Comparative Evaluation of Commercial DNA Isolation Approaches for Nanopore-only Bacterial Genome Assembly and Plasmid Recovery

Worarat Kruasuwan (worarat.kru@mahidol.edu)
Mahidol University

Pongpun Sawatwong (hps5@cdc.gov)
Division of Global Health Protection, Thailand Ministry of Public Health-U.S. Center of Diseases Control and Prevention

Piroon Jenjaroenpun (piroon.jen@mahidol.edu)
Mahidol University

Natnicha Wankaew (natnicha.wan@mahidol.edu)
Mahidol University

Tantip Arigul (tantip.ari@mahidol.edu)
Mahidol University

Suganya Yongkiettrakul (suganya.yon@biotec.or.th)
National Science and Technology Development Agency (NSTDA)

Kamonwan Lunha (kamonwan.lun@ncr.nstda.or.th)
National Science and Technology Development Agency (NSTDA)

Aunthikarn Sudjai (aunthikarn.sud@mahidol.edu)
Mahidol University

Duangkamon Siludjai (uwz5@cdc.gov)
Division of Global Health Protection, Thailand Ministry of Public Health-U.S. Center of Diseases Control and Prevention

Beth Skaggs (bgs7@cdc.gov)
Division of Global Health Protection, Thailand Ministry of Public Health-U.S. Center of Diseases Control and Prevention

Thidathip Wongsurawat (thidathip.won@mahidol.edu)
Mahidol University

Article

Keywords: Bead-beating, Enzymatic lysis, DNA extraction, Long-read sequencing, Pathogen, GridION

Posted Date: February 2nd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3881497/v1

Additional Declarations: No competing interests reported.

Comparative Evaluation of Commercial DNA Isolation Approaches for Nanopore-only Bacterial Genome Assembly and Plasmid Recovery

Worarat Kruasuwan[1,2], Pongpun Sawatwong[3], Piroon Jenjaroenpun[1,2,4], Natnicha Wankaew[1,2], Tantip Arigul[1,2], Suganya Yongkiettrakul[5], Kamonwan Lunha[5], Aunthikarn Sudjai[1], Duangkamon Siludjai[3], Beth Skaggs[3], and Thidathip Wongsurawat[1,2,4,*]

[1]Division of Medical Bioinformatics, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[2]Siriraj Long-read Lab (Si-LoL), Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[3]Division of Global Health Protection, Thailand Ministry of Public Health-U.S. Center of Diseases Control and Prevention, Nonthaburi, Thailand
[4]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR, USA
[5]National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani, Thailand

***Correspondence:**
Thidathip Wongsurawat, Ph.D.
Email: thidathip.won@mahidol.edu

Keywords: Bead-beating, Enzymatic lysis, DNA extraction, Long-read sequencing, Pathogen, GridION

Abstract

The advent of Oxford Nanopore Technologies has undergone significant improvements in terms of sequencing costs, accuracy, and sequencing read lengths, making it a cost-effective, and readily accessible approach for analyzing microbial genomes. A major challenge for bacterial whole genome sequencing by Nanopore technology is the requirement for a higher quality and quantity of high molecular weight DNA compared to short-read sequencing platforms. In this study, using eight pathogenic bacteria, we evaluated the quality, quantity, and fragmented size distribution of extracted DNA obtained from three different commercial DNA extraction kits, and one automated robotic platform. Our results demonstrated significant variation in DNA yield and purity among the extraction kits. The ZymoBIOMICS DNA Miniprep Kit (ZM) provided a higher purity of DNA compared to other kit-based extractions. All kit-based DNA extractions were successfully performed on all twenty-four samples using a single MinION flow cell, with the Nanobind CBB Big DNA kit (NB) yielding the longest raw reads. The Fire Monkey HMW-DNA Extraction Kit (FM) and the automated Roche MagNaPure 96 platform (RO) outperformed in genome assembly, particularly in gram-negative bacteria. A minimum of 30× to 50× read coverages is recommended for genome assembly and plasmid recovery. Our evaluation indicated that the RO platform gave the best overall performance compared to other kits. The RO platform has the additional advantages of full automation and high throughput. However, consideration of upfront costs associated with instruments and reagents is crucial. In conclusion, our study provides valuable guidance for selecting effective kit-based DNA extraction methods for bacterial whole genomes and plasmids recovery.

Introduction

Over the past two decades, the implementation of microbial whole genome sequencing (WGS) has been considerably advanced in the field of infectious disease epidemiology[1]. WGS has emerged as a critical tool for species identification, sub-species-level typing, outbreak investigation,

and gene function identification. Indeed, this approach has been proven to be a comprehensive and efficient approach for investigating and characterizing antimicrobial resistance (AMR) genes. Additionally, when combined with phenotypic antimicrobial susceptibility testing data, it can effectively identify novel AMR genes and mutations, particularly those mediated by mobile genetic elements like plasmids[2]. Therefore, the Global Antimicrobial Resistance and Use Surveillance System (GLASS), led by the World Health Organization (WHO), advocates for the use of WGS in global antimicrobial resistance surveillance to facilitate the timely development of AMR control strategies[3].

Illumina short-read sequencing, which produces millions of low-error paired-end reads (100–300 bp), has been used for sequencing pathogenic bacteria, and is commonly used for conventional molecular typing, which relies on specific genes as biomarkers[4]. This sequencing platform has limitations in accurately reconstructing complex genome structures, particularly repetitive sequences and mobile genetic elements, which can result in missing or fragmented genes and/or loss of plasmid recovery[5,6]. On the other hand, reconstructing complex genome structure can be addressed through the application of single-molecule sequencing based on Nanopore technology, which allows for the sequencing of repetitive regions such as the rRNA gene operon (range in size between 5 and 7 kb) in the case of bacteria[7,8]. Nanopore is more cost-effective for small batches, has a lower capital cost, and can provide quicker results than Illumina sequencing because Nanopore's flow cells can be washed and reused until all pores are unavailable. Furthermore, improvements in Nanopore sequencing chemistry and base-calling models have improved significantly, with ~6% for R9.4.1 flow cell, resulting in a gradual reduction of error rates over time[9]. However, obtaining sufficient amounts of high-quality input DNA is crucial for successful Nanopore long-read sequencing.

Sequencing of low-quality nucleic acid templates can lead to suboptimal performance or even unsuccessful sequencing runs and high-quality WGS construction. Therefore, it is essential to optimize the DNA

extraction process to obtain high molecular weight (HMW) DNA suitable for long-read sequencing. Even numerous commercially available DNA isolation kits are generally employed in DNA preparation; however, they have not been optimized and applied for ubiquitous bacteria since it differs in the properties of bacterial cell wall types and the efficiency of the kit. Commercial DNA extraction kits mostly emerged during the era of short-read sequencing which typically utilized a combination of mechanical (bead-beating) and chemical (enzymatic lysis) methods to extract DNA, with subsequent purification and elution steps. An example of such a kit is the ZymoBIOMICS DNA Miniprep kit which recently reported to recover bacterial DNA and perform Nanopore long-read whole genome sequencing for characterization of strain, virulence, and antimicrobial resistance genes in *Actinobacillus equuli*[10].

Several methods have been developed to extract high-molecular weight (HMW) DNA from bacteria that are suitable for long-read sequencing. These include a novel magnetic disk in the Nanobind CBB Big DNA kit (PacBio, USA) and a spin-column-based protocol to extract HMW DNA using a high *g*-force of the Fire Monkey High Molecular Weight (HMW) DNA Extraction kit (Revolugen, UK). These methods have been reported to extract HMW DNA from either pathogenic *Escherichia coli* O157:H7, *Klebsiella michiganensis* or *Salmonella* Thyphi and subject to long-read nanopore sequencing to confirm genome rearrangement[11-13]. Despite the development of numerous protocols, DNA extraction remains a bottleneck step in clinical applications due to being labor-intensive and time-consuming features. Moreover, the involvement of multiple steps in these procedures increases the risk of DNA degradation or cross-contamination, particularly when processing a large number of samples simultaneously. Consequently, the utilization of automated robotic platforms for DNA extraction and purification has emerged as a promising solution. Platforms like the MagMAX™ Express Magnetic Particle Processors (Thermo Fisher Scientific Inc., Waltham, USA) and the Roche MagNaPure 96 system (Roche, Switzerland), have the potential to offer

several advantages, including reduced hands-on time, user-friendliness, reproducibility, and the ability to achieve higher throughput levels[14].

In the field of microbial WGS, several commercial kits have been compared for bacterial DNA extraction and their performance has been evaluated by either short- or long-read sequencing. Nonetheless, most studies have been focused on a one specific bacterial species, such as Shiga toxin-producing *E. coli*, *Klebsiella pneumoniae*, or *Salmonella enterica*, and evaluated the performance of the kits in term of in genome assembly[15-17]. There have been relatively fewer studies conducted to assess the effectiveness of automated robotic platforms for DNA extraction. Our goal is to evaluate the performance of commonly available commercial DNA extraction kits, ZymoBIOMICS DNA Miniprep Kit, Nanobind CBB Big DNA Kit, and Fire Monkey High Molecular Weight DNA Extraction Kit, and one automated robotic platform, Roche MagNaPure 96 system for nanopore long-read sequencing of eight pathogenic bacteria. The evaluation focused on their impact on DNA quantity, quality and integrity, as well as subsequent genome assembly and plasmid recovery.

Results

DNA yield and quality of extracted DNA

In our assessment, a consistent starting cell input (~1.2□10^9 CFU mL^{-1}) was used for all eight pathogenic bacteria (Table S1). Each bacterial sample was processed for DNA extraction using the following four different DNA isolation kits (Table 1). The quantity and quality of purified DNA were measured and shown in Fig. 1. Overall, ZymoBIOMICS™ DNA Miniprep Kit (ZM) with bead-beating step demonstrated a significant increase in DNA yield across most of the tested pathogenic bacterial strains, ranging from 20.6–235.3 ng µL^{-1} (Table S2). However, it should be noted that for *Enterococcus faecium* (Efa) and *Streptococcus suis* (Ssu), the Nanobind CBB Big DNA Kit (NB) yielded higher amounts of DNA, with values of 63.5 ± 10.5 ng µL^{-1} and 116.8 ± 17.3 ng µL^{-1}, respectively, compared to ZM, Fire Monkey HMW-DNA Extraction Kit (FM) and Roche MagNaPure 96 system (RO). Notably, there is no significant difference in

the DNA concentration from *Streptococcus agalactiae* (Sag) among the DNA isolation kits (Fig. 1, Table S2).

The absorption spectra were subsequently investigated to assess the purity of DNA samples and any contaminants. Acceptable values for pure DNA typically are within the range of 1.8–2.2 for A260/A280 and ≥ 2.0 for A260/A230 ratio. The results showed that all extraction kits exhibited lower values within the desired range (1.2–1.7 of A260/280) for all gram-positive pathogenic bacteria. The extraction kit, FM resulted in A260/A280 ratio ≤ 1.8 for gram-negative bacteria, suggesting a potential presence of contaminants in the DNA samples. Nonetheless, three extraction kits, ZM, NB, and RO, demonstrated acceptable A260/A280 ratios for all gram-negative pathogenic bacteria, except FM which resulted in ≤ 1.8. Among the tested extraction kits, ZM only achieved an acceptable ratio of A260/230 ratio (≥ 2.0) for both gram-positive and negative pathogenic bacteria when compared to three extraction kits (Fig. 1, Table S3).

Following DNA extraction, gel electrophoresis was performed using TapeStation to visually examine the size distribution of the obtained DNA fragments. In general, all of the DNA extraction kits provided a single DNA band that corresponded to the reference 48.5 kb gDNA ladder. Genomic DNA extracted by the FM resulted in the same size DNA fragments as the RO extraction kit, while the ZM generated the smallest size distribution and a faint smear of extracted DNA. Remarkably, the NB extraction kit resulted in greater DNA size than the reference gDNA ladder for both gram-positive and gram-negative pathogenic bacteria, except for Sag and *Stapylococcus aureus* (Sau) strains which the ZM gave the single DNA band (Fig. S1).

Sequencing statistics and assembled genomes evaluation

To evaluate the influence of commercial DNA isolation kits on bacterial genome assembly, a total of twenty-four samples, comprising three independent replicates of eight pathogenic bacteria, were pooled together. The pooled samples were then sequenced on the same flow cell (R9.4.1) using GridION sequencer. The sequencing process resulted in the

generation of twelve assemblies for each strain, based on the utilization of four DNA extraction kits with three replicates each. The sequencing runs generated 10.89, 7.71, 13.96, and 14.01 GB of sequenced DNA, which was extracted by ZM, NB, FM and RO, respectively. Among the four strains of gram-positive pathogenic bacteria, the NB generate read length *N50* in excess of 6,000 bp for both Efa (*N50* = 8,036 bp) and Ssu (*N50* = 6,304 bp) strains while FM, ZM and RO gave the maximum read length of 5,978 bp, 3,528 bp, and 4,146 bp for Efa, respectively. Similarly, both NB and FM extraction kits yielded the longest raw read length in both Aba and Pae, except for Kpn and Sgd of which either FM- or RO-extracted DNA respectively showed the highest *N50* at 8,751 and 9,405 bp (Table S4).

Flye assembled genome using all sequencing reads was measured based on total length, the number of contigs and contig *N50*. In this study, the number of contigs varied depending on the extraction kits used and the bacterial strains. Remarkably, the RO failed to produce sufficient sequencing reads for successful genome assembly in the Sag strain, while the NB had the same outcome for the Sau strain. However, both the RO and FM exhibited a noteworthy reduction in the number of contigs for most pathogenic bacteria (6/8 strains, 75%), followed by ZM (5/8 strains, 62.5%) when comparing to reference genomes (Table 2, Fig. S2). The NB, on the other hand, exhibited lower success (3/8 strains, 37.5%) in terms of genome assembly performance as indicated in Tables 2 and Tables S5-S6. Despite ZM and NB demonstrating the lower performance in genome assembly, NB exhibited a significant advantage in specific instances, particularly with the Efa and Kpn strains of which the DNA extracted by NB achieved a considerably higher contig *N50* value of 2,255,486 bp and 3,672,441 bp, surpassing ZM-extracted DNA assembly with contig *N50* of 1,731,151 bp and 285,365 bp, respectively (Table 2).

In terms of plasmid recovery, the number of plasmids observed in gram-positive bacteria varied depending on the specific extraction kit used. Notably, Efa yielded the same number of plasmids as the reference genome (5 contigs) when using either NB or FM kits. While, higher plasmid contig numbers were found in RO (6 contigs) and ZM (11 contigs), due to

fragmentation or replication of plasmid contigs by the Flye assembler (Table 2, Table S6). On the other hand, the FM and RO extraction kits exhibited inferior recovery of plasmid in most gram-negative bacteria, particularly noticeable in the case of Aba and Kpn, where only 2 and 3 plasmids were obtained for the NB kit, respectively. However, the plasmid contig numbers for Knp of Sgd were higher when extracted using ZM and NB kits, respectively (Table 2, Table S6). All assembled genomes were compared against the Genome Taxonomy Database (GTDB-Tk) taxonomic classification based on genomes comparison, using a 95% average nucleotide identity (ANI) values cutoff to group genomes belonging to the same species. The results of the genome-based taxonomic assignment revealed that all genomes displayed ANI values >97%, indicating a high similarity to the described strains (Table S7).

Long-read coverage on bacterial genome assembly statistics

Flye genome assembly of subsampled read coverage, including 20×, 30×, 50×, 80×, and 100×, was evaluated by analyzing the raw read *N50*, number of contig, completeness, and contig contiguity through the observation of the *N50* values, representing the minimum contig length needed to cover 50% of the genome. Overall, sequencing reads, particularly in Aba and Pae strains, generated by NB- (12 and 15 kb) and FM-extracted kits (10 and 11 kb) showed greater read lengths compared to those from ZM (4.3 and 5.9 kb) and RO (6.4 and 7.5 kb) kits (Fig. 3, Table S8). As the read coverage approached 30× to 50×, there was a reduction in the number of contigs for all assembled genomes. Furthermore, increasing the read coverage resulted in improved genome completeness, with values exceeding 97%. The reduction in the number of contigs for all assembled genomes depends on extraction kits. For instance, the chromosome numbers of Ssu, Aba, Kpn, and Pae were reduced to a single chromosome even with 20× read coverage when using extraction kits from FM and RO (Figs. 2 and 3).

For gram-positive bacteria, the assembly contiguity of assemblies improved substantially when the read coverage exceeded 30× coverages,

with the exception for the ZM (31 contigs), which had a lower number of contig for DNA extracted by NB (1 contig), FM (1 contig) and RO (1 contig) for Ssu, for example. It is noteworthy that NB improved the contig *N50* value from 1.43 Mb (20×) to 2.07 Mb (30×), while FM (2.07 Mb) and RO (2.07 Mb) demonstrated in achieving 99–100% genome coverage at the 20× read coverages. The RO yielded suboptimal reads for genome assembly in Sag strain, with ~25% genome completeness, while the NB extraction kit resulted in 0% genome completeness in the Sau strain (Figs. 2b, 2d and Table S8). In contrast, increasing the read coverage led to an improvement in the proportion of genome completeness in other gram-positive genomes across all extraction kits, ranging from approximately 98–100%. However, regardless of the increased coverage to 80×, ZM kit yielded unusually low contig *N50* value of the Efa, Sag, and Ssu genomes compared to Sau strain (Fig. 2, Table S8).

For gram-negative bacteria, ZM and NB kits resulted in a high number of assembled contigs for Kpn and Sgd, even with an increase in read coverage to 80× (43 contigs and 3 contigs, respectively), compared to the other extraction kits. Nonetheless, both FM and RO demonstrated incremental improvement in genome completeness (99.4–100%), contig *N50* values, and a reduced in the numbers of contigs for all four gram-negative pathogenic bacteria when compared to ZM and NB at the coverage of 30×. Unexpectedly, the RO extraction provided a high-quality genome of all gram-negative bacteria, even with a read coverage as low as 20× (Fig. 3, Table S8).

Long-read coverage on recovered plasmid number

Overall, a minimum read coverage of ≥50× was found to be sufficient for accurate plasmid recovery from most subsampled assemblies obtained using the four different extraction kits. However, the NB kit failed to recover the plasmid of Sau. When comparing the assembled genomes obtained from various extraction kits, similar numbers of recovered plasmids were observed. However, there were exceptions for Efa and Kpn genomes extracted by either the ZM extraction kit which resulted in a high

number of the plasmids, except for the Aba and Sgd strains. Interestingly, an increase in read coverage resulted in a decreased number of reconstructed plasmids in the genome, such as particularly in the case of Kpn, where the number of contigs decreased from 11 contigs at 30× coverage to 5 contigs at 80× coverage when extracted using the ZM (Fig. 4, Table S9).

Discussion

This study aims to evaluate the efficacy of different DNA extraction kits and an automated robotic platform in term of their impact on the performance of long-read nanopore sequencing and influence on the subsequent processes, genome assembly and plasmid recovery. DNA quality is a significant factor contributing to inadequate genome assembly. To enhance the quality of DNA extraction, several commercially available DNA extraction kits have been employed, aiming to identify the most suitable kit that is applicable to all bacterial species. Our results demonstrated that most of the tested DNA extraction kits provided a sufficient amount of high molecular weight (HMW) DNA (50 ng per sample) for DNA library construction of SQK-RBK110.96 kit. However, in this study, none of the tested extraction kits provided enough DNA for Sag while all the kits, except for the ZM, yielded the lowest DNA amount. The efficiency of ZM kit may be attributed to the manufacturer's recommended bead beating protocol, which differs from the other kits. The utilization of enzymatic lysis in combination with the bead-beating method notably enhanced in DNA yield. This approach facilitated the lysis of gram-positive bacterial cells, particularly in Sag and other gram-negative bacteria. Our findings are correspondent with prior studies, emphasizing the importance of bead beating in combination with the enzymatic lysis for gram-positive bacteria, resulting in higher DNA yields and improved performance of long-read nanopore sequencing. This improvement applies not only to the single strain investigated in this study but also to microbial communities as a whole, for instance human gut microbial community[18,19]. However,

the sensitivity to bead beating varies among species, as revealed by the recent report and present study[20].

Regarding DNA quality, the FM kit produced DNA samples with very low A260/280 and A260/230 ratios, indicating the potential presence of protein contamination, organic solvents, or residual reagents from the purification process. Conversely, the ZM kit which employed beat-beating, resulted in acceptable A260/A230 ratios across all tested pathogens compared to other extraction kits. However, this method led to increased fragmentation of DNA (Table S2, Fig. S1). Despite the potential of the ZM extraction kit to yield sufficient amount of DNA from most of the tested bacteria, our results suggest the importance of optimizing the specific duration of bead-beating process. This optimization is crucial to strike the right balance between maximizing DNA yield and minimizing DNA fragmentation, ensuring optimal conditions for nanopore long-read sequencing application.

Nanopore long-read sequencing confirmed the success of combining twenty-four samples in a single run and utilizing them for genome assembly and species identification (Tables S5–S7). We found that the NB and FM extraction kits produced the longest filtered read *N50* values across most of the pathogenic bacteria, while ZM exhibited the shortest filtered read *N50* (Table S4). Nonetheless, the total number of reads produced by NB kit was notably lower than that of FM kit, particularly in gram-positive pathogenic bacteria, Sag and Sau. This lower read count led to lower success rates for genome assembly. On the other hand, the genome assembly statistic such as total length of the genome, particularly in gram-negative bacteria, did not show any difference among the extraction kits used, except for the Kpn which were extracted by ZM kit (Table S5). Our results suggest that either HMW DNA extraction kits (NB and FM) or an automated RO platform could be effectively employed for long-read sequencing, enabling both nearly complete genome assembly and species identification in most pathogenic bacteria.

For bacterial genome assembly and plasmid recovery, considerable variability was observed in sequencing read coverages for complete

genome assembly when relying solely on nanopore long-read sequences. This variability was found to be dependent on the complexity of each genome. In this study, we observed minimal improvement in contig *N50* beyond a depth of 30× for both gram-positive and gram-negative bacteria across DNA extraction kits indicating that a sequencing depth of 30× was sufficient to achieve satisfactory genome assembly. Our result correlates with previous reports suggest that the depth of 30× is sufficient for *de novo* assembly of the complete genome and reliably determine single-nucleotide variations in the genome of *Escherichia coli*[21]. However, it is noted that other studies have suggested that, for larger bacterial genomes like *Pseudonocadia*, a coverage depth of 40× to 50× may be required for sufficient coverage[22]. Furthermore, this was prominently demonstrated by our long-read coverage on plasmid recovery, where most of the extraction kits yielded the numbers of plasmid closely to reference assembled genomes. Notably, FM and RO extraction kits proved to be particularly effective in generating accurate and contiguous microbial genome assemblies, as evidenced by their performance in plasmid recovery at 50× coverage for gram-positive and gram-negative bacteria (Table S9). However, the number of plasmids varied among tested bacteria particularly when using ZM extraction kit for the assembled genomes of Efa and Kpn. This difference is consistent with the shorter raw reads (*N50* = 3,528 bp for Efa and *N50* = 3,787 bp for Kpn, Table S4), resulting to generate lower-quality genomes and plasmids compared to other kit-based extraction. Recently report demonstrated that using long-read-only genome assemblers such as Flye, Miniasm, Canu, and Raven encounters difficulties when dealing with small plasmids, particularly those smaller than 10 Kb. Even though the reason remains uncertain, the small plasmids were absent in approximately one-third of all repeated assemblies, and they had noticeably greater average read depths, which suggested that this could be related to differences in sequencing depths[23]. Thus, increasing sequencing read depth could possibly result in a lower number of reconstructed plasmids, especially in the case of Efa when extracted by ZM and RO as shown in this work (Fig. 4, Table S9).

Regarding cost and time effectiveness, the RO method demonstrated superior performance, compared to the other evaluated extraction kits in this work (Table 1). The utilization of robotic extraction platforms can further enhance efficiency in both analyses and reduce potential analytical errors. This is beneficial especially when handling a large number of samples in batches, as previously demonstrated in this work and also reported in other studies involving dietary samples[24]. Nonetheless, additional costs of equipment and infrastructure can lead to doubling of the overall setup cost when an automated robotic DNA extraction platform is implemented, compared to other kit-based extractions. The HMW DNA extraction kits, the FM and NB extraction kits, exhibit a considerably higher cost, amounting to approximately six times the cost of the ZM extraction kit (approximately $61 USD per sample compared to $9 USD per sample). It is worth noting that the ZM kit does not require a pre-lysis step, leading to a simple, more effective, and time-saving than the other kit-based extraction methods.

Conclusions

Our findings revealed that the ZM kit, which combines enzymatic lysis and bead-beating steps, outperformed other kit-based extractions methods in term of yielding high-purity DNA. The NB kit generated the longest raw sequences and showed comparable performance to the FM kit and the automated RO platform in terms of genome assembly, particularly in gram-negative bacteria. Additionally, because multiplex genomes (24 genomes) can be sequenced on a single MinION flowcell, then, we recommend a read coverage of 30× to 50× to sufficiently minimize the number of contigs for all assembled genomes and increased the genome completeness, including plasmid recovery. Although both the NB and FM kits required more hands-on time, they offer the benefit of generating longer DNA molecular weight sizes, which can be advantageous for obtaining longer sequencing read lengths and improving the quality of genome assembly. Conversely, the RO kit demonstrated superiority in terms of reduced processing time and labor compared to other DNA

extraction kits. However, it is important to consider the additional upfront cost for instruments and reagents, as well as the cost per run to ensure technical reproducibility. In summary, our findings provide valuable insights for laboratories seeking to make informed decisions regarding the selection of DNA extraction kits for genome assembly and plasmid recovery.

Materials and methods

Pathogenic bacteria samples

Eight pathogenic bacteria, including four strains of gram-positive bacteria; *Enterococcus faecium* SF01961 (Efa), *Streptococcus agalactiae* SF04137 (Sag), *Streptococcus suis* NF06446 (Ssu), *Staphylococcus aureus* SFP009 (Sau), and four strains of gram-negative bacteria; *Acinetobacter baumannii* SPP007 (Aba), *Klebsiella pneumoniae* SF05210 (Kpn), *Pseudomonas aeruginosa* SF01204 (Pae), *Salmonella* spp. Group D SA8854 (Sgd), obtained from the Division of Global Health Protection, Thailand Ministry of Public Health-U.S. Center of Diseases Control and Prevention (Nonthaburi, Thailand) were used for bacterial genomic DNA extraction in this work (Table S1). All bacterial culture was maintained on Colombia 5% Sheep Blood Agar (Scharlau, Spain) at 30°C for 18–24 h before further genome extraction step.

Initial bacterial cell density preparation

The initial of bacterial cell suspensions was adjusted to a cell density of McFarland = 4 (~1.2×10^9 CFU mL^{-1}) by resuspending the bacterial cell with 0.1 M phosphate buffer solution (PBS, pH 7.2; Gibco™, ThermoFisher Scientific, MA, USA). The cell pellet was collected by centrifuging of 1 mL cell suspension at 16,000 ×*g* for 1 min. The experiment was performed in three independent replicates per treatment.

Evaluation of bacterial gDNA isolation procedures

In this work, we initiated our investigation by an evaluation of three commercial DNA extraction kits: 1) ZM: ZymoBIOMICS™ DNA Miniprep

Kit (D4300, Zymo Research, USA), 2) NB: Nanobind CBB Big DNA Kit (Circulomics, USA), 3) FM: Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (Revolugen, UK), and (4) RO: one robot-based extraction system (MagNaPure 96 system; Roche, Switzerland). Manufacturers' instructions were followed for all methods except where noted (Supplementary Methods). In brief, DNA extraction using ZM was performed on 250 µL of cell pellet resuspended with 0.1 M PBS according to the manufacturer's protocol, including a modified bead beating step of 3 mins. DNA extraction using NB was performed according to the manufacturer's protocol. However, lysostaphin was not substituted for lysozyme as recommended for the pre-digestion step of *Staphylococcus aureus*, and FM, bacterial DNA was isolated as the manufacturer described with the following modification by using the eluted DNA from the Fraction A for further analysis. For RO, a MagNA Pure 96 DNA and Viral NA Small Volume Kit were applied for this experiment. Most extracted DNA obtained from three commercial kits was finally eluted using 100 µL of either nuclease-free water or elution buffer as recommended except Roche system which was eluted at 50 µL. Then, all extracted DNA were finally purified using 0.8× AMPure XP beads (Beckman Coulter, USA) and eluted at 25 µL of nuclease-free water.

Determination of DNA yield, purity metrics and fragment size distribution

The DNA yield was quantified on a Qubit™ 4.0 Fluorometer (Invitrogen, USA) using the dsDNA Broad Range Assay kit according to the manufacturer's protocols. The purity of the extracted DNA with the A260/280 and A260/230 absorbance ratios was obtained using a NanoDrop spectrophotometer (ThermoFisher Scientific, USA). The DNA fragment size distribution was analyzed by 2200 TapeStation with Genomic DNA ScreenTape Assay according to the manufacturers' instructions (Agilent Technologies, USA).

Library preparation and sequencing

For long-read sequencing, the library was prepared from 50 ng input DNA using the SQK-RBK110.96 kit (Oxford Nanopore Technologies, UK). The library was loaded into the R9.4.1 flow cell (FLO-MIN106; Oxford Nanopore Technologies, UK) and sequenced using GridION with the default setting. Guppy v6.0.1 with the SUP (super accuracy) mode was used for base calling and quality control studies[25]. For short-read sequencing, the DNA library was constructed using MiSeq Reagent Kit v3 (Illumina, USA). Illumina libraries were sequenced in pair-end mode using the Illumina MiSeq platform (Illumina, USA).

Raw read processing and genome assembly

The quality and adapter trimming of raw sequenced reads obtained from ONT is possessed by Porechop v0.2.4 (https://github.com/rrwick/Porechop) and Filtlong v0.2.1 for filtering, keeping only reads over 1,000 base pairs and with a quality score (Q) above 9. NanoPlot v1.38.0 was used to evaluate the resulting reads[25]. Illumina reads were quality checked using FastQC v0.11.9[26], adapters were removed, and low-quality reads (Q ≤ 30) were filtered out using fastp v0.23.2[27] with default parameters. To construct the reference genome of eight isolates, hybrid assembly of both Nanopore long-read and Illumina short-read were assembled using Unicycler v0.4.8[28]. Consecutively, the assembled genome was then checked for completeness and contamination using CheckM v1.2.1 (lineage_wf -r)[29] and MOB-suite v3.1.5 (--run_typer) was used for plasmid typing[30]. The genome features were evaluated by QUAST v5.0.2[31] and plasmid contigs were verified by searching against PSLDB database[32].

For only-long read genome assembly, both all filtered reads and read subsets (20×, 30×, 50×, 80×, and 100× coverages), generated by seqtk v1.3 (https://github.com/lh3/seqtk), were assembled using Flye v2.9.2-b1786 (--meta)[33], and subsequently polished in one round of Medaka v1.8.0 (-m r941_min_sup_g507) (https://github.com/nanoporetech/medaka) with default settings in order to facilitate highly accurate assemblies. Assembly quality was assessed following aforementioned

described. Next, the bacterial chromosome was then identified by aligning against all identified marker genes in the GTDB-Tk database (R207_v2). The average nucleotide identity (ANI) and alignment fraction (AF) are calculated using GTDB-Tk v2.1.1[34].

The genome assemblies obtained solely from only long-read assembly were aligned to either the genome or plasmid contigs of the reference genome using Minimap2 v2.2.21[35] with provided parameters (--secondary=no --cs -cx asm5) to validate the genome reconstruction. Either chromosome or plasmid sequences were considered present if the total draft assembly alignment length exceeded 90% of the reference contig length. In the case that more than one draft contig aligned to a reference contig, the total length of all aligned draft contigs was considered. Standard assembly quality metrics (genome size, total number of contigs, contig length, and *N50*) and number of either chromosome or plasmid recovery were used for each extraction kit performance assessment.

Estimation of time and cost

The comprehensive time and cost of four selected commercial DNA extraction kits were estimated in terms of time and material expenses. The cost of one extraction for each method was calculated based on the list price for necessary supplies and DNA extraction kits (as of January 2023). Start-up costs for the Roche MagNaPure 96 system, as well as material supplies, were excluded. Estimated processing times were calculated based on processing 24 samples and included time taken to pre-treat samples with enzymatic digestion. Then, Comprehensive cost and time were calculated as: (estimated cost per extraction of any one method / maximum estimated cost among four methods) × (estimated time per extraction of any one method / maximum estimated time among four methods) as previously described by Wang, et al. [36].

Statistical analysis and data visualization

Data were subjected to statistical analysis using either one-way ANOVA using post hoc correction by Duncan’s multiple range test (IBM

Statistic SPSS, version 23). Data were presented as mean ± S.D. calculated from three different replicates, with a different letter indicating statistical significance at $p<0.05$. The visual representations of the data, encompassing all graphical depictions, were generated utilizing the ggplot2 plotting library within the R programming language's package system. The reference complete genomes resent in this study were visualized by Bangdage[37].

References

1 Price, V. *et al.* A systematic review of economic evaluations of whole-genome sequencing for the surveillance of bacterial pathogens. *Microb Genom.* **9**, 000900, doi:10.1099/mgen.0.000947 (2023).

2 Weinmaier, T. *et al.* Validation and application of long-read whole-genome sequencing for antimicrobial resistance gene detection and antimicrobial susceptibility testing. *Antimicrob Agents Chemother.* **67**, e0107222, doi:10.1128/aac.01072-22 (2023).

3 Glass. GLASS Whole-genome sequencing for surveillance of antimicrobial resistance., (World Health Organization, New York, 2020).

4 Bogaerts, B. *et al.* Evaluation of WGS performance for bacterial pathogen characterization with the Illumina technology optimized for time-critical situations. *Microb. Genom.* **7**, 000699, doi:https://doi.org/10.1099/mgen.0.000699 (2021).

5 De Maio, N. *et al.* Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. *Microbial Genomics* **5**, 000294, doi:https://doi.org/10.1099/mgen.0.000294 (2019).

6 George, S. *et al.* Resolving plasmid structures in Enterobacteriaceae using the MinION nanopore sequencer: assessment of MinION and MinION/Illumina hybrid data assembly approaches. *Microb. Genom.* **3**, 000118, doi:https://doi.org/10.1099/mgen.0.000118 (2017).

7 Treangen, T. J., Abraham, A.-L., Touchon, M. & Rocha, E. P. C. Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS*

Microbiol. Rev. **33**, 539–571, doi:10.1111/j.1574-6976.2009.00169.x (2009).

8 Taylor, T. L. *et al.* Rapid, multiplexed, whole genome and plasmid sequencing of foodborne pathogens using long-read nanopore technology. *Sci. Rep.* **9**, 16350, doi:10.1038/s41598-019-52424-x (2019).

9 Delahaye, C. & Nicolas, J. Sequencing DNA with nanopores: Troubles and biases. *PLoS ONE* **16**, e0257521, doi:10.1371/journal.pone.0257521 (2021).

10 Vereecke, N., Vandekerckhove, A., Theuns, S., Haesebrouck, F. & Boyen, F. Whole genome sequencing to study antimicrobial resistance and RTX virulence genes in equine *Actinobacillus* isolates. *Vet. Res.* **54**, 33, doi:10.1186/s13567-023-01160-2 (2023).

11 Kang, M., Chmara, J., Duceppe, M.-O., Phipps-Todd, B. & Huang, H. Complete genome sequence of a Canadian K*lebsiella michiganensis* strain, obtained using Oxford Nanopore Technologies sequencing. *Microbiol Resour Announc.* **9**, e00960–00920, doi:doi:10.1128/mra.00960-20 (2020).

12 Greig, D. R., Jenkins, C., Gharbia, S. E. & Dallman, T. J. Analysis of a small outbreak of Shiga toxin-producing *Escherichia coli* O157:H7 using long-read sequencing. *Microb Genom.* **7**, 000545, doi:https://doi.org/10.1099/mgen.0.000545 (2021).

13 Waters, E. V., Tucker, L. A., Ahmed, J. K., Wain, J. & Langridge, G. C. Impact of *Salmonella* genome rearrangement on gene expression. *Evol. Lett.* **6**, 426–437, doi:10.1002/evl3.305 (2022).

14 Boyd, J. Robotic laboratory automation. *Science* **295**, 517–518, doi:doi:10.1126/science.295.5554.517 (2002).

15 Becker, L., Steglich, M., Fuchs, S., Werner, G. & Nübel, U. Comparison of six commercial kits to extract bacterial chromosome and plasmid DNA for MiSeq sequencing. *Sci. Rep.* **6**, 28063, doi:10.1038/srep28063 (2016).

16 Jaudou, S., Tran, M.-L., Vorimore, F., Fach, P. & Delannoy, S. Evaluation of high molecular weight DNA extraction methods for

long-read sequencing of Shiga toxin-producing *Escherichia coli*. *PLoS ONE* **17**, e0270751, doi:10.1371/journal.pone.0270751 (2022).

17 Eagle, S. H. C., Robertson, J., Bastedo, D. P., Liu, K. & Nash, J. H. E. Evaluation of five commercial DNA extraction kits using *Salmonella* as a model for implementation of rapid Nanopore sequencing in routine diagnostic laboratories. *Access Microbiol.* **5**, 000468, doi:https://doi.org/10.1099/acmi.0.000468.v3 (2023).

18 de Boer, R. *et al.* Improved detection of microbial DNA after bead-beating before DNA isolation. *J. Microbiol. Methods* **80**, 209–211, doi:https://doi.org/10.1016/j.mimet.2009.11.009 (2010).

19 Lim, M. Y., Song, E.-J., Kim, S. H., Lee, J. & Nam, Y.-D. Comparison of DNA extraction methods for human gut microbial community profiling. *Syst. Appl. Microbiol.* **41**, 151–157, doi:https://doi.org/10.1016/j.syapm.2017.11.008 (2018).

20 Zhang, B. *et al.* Impact of Bead-Beating Intensity on the Genus- and Species-Level Characterization of the Gut Microbiome Using Amplicon and Complete 16S rRNA Gene Sequencing. *Front. Cell Infect Microbiol.* **11**, 678522, doi:10.3389/fcimb.2021.678522 (2021).

21 Khrenova, M. G. *et al.* Nanopore sequencing for *de novo* bacterial genome assembly and search for single-nucleotide polymorphism. *Int J Mol Sci* **23**, 8569, doi:10.3390/ijms23158569 (2022).

22 Goldstein, S. L. & Klassen, J. L. *Pseudonocardia* symbionts of fungus-growing nnts and the evolution of defensive secondary metabolism. *Front. Microbiol.* **11**, 621041, doi:10.3389/fmicb.2020.621041 (2020).

23 Johnson, J., Soehnlen, M. & Blankenship, H. M. Long read genome assemblers struggle with small plasmids. *Microb. Genom.* **9**, 001024, doi:https://doi.org/10.1099/mgen.0.001024 (2023).

24 Wallinger, C. *et al.* Evaluation of an automated protocol for efficient and reliable DNA extraction of dietary samples. *Ecol. Evol.* **7**, 6382–6389, doi:10.1002/ece3.3197 (2017).

25 De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669, doi:10.1093/bioinformatics/bty149 (2018).

26 Andrews, S. (2010).

27 Chen, S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**, e107, doi:https://doi.org/10.1002/imt2.107 (2023).

28 Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595, doi:10.1371/journal.pcbi.1005595 (2017).

29 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055, doi:10.1101/gr.186072.114 (2015).

30 Robertson, J. & Nash, J. H. E. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb Genom.* **4**, 10.1099/mgen.1090.000206, doi:10.1099/mgen.0.000206 (2018).

31 Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075, doi:10.1093/bioinformatics/btt086 (2013).

32 Galata, V., Fehlmann, T., Backes, C. & Keller, A. PLSDB: a resource of complete bacterial plasmids. *Nucleic Acids Res.* **47**, D195–D202, doi:10.1093/nar/gky1050 (2018).

33 Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol.* **37**, 540–546, doi:10.1038/s41587-019-0072-8 (2019).

34 Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, 5315–5316, doi:10.1093/bioinformatics/btac672 (2022).

35 Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100, doi:10.1093/bioinformatics/bty191 (2018).
36 Wang, Y.-S., Dai, T.-M., Tian, H., Wan, F.-H. & Zhang, G.-F. Comparative analysis of eight DNA extraction methods for molecular research in mealybugs. *PLoS ONE* **14**, e0226818, doi:10.1371/journal.pone.0226818 (2020).
37 Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350–3352, doi:10.1093/bioinformatics/btv383 (2015).

Acknowledgement

The authors would like to thank the Division of Global Health Protection, Thailand Ministry of Public Health-U.S. Center of Diseases Control and Prevention (Nonthaburi, Thailand) for providing all pathogenic bacteria used in this study, automated DNA extraction machine, and supporting the Nanopore sequencing facility. For computing facility, we thank Mahidol University and the Office of the Ministry of Higher Education, Science, Research, and Innovation under the Reinventing University project: the Center of Excellence in AI-Based Medical Diagnosis (AI-MD) sub-project.

Funding

This work was supported by Health Systems Research Institute of Thailand under the Genomics Thailand Initiative (HSRI 65-118). TW and PJ were partially supported by the National Research Council of Thailand (NRCT) Project ID N42A660897. TA and NW have received funding support from the NSRF via the Program Management Unit for Human Resources & Institutional Development, Research Innovation (Grant No. B13F660073).

Author contributions

WK, PS, SY, BS, PJ, and TW conceptualized, conceived and designed the study. WK, PS, DS carried out the experimental work and interpretation of data. TA, NW, PJ performed bioinformatics. WK wrote the original draft of the manuscript. All participated in the review and editing of the manuscript. All authors contributed to the article and approved the submitted version.

Data Availability Statement

The original contributions presented in the study are included in the article or supplementary material, further inquiries can be directed to the corresponding authors. All Nanopore sequencing data used in this study have been uploaded to the sequence read archive (SRA) numbers under the BioProject number PRJNA909850.

Competing interest statement

The authors declare that they have no competing interests. Use of trade names is for research only and does not imply endorsement by all authors and the Division of Medical Bioinformatics, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Figures Legends

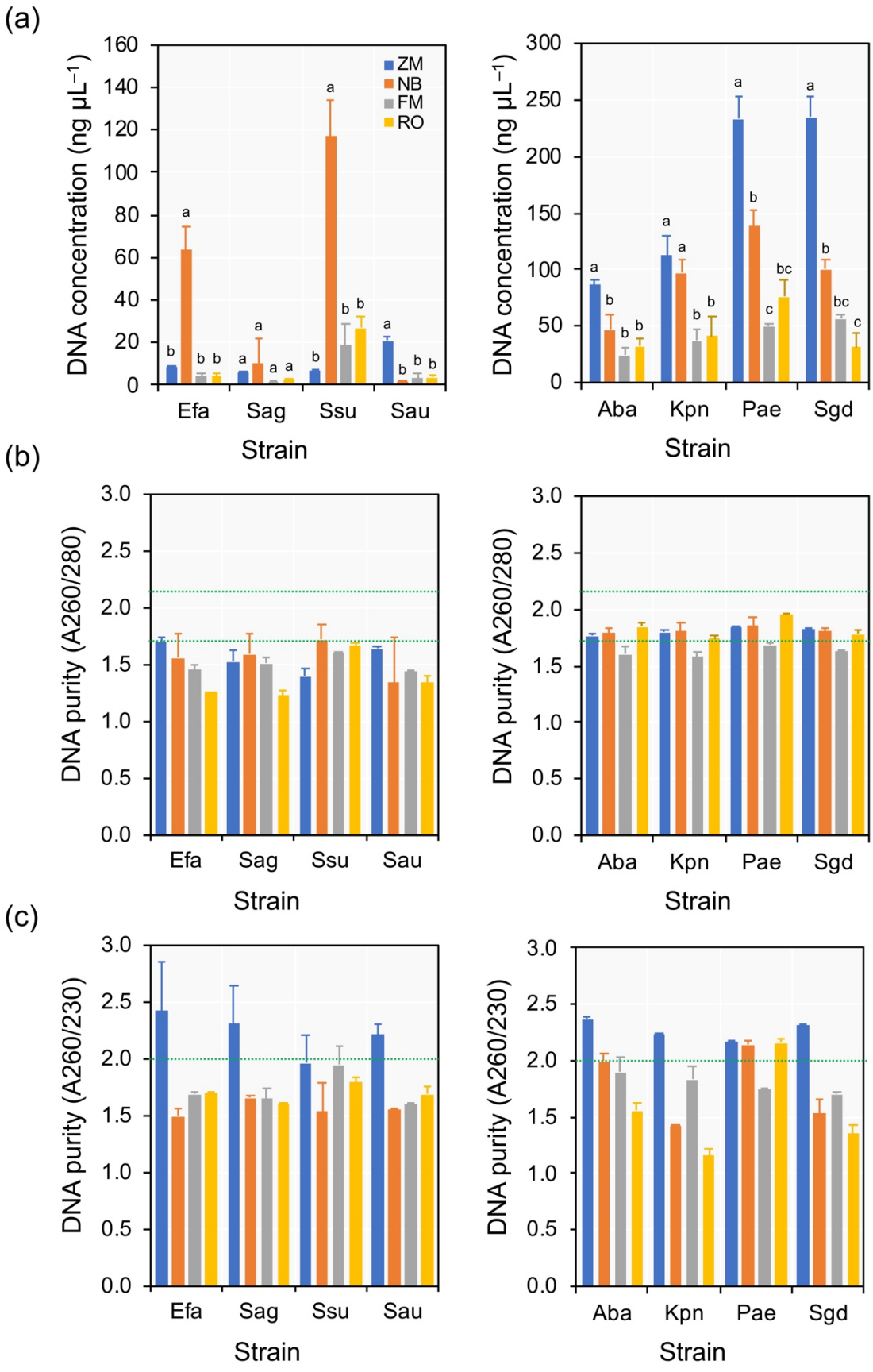
Fig. 1 Influence of commercial DNA extraction kits; ZymoBIOMICS™ DNA Miniprep Kit (ZM, Blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and MagNaPure 96 system (RO, yellow) on DNA concentration (ng μL^{-1}) (a) and DNA purity in ratio of A260/280 (b) and A260/230 (c) of eight pathogenic bacteria. Hatched green lines in (b) and (c) indicate recommended intervals.

Fig. 2 Influence of genome coverage on raw read length, number of contigs, percentage of completeness and contig *N50* of *Enterococcus*

faecium (Efa, a), *Streptococcus agalactiae* (Sag, b), *Streptococcus suis* (Ssu, c), and *Staphylococcus aureus* (Sau, d) extracted by ZymoBIOMICS™ DNA Miniprep Kit (ZM, blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and MagNaPure 96 system (RO, yellow).

Fig. 3 Influence of genome coverage on raw read length, number of contigs, contig *N50*, and percentage of completeness of *Acinetobacter baumannii* (Aba, a), *Klebsiella pneumoniae* (Kpn, b), *Pseudomonas aeruginosa* (Psu, c), and *Salmonella* sp. group D (Sgd, d) extracted by ZymoBIOMICS™ DNA Miniprep Kit (ZM, blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and MagNaPure 96 system (RO, yellow).

Fig. 4 Influence of genome coverages on number of plasmid of *Enterococcus faecium* (Efa, a), *Staphylococcus aureus* (Sau, b), *Acinetobacter baumannii* (Aba, c), *Klebsiella pneumoniae* (Kpn, d), and *Salmonella* sp. group D (Sgd, e) extracted by ZymoBIOMICS™ DNA Miniprep Kit (ZM, blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and MagNaPure 96 system (RO, yellow). Hatched black lines indicate numbers of plasmid in reference genomes.



774

775 **Fig. 1**

776

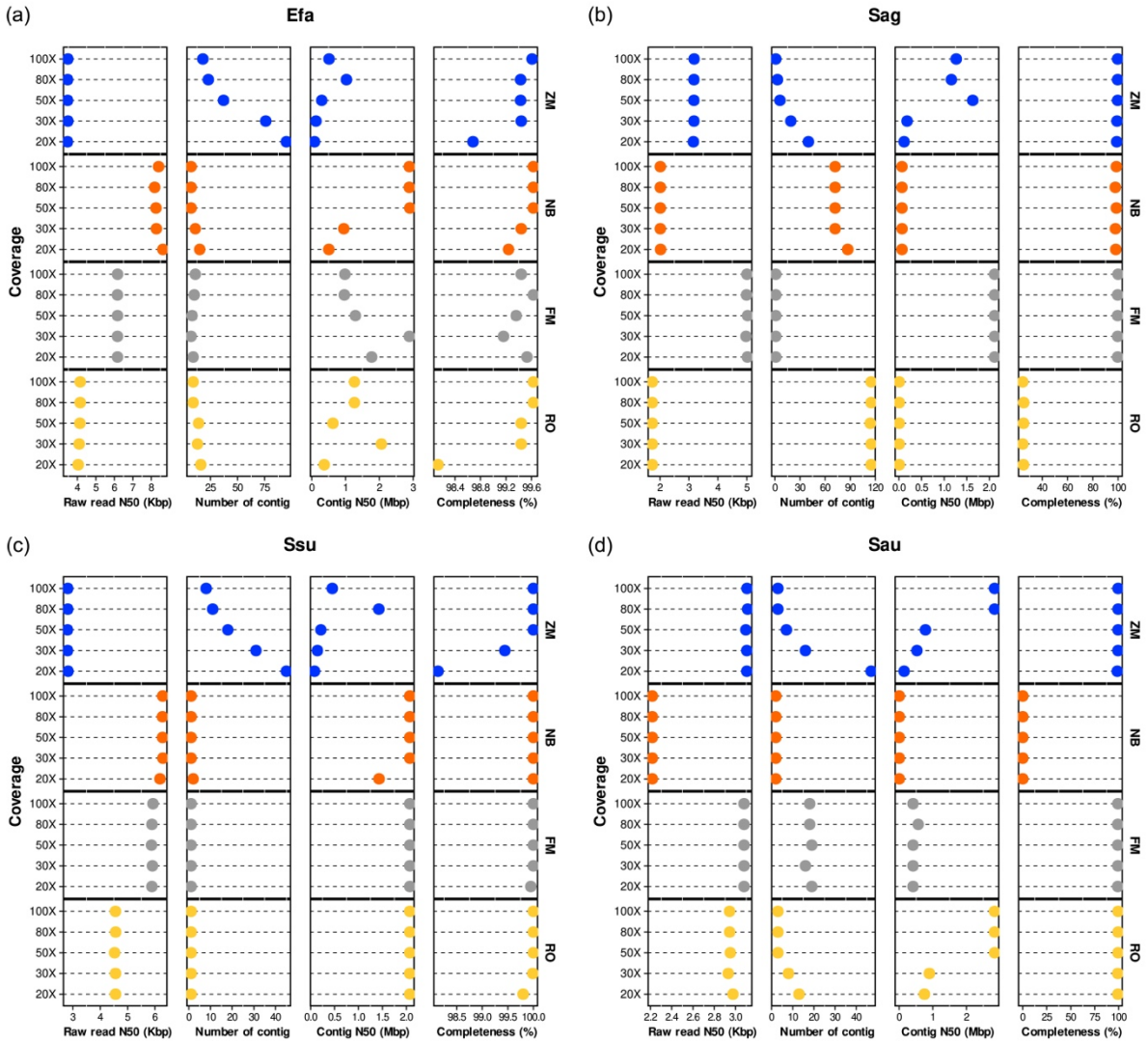

(a)
Efa
(b)
Sag
(c)
Ssu
(d)
Sau
Coverage
100X
80X
50X
30X
20X
ZM
NB
FM
RO
Raw read N50 (Kbp)
Number of contig
Contig N50 (Mbp)
Completeness (%)

Fig. 2

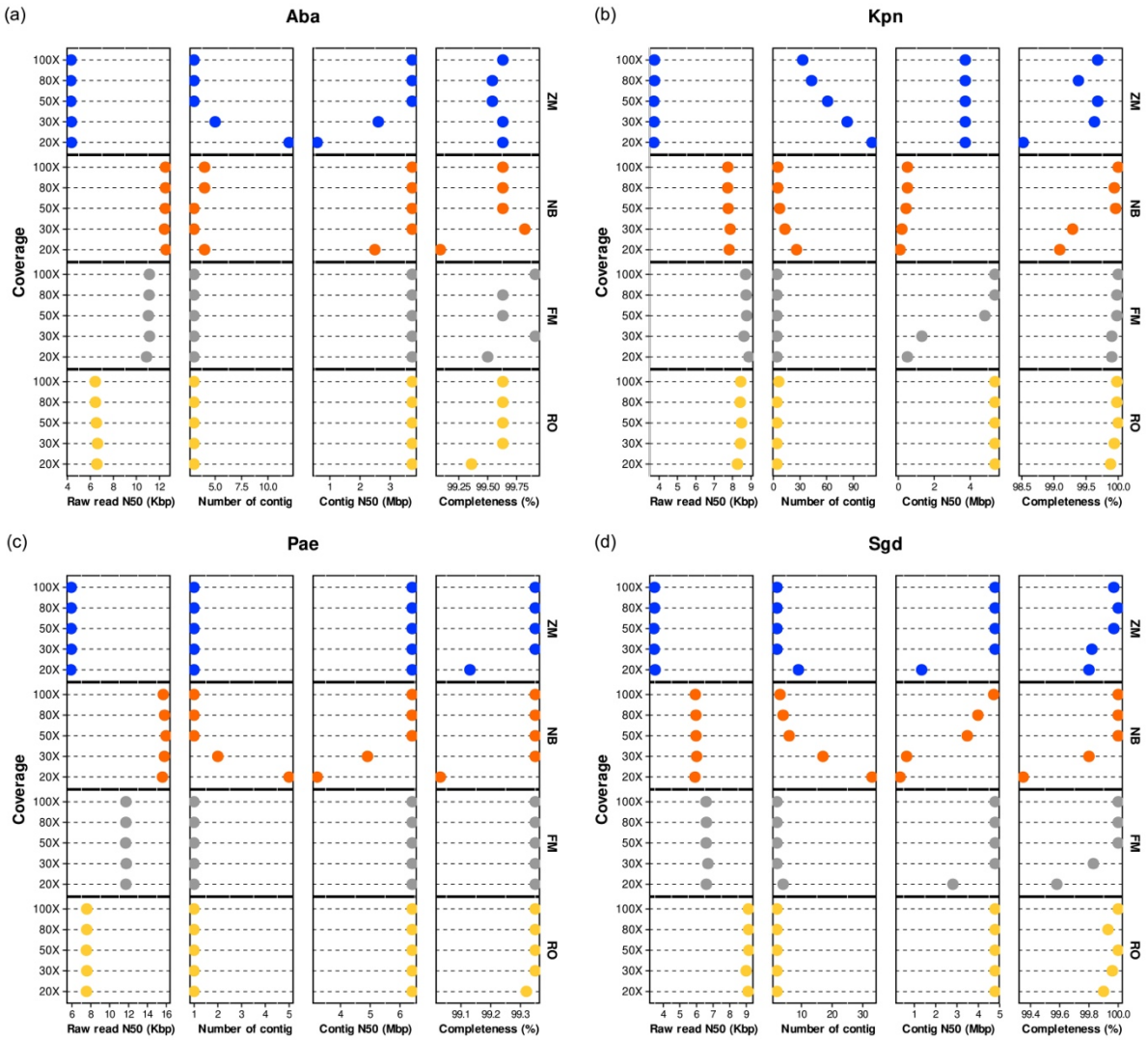

(a)
Aba
(b)
Kpn
(c)
Pae
(d)
Sgd
Coverage
100X
80X
50X
30X
20X
ZM
NB
FM
RO
Raw read N50 (Kbp)
Number of contig
Contig N50 (Mbp)
Completeness (%)

Fig. 3

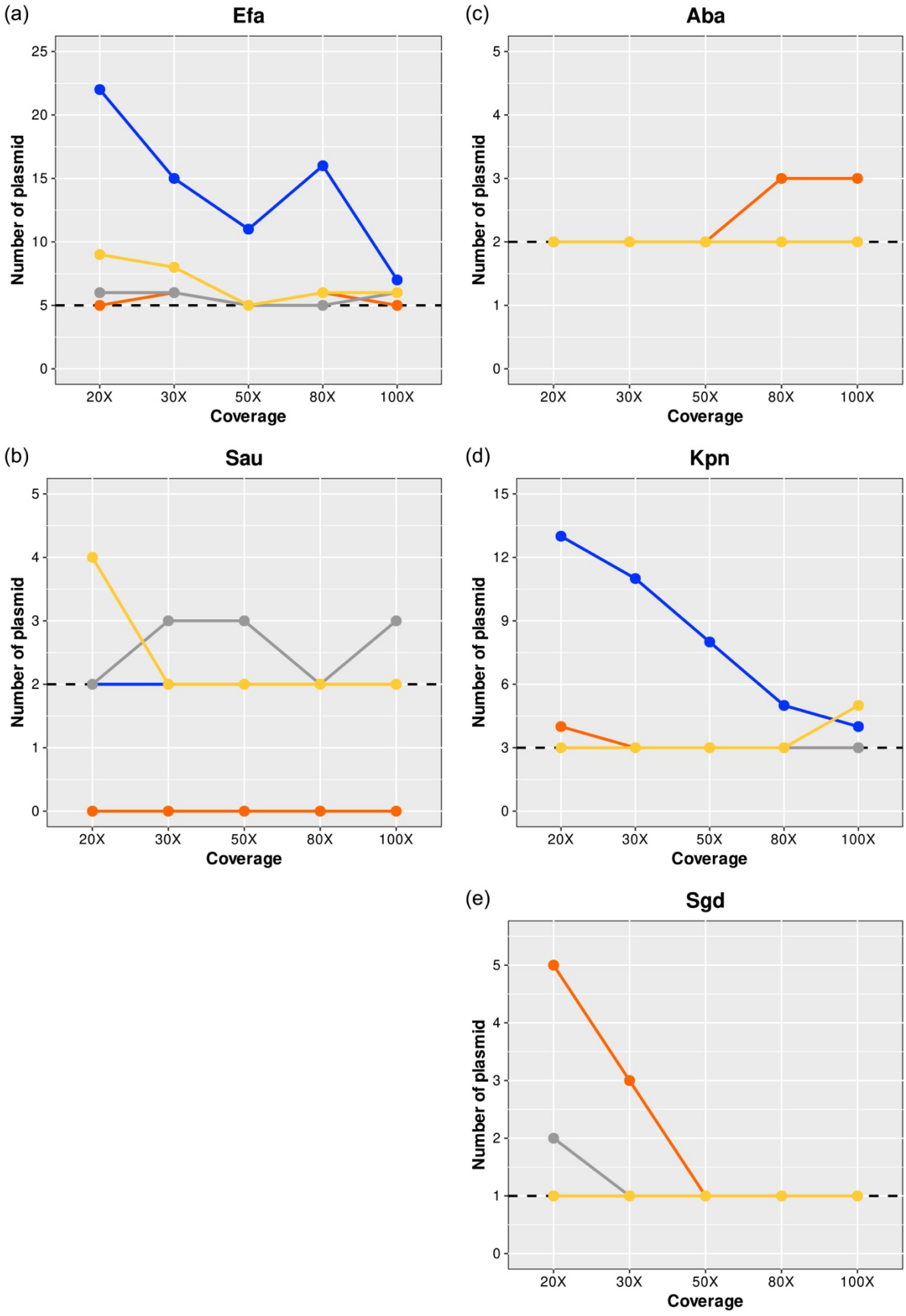

(a)
Efa
Number of plasmid
Coverage
20X
30X
50X
80X
100X
(c)
Aba
(b)
Sau
(d)
Kpn
(e)
Sgd

Fig. 4

Table Legends

Table 1 Summary of DNA isolation kit features used in this study

Table 2 Genome assembly statistics for all sequencing reads of eight pathogenic bacteria extracted by four different benchmark DNA extraction kits; ZymoBIOMICS™ DNA Miniprep Kit (ZM), Nanobind CBB Big DNA Kit (NB), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM) and MagNaPure 96 system (RO). The best assembly statistics from three independent replicates are shown. T, total contig; C, chromosome (chromosome contig); and P, plasmid (plasmid contig).

Figures

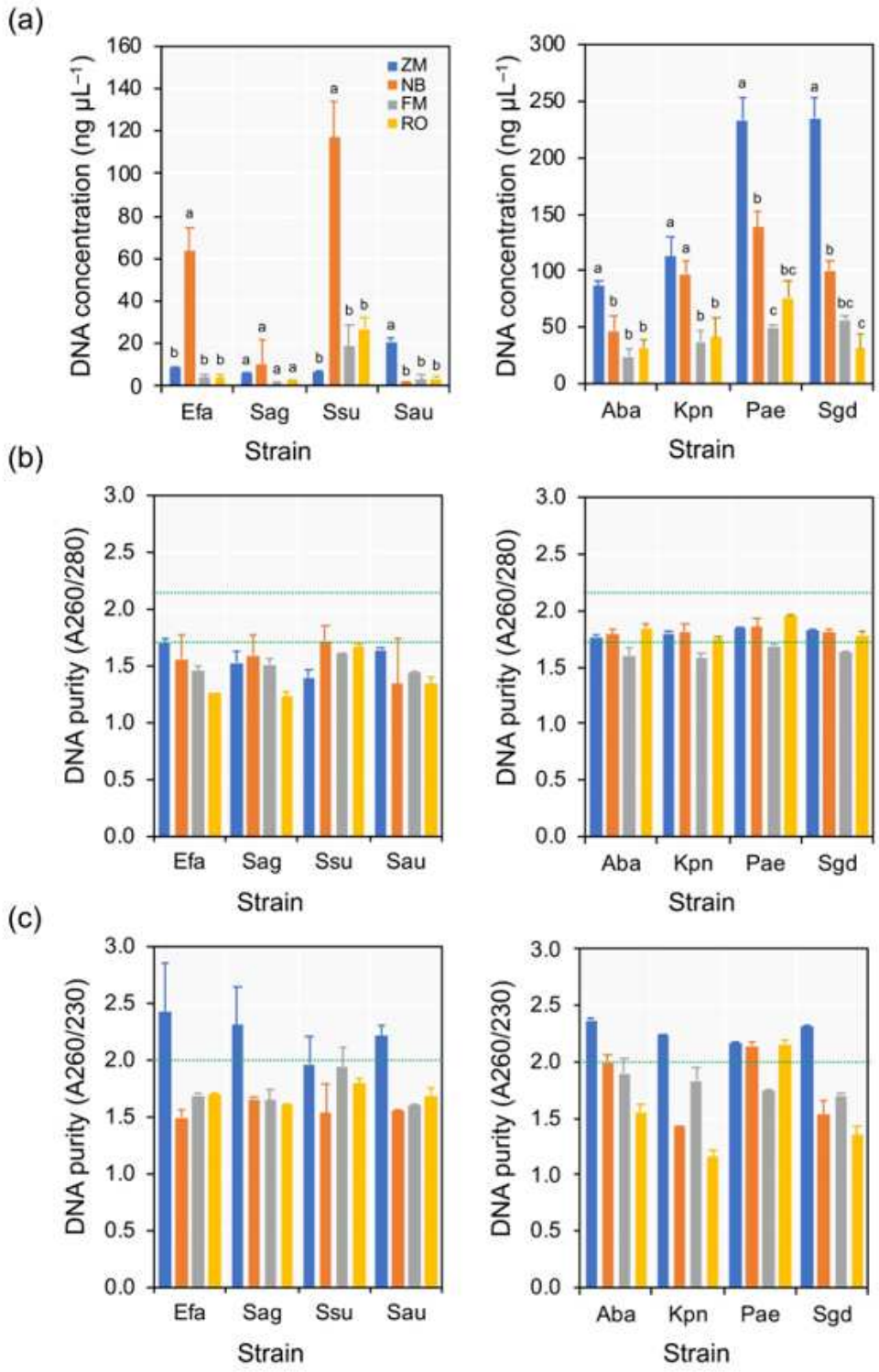


Figure 1

Influence of commercial DNA extraction kits; ZymoBIOMICS™ DNA Miniprep Kit (ZM, Blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and

MagNaPure 96 system (RO, yellow) on DNA concentration (ng μL^{-1}) (a) and DNA purity in ratio of A260/280 (b) and A260/230 (c) of eight pathogenic bacteria. Hatched green lines in (b) and (c) indicate recommended intervals.

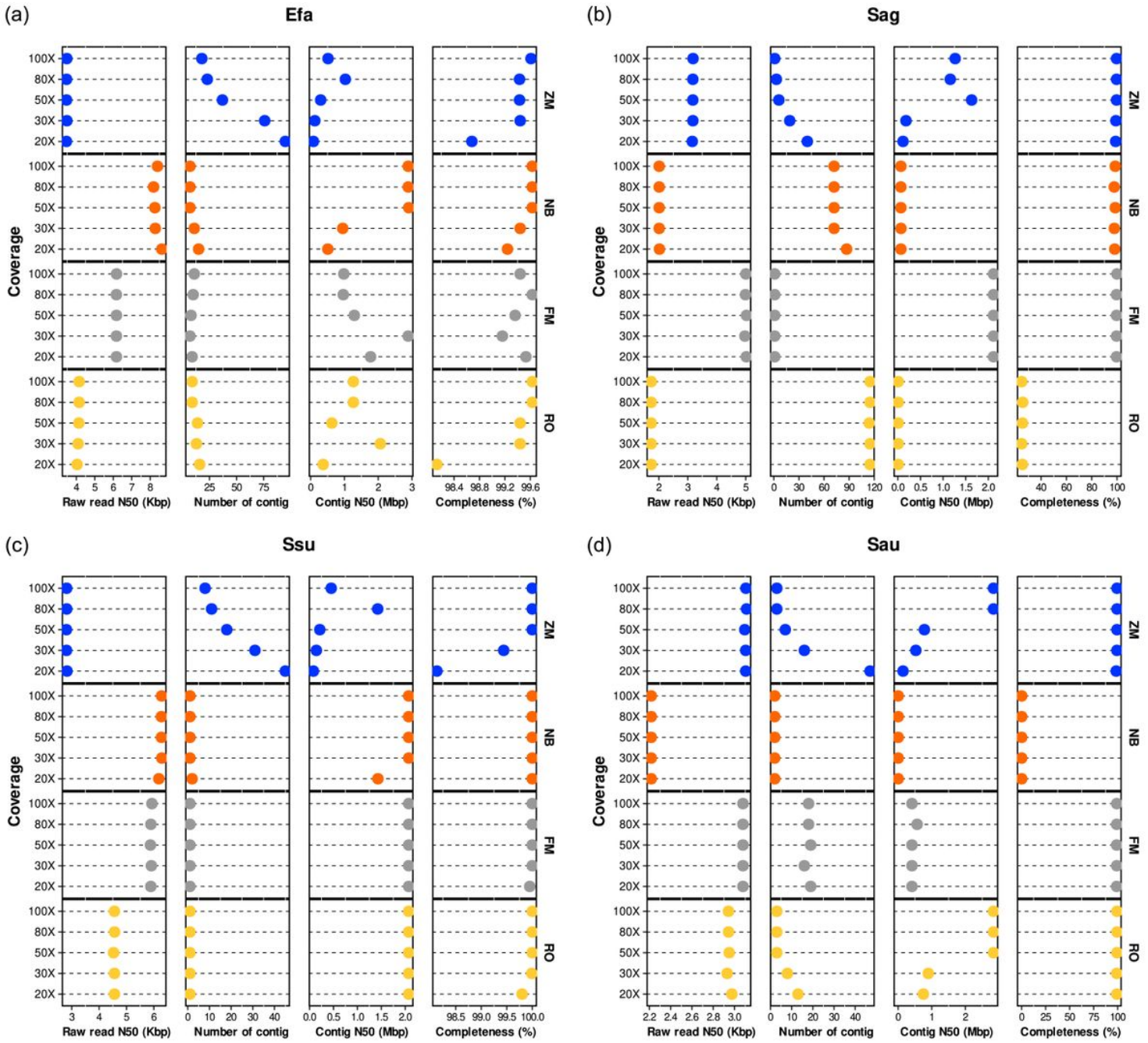


Figure 2

Influence of genome coverage on raw read length, number of contigs, percentage of completeness and contig *N50* of *Enterococcus faecium* (Efa, a), *Streptococcus agalactiae* (Sag, b), *Streptococcus suis* (Ssu, c), and *Staphylococcus aureus* (Sau, d) extracted by ZymoBIOMICS™ DNA Miniprep Kit (ZM, blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and MagNaPure 96 system (RO, yellow).

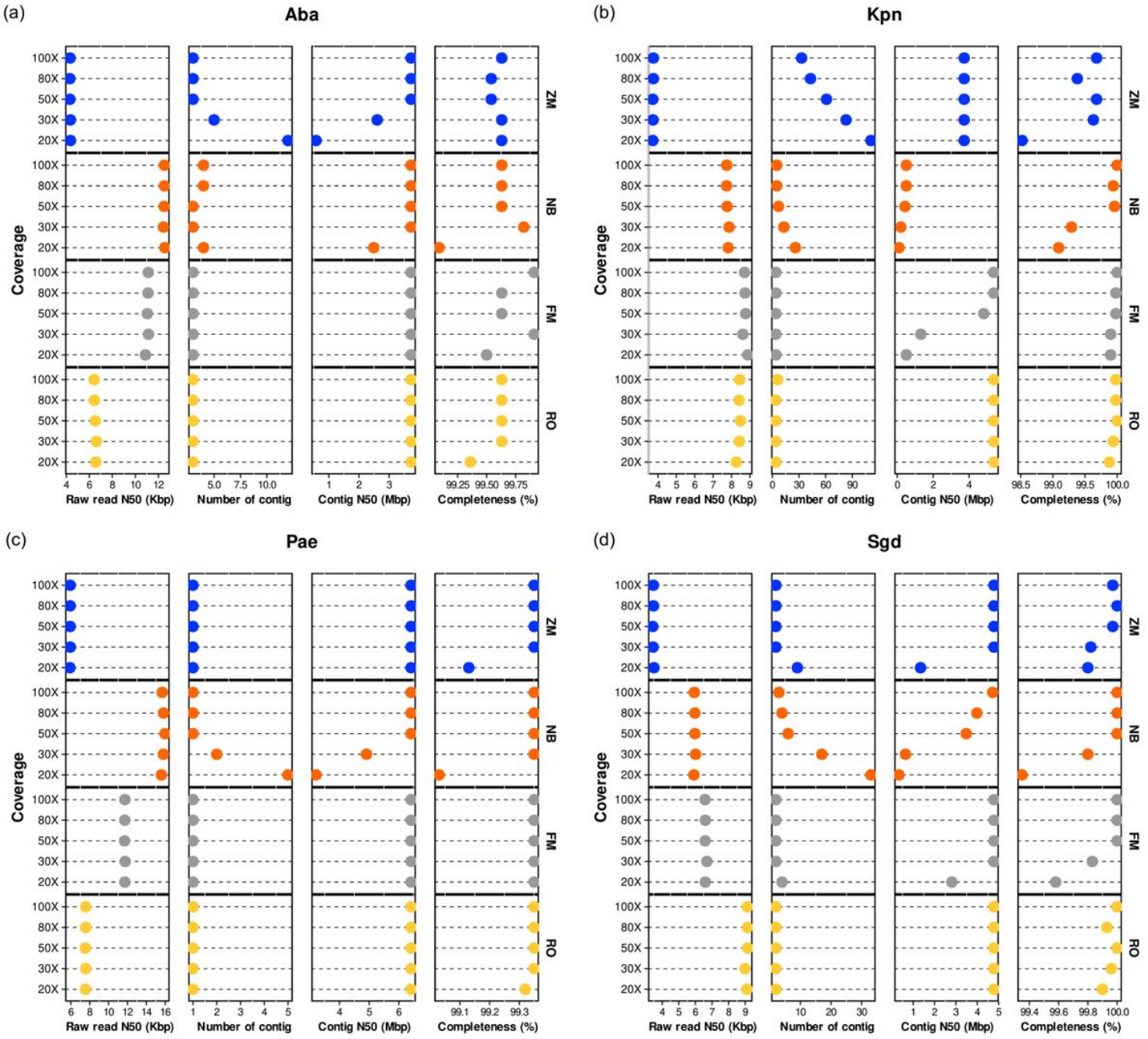


Figure 3

Influence of genome coverage on raw read length, number of contigs, contig *N50*, and percentage of completeness of *Acinetobacter baumannii* (Aba, a), *Klebsiella pneumoniae* (Kpn, b), *Pseudomonas aeruginosa* (Psu, c), and *Salmonella* sp. group D (Sgd, d) extracted by ZymoBIOMICS™ DNA Miniprep Kit (ZM, blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and MagNaPure 96 system (RO, yellow).

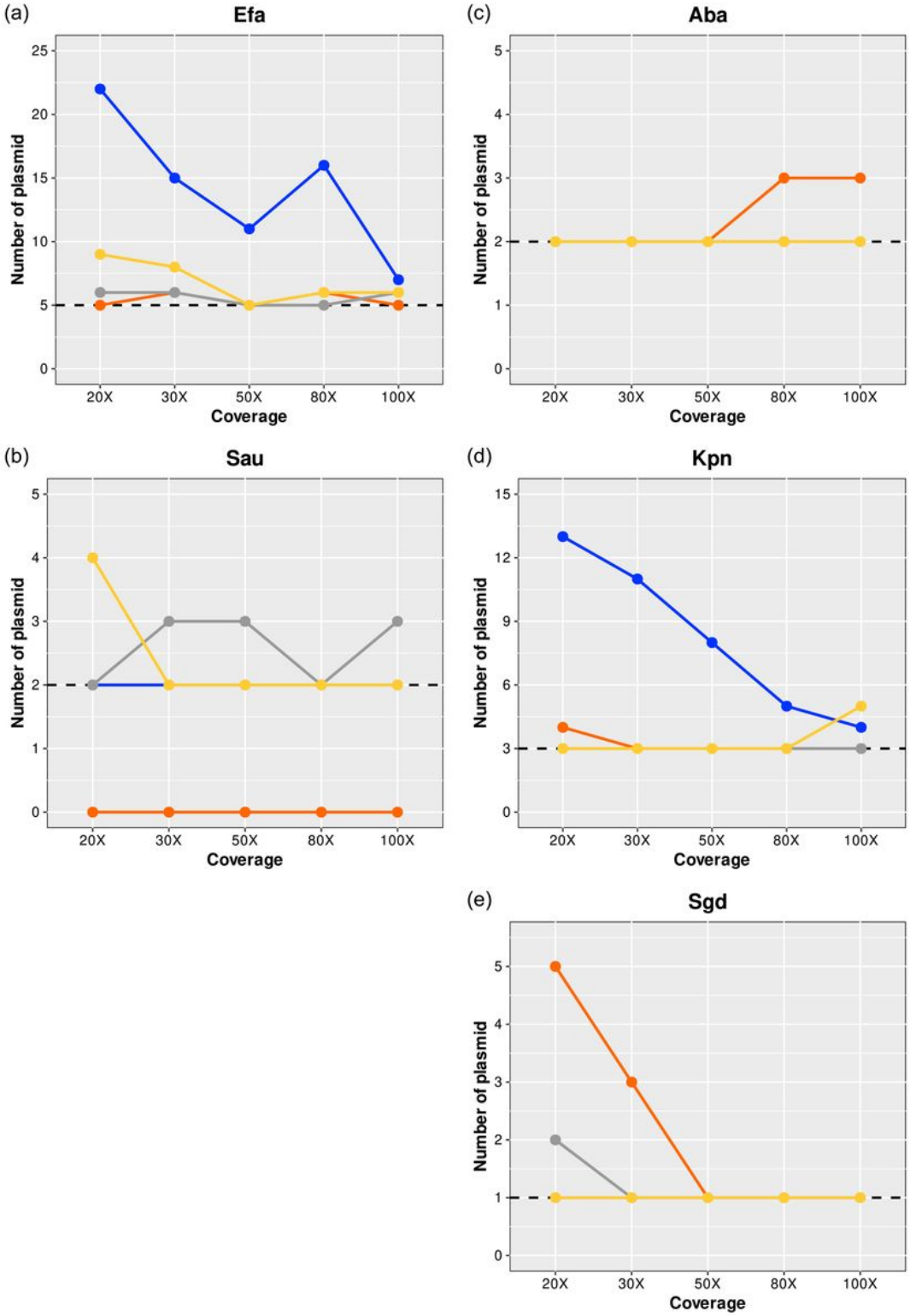


Figure 4

Influence of genome coverages on number of plasmid of *Enterococcus faecium* (Efa, a), *Staphylococcus aureus* (Sau, b), *Acinetobacter baumannii* (Aba, c), *Klebsiella pneumoniae* (Kpn, d), and *Salmonella* sp. group D (Sgd, e) extracted by ZymoBIOMICS™ DNA Miniprep Kit (ZM, blue), Nanobind CBB Big DNA Kit (NB, orange), Fire Monkey High Molecular Weight (HMW) DNA Extraction Kit (FM, grey) and MagNaPure 96 system (RO, yellow). Hatched black lines indicate numbers of plasmid in reference genomes.

Supplementary Files

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

- 3SupplementaryInformation.xlsx
- SubmittingsystemLegendsforsupplementaryfiguresandtables.docx