiguities. Previous studies have shown that when coverage is low or species are highly similar, classification tools may fail to separate co-occurring taxa or may assign reads to only one species[18,19]. Nonetheless, genomic-based classification remains more accurate at the nucleotide level and less

dependent on phenotypic variation than conventional biochemical tests. For example, genome-based analysis confirmed *E. faecium* for Si01 (previously reported as *Enterococcus spp.* by VITEK® 2), and *C. jejuni* Si31 was correctly classified beyond genus level, demonstrating the added value of genomics for surveillance-oriented pathogen characterization and AMR tracking.

Comparison of AMR profiles revealed both concordance and quantifiable discordance between phenotypic AST and genomic predictions. For carbapenems, genomic data identified carbapenemase genes in 4 isolates (*K. pneumoniae* Si04 carrying $bla_{\text{OXA-232}}$; *K. pneumoniae* Si34 carrying $bla_{\text{NDM-1}}$ and $bla_{\text{OXA-48}}$; *A. baumannii* Si09 carrying $bla_{\text{NDM-5}}$ and $bla_{\text{OXA-23}}$; and *R. mannitolilytica* Si40 carrying $bla_{\text{OXA-443}}$), all of which are consistent with phenotypic carbapenem resistance. In contrast, 3 isolates of *P. aeruginosa* lacked acquired carbapenemase genes despite reduced carbapenem susceptibility, likely attributable to non-enzymatic mechanisms such as overexpression of multidrug efflux pumps (e.g., MexAB-OprM, MexXY) and chromosomal AmpC β-lactamase activity.

For extended-spectrum β-lactam resistance, ESBL genes ($bla_{\text{CTX-M-15}}$ and $bla_{\text{CTX-M-55}}$) were detected in three *E. coli* and two *K. pneumoniae* genomes, showing strong concordance with phenotypic resistance to third-generation cephalosporins. ONT-based metagenomic sequencing additionally identified resistance markers not routinely captured by AST, including efflux systems and minor aminoglycoside-modifying enzymes, as well as clinically relevant β-lactamases such as $bla_{\text{NDM-5}}$ in *A. baumannii* Si09 and $bla_{\text{OXA-605}}$ in *C. jejuni* Si31, highlighting the added value of metagenomics for contextual AMR characterization[20]. Conversely, a limited number of phenotypic resistances were not explained by major resistance genes, plausibly due to gene expression variability[21], regulatory mutations and other phenotypic resistance mechanisms[22], or limitations related to fragmented metagenomic assemblies[23]. Together, these findings demonstrate that ONT-based metagenomics enables faster and concordant genomic inference of AMR profiles while also revealing biologically meaningful discordances.

Our genomic findings align with recent regional and global reports on MDR Gram-negative pathogens. We observed widespread ESBL-associated genes ($bla_{\text{CTX-M}}$, bla_{TEM}) in *E. coli*, diverse carbapenemases ($bla_{\text{NDM-1}}$, $bla_{\text{OXA-48}}$) in *K. pneumoniae*, extensive resistance and biofilm-related genes in *A. baumannii*, and prominent efflux and secretion systems in *P. aeruginosa*. Plasmid- and insertion element-mediated dissemination of $bla_{\text{CTX-M}}$ and related ESBLs has been repeatedly documented and accounts for much of the current ESBL burden[24]. Carbapenemase genes such as bla_{NDM} and $bla_{\text{OXA-48}}$ are increasingly reported in *K. pneumoniae* and represent a major driver of clinical treatment failure in the region[25]. The persistence of *A. baumannii* in hospital environments is well described and reflects the combination of carbapenemases, aminoglycoside-resistance methylases (e.g., *armA*), and biofilm gene markers that promote nosocomial transmission[26]. Likewise, the central role of RND-family efflux pumps and type III secretion effectors in *P. aeruginosa* virulence and intrinsic MDR aligns with prior mechanistic reviews[27].

Methodologically, metagenomic assembly and species resolution can be imperfect in mixed or low-coverage samples due to uneven species abundance, conserved inter-species regions (including shared MGEs), and limited coverage of minor community members, which can fragment assemblies or collapse closely related genomes[28]. Nonetheless, ONT long reads improve assembly contiguity and plasmid context, strengthening inference about AMR gene carriers and horizontal gene transfer even when mixed-sample complexity limits full strain resolution[29,30].

Our plasmid-analysis findings reflect the growing body of evidence that epidemic plasmids play a central role in resistance gene dissemination. For instance, IncF-type plasmids have been shown to drive the global success of MDR *E. coli* lineages and mediate spread of ESBL genes such as $bla_{\text{CTX-M}}$[31]. In *K. pneumoniae*, heterogeneous Inc and Col plasmids frequently carry carbapenemase gene markers (e.g., bla_{NDM}, bla_{OXA}) and virulence loci, thereby facilitating horizontal gene transfer and convergent evolution of resistance and virulence[32,33]. The detection of Rep-family plasmids among Gram-positive isolates such as *E. faecium* and *S. aureus* is consistent with prior studies highlighting distinct plasmid backbones in hospital-adapted lineages. The absence of detectable plasmids in 14/43 isolates may reflect chromosomal integration of AMR genes or the presence of novel plasmid scaffolds not yet covered in current databases.

Accurate and timely diagnosis of bloodstream infections is a key component of clinical microbiology practice. Conventionally, this diagnostic process involves three sequential steps: detection of bacterial growth, taxonomic identification, and AST. Detection in blood culture bottles may take up to five days but typically occurs within the first 24 h of incubation. Subsequent bacterial identification requires an additional 24 h, followed by 4–24 h for AST on pure colonies[34]. In contrast, our ONT-based metagenomic workflow markedly reduced this timeline—preliminary read classification was achievable within the first hour of sequencing, and species-level identification with AMR gene prediction could be completed within 6–8 h, even though sequencing runs were extended to 72 h to improve genome coverage, assembly contiguity, recovery of low-abundance taxa, and MGE context. This fast turnaround is consistent with findings from Kamau and Yang, who demonstrated that nanopore metagenomic sequencing of positive blood culture fluids accurately identified bacterial and fungal pathogens within a single working day[35]. Such accelerated diagnostics have the potential to support earlier clinical decision-making and infection control measures, although this study did not directly evaluate clinical outcomes. Beyond speed, direct metagenomic sequencing preserves the native microbial composition of bloodstream infections, avoiding potential biases introduced by subculturing, such as overgrowth of dominant species or loss of fastidious or minor taxa[36]. Therefore, evaluation of this approach alongside routine workflows may enhance diagnostic resolution and support future genomic surveillance efforts for bloodstream infections in hospital settings.

Conclusion

This pilot study highlights the feasibility and clinical relevance of integrating ONT–based metagenomic sequencing into bloodstream infection diagnostics in Thailand. The approach enables comprehensive genomic profiling—encompassing species identification, antimicrobial resistance, virulence, and plasmid content—directly from positive blood cultures without subculture. By complementing conventional methods, ONT metagenomics provides a scalable platform for genomic surveillance, antimicrobial stewardship, and real-time genomic surveillance frameworks. Broader adoption could strengthen outbreak detection and regional AMR surveillance across Southeast Asia.

Limitations

This study is limited by its small sample size and pilot design, which primarily assess workflow feasibility. Metagenomic sequencing may underrepresent polymicrobial infections due to read imbalance and sequence similarity among related taxa. Detection of AMR and virulence genes depended on existing databases (CARD, VFDB, PlasmidFinder), potentially missing novel local variants. Additionally, the lack of clinical correlation limits assessment of diagnostic impact. We did not perform a formal cost-effectiveness analysis; therefore, the economic feasibility relative to conventional diagnostics was not assessed. In addition, sequencing was conducted on positive blood culture material rather than directly from whole blood, and the performance of direct-from-blood metagenomic sequencing was beyond the scope of this study. Although different ONT library preparation kits and flow cell chemistries were used across sequencing periods, all samples were processed using a fully standardized bioinformatic pipeline; however, minor chemistry-related effects on read length distribution or yield cannot be completely excluded. Finally, incremental gains in AMR detection beyond the 6–8-h time frame were not formally quantified. Future work integrating patient metadata and larger cohorts will be necessary to validate the predictive and epidemiological utility of this workflow.

Data availability

Sequence data was made publicly available in NCBI databases (BioProject no. PRJNA1354200).

Received: 7 December 2025; Accepted: 18 February 2026
Published online: 07 March 2026

References

1. Khanal, B., Shrestha, L. B., Sharma, A. & Siwakoti, S. Bloodstream infections: Trends in etiology and antimicrobial resistance in 10 years in Eastern Nepal (in Eng). *BMC Infect. Dis.* **25**(1), 1001 (2025).
2. A. M. Nafea *et al.* Application of next-generation sequencing to identify different pathogens (in English). *Front. Microbiol.* 14 (2023).
3. W. Zhao *et al.* Oxford nanopore long-read sequencing enables the generation of complete bacterial and plasmid genomes without short-read sequencing (in English). *Front. Microbiol.* Volume 14 (2023).
4. Kamau, E. & Yang, S. Metagenomic sequencing of positive blood culture fluid for accurate bacterial and fungal species identification: A pilot study. *Microorganisms* https://doi.org/10.3390/microorganisms11051259 (2023).
5. C. a. L. S. Institute, J. S. L. I. (Chairholder), Ed. *Performance standards for antimicrobial susceptibility testing* (CLSI supplement M100). Wayne, Pennsylvania: *Clinical and Laboratory Standards Institute*, 2025.
6. Kolmogorov, M. et al. metaFlye: Scalable long-read metagenome assembly using repeat graphs. *Nat. Methods.* **17**(11), 1103–1110 (2020).
7. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. "QUAST: Quality assessment tool for genome assemblies," (in eng). *Bioinformatics* **29**(8), 1072–1075 (2013).
8. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2 (in eng). *Genome Biol.* **20**(1), 257 (2019).
9. Johnson, M. et al. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **36**(suppl_2), W5–W9 (2008).
10. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. Gtdb-tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* **36**(6), 1925–1927 (2020).
11. T. Seemann. (2016). *ABRicate*. Available: https://github.com/tseemann/abricate
12. Carattoli, A. & Hasman, H. PlasmidFinder and *in silico* pMLST: Identification and typing of plasmid replicons in whole-genome sequencing (WGS). In *Horizontal Gene Transfer: Methods and Protocols* (ed. de la Cruz, F.) 285–294 (Springer US, 2020).
13. Parcell, B. J. et al. Applications and benefits of the British society for antimicrobial chemotherapy resistance surveillance project—legacy and future. *J. Antimicrob. Chemother.* **80**(Supplement_4), iv87–iv95 (2025).
14. Kern, W. V. & Rieg, S. Burden of bacterial bloodstream infection—a brief update on epidemiology and significance of multidrug-resistant pathogens. *Clin. Microbiol. Infect.* **26**(2), 151–157 (2020).
15. Saparamadu, A. A. D. N. S. & Ratnayake, L. Epidemiology of *Escherichia coli* and *Klebsiella pneumoniae* bloodstream infections in a general hospital in Singapore: A retrospective cohort study. *Singapore Med. J.* https://doi.org/10.11622/smedj.2022094 (2023).
16. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses (in eng). *BMC Biol.* **12**, 87 (2014).
17. Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B. & Chiodini, R. J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog.* **8**, 24 (2016) ((**in eng**)).
18. Kim, C., Pongpanich, M. & Porntaveetus, T. Unraveling metagenomics through long-read sequencing: A comprehensive review. *J. Transl. Med.* **22**(1), 111 (2024).
19. Marshall, A. et al. Application of nanopore sequencing for accurate identification of bioaerosol-derived bacterial colonies. *Environ. Sci. Atmos.* **4**(7), 754–766 (2024).
20. Olsen, N. S. & Riber, L. Metagenomics as a transformative tool for antibiotic resistance surveillance: Highlighting the impact of mobile genetic elements with a focus on the complex role of phages. *Antibiotics* https://doi.org/10.3390/antibiotics14030296 (2025).
21. Suzuki, S., Horinouchi, T. & Furusawa, C. Prediction of antibiotic resistance by gene expression profiles. *Nat. Commun.* **5**(1), 5792 (2014).
22. Corona, F. & Martinez, J. L. "Phenotypic resistance to antibiotics," (in eng). *Antibiotics* **2**(2), 237–255 (2013).
23. A. Abramova, A. Karkman, and J. Bengtsson-Palme, "Metagenomic assemblies tend to break around antibiotic resistance genes," *bioRxiv,* p. 2023.12.13.571436, 2023.

24. Zhang, S. et al. Threats across boundaries: The spread of ESBL-positive Enterobacteriaceae bacteria and its challenge to the "one health" concept. *Front. Microbiol.* **16**, 1496716 (2025) ((**in eng**)).
25. Jayathilaka, N., Shehana, S., Nakkawita, D. & Senaratne, T. Prevalence and molecular epidemiology of carbapenem resistance in Asia: A systematic review and meta-analysis. *Systematic Rev.* **14**(1), 123 (2025).
26. S. Fahy, J. A. O'Connor, B. Lucey, and R. D. Sleator. Hospital reservoirs of multidrug resistant acinetobacter species—The elephant in the room! (in English). *Br. J. Biomed. Sci.* Volume 80 (2023).
27. A. Avakh, G. D. Grant, M. J. Cheesman, T. Kalkundri, and S. Hall. The art of war with pseudomonas aeruginosa: Targeting mex efflux pumps directly to strategically enhance antipseudomonal drug efficacy. (in eng). *Antibiotics (Basel)*. **12**(8) (2023).
28. A. L. Lapidus and A. I. Korobeynikov. Metagenomic data assembly – The way of decoding unknown microorganisms (in English). *Front. Microbiol*. Volume 12 (2021).
29. Zhao, W. et al. Oxford nanopore long-read sequencing enables the generation of complete bacterial and plasmid genomes without short-read sequencing. *Front. Microbiol.* **14**, 1179966 (2023) ((**in eng**)).
30. S. Peter *et al.*, "Tracking of antibiotic resistance transfer and rapid plasmid evolution in a hospital setting by nanopore sequencing," (in eng), *mSphere,* vol. 5, no. 4, Aug 19 2020.
31. Pitout, J. D. D. & Chen, L. The significance of epidemic plasmids in the success of multidrug-resistant drug pandemic extraintestinal pathogenic *Escherichia coli*. *Infect. Dis. Ther.* **12**(4), 1029–1041 (2023) ((**in eng**)).
32. Jiang, J. et al. Carbapenem-resistant, virulence plasmid-harboring *Klebsiella pneumoniae*, United States. *Emerg. Infect. Dis.* **31**(4), 761–771 (2025) ((**in eng**)).
33. Du, P., Liu, C., Fan, S., Baker, S. & Guo, J. The role of plasmid and resistance gene acquisition in the emergence of ST23 multi-drug resistant, hypervirulent *Klebsiella pneumoniae*. *Microbiol. Spectr.* **10**(2), e01929-e2021 (2022).
34. Reszetnik, G. et al. Next-generation rapid phenotypic antimicrobial susceptibility testing. *Nat. Commun.* **15**(1), 9719 (2024).
35. E. Kamau and S. Yang. Metagenomic sequencing of positive blood culture fluid for accurate bacterial and fungal species identification: A pilot study. (in eng). *Microorganisms*. 11(5) (2023).
36. Fu, Y. et al. Clinical performance of nanopore targeted sequencing for diagnosing infectious diseases. *Microbiol. Spectr.* **10**(2), e00270-e322 (2022).

Author contributions

TY: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft; TW: Conceptualization, Methodology, Investigation, Supervision, Project administration; PJ: Conceptualization, Formal analysis, Software, Supervision; IT: Conceptualization, Resources, Data curation; MC: Conceptualization, Data curation, Supervision; CT: Conceptualization, Resources, Data curation; PN: Investigation, Validation; MTP: Investigation, Validation; ST: Methodology, Validation; WK: Conceptualization, Writing – review & editing, Data curation; KS: Conceptualization, Software, Writing – review & editing, Supervision.

Funding

This research was supported by the Postdoctoral Fellowship Program, Prince of Songkla University, Thailand. Additional funding was provided by the National Science Research and Innovation Fund (NSRF) and Prince of Songkla University, Thailand (grant number MED6801076S), as well as by the NSRF through the Program Management Unit for Human Resources & Institutional Development, Research and Innovation (grant number B13F660074).

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval and consent to participate

This study was approved by the Siriraj Institutional Review Board, Faculty of Medicine Siriraj Hospital, Mahidol University (COA no. Si 768/2025, SIRB Protocol No.778/2568(IRB02)) under the expedited review process. Informed consent was waived as the study used leftover clinical specimens with minimal risk to participants.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-41247-2.

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

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/.

© The Author(s) 2026