Asia-Pacific Journal of Science and Technology: Volume: 29. Issue: 02. Article ID.: APST-29-02-08. Research Article

APST

Asia-Pacific Journal of Science and Technology

https://www.tci-thaijo.org/index.php/APST/index

Published by the Research and Graduate Studies Division,
Khon Kaen University, Thailand

Detecting microbial communities and pathogens in gastrointestinal tracts of commercial broilers in Thailand using high-throughput sequencing technology with different bioinformatic pipelines

Athisri Sitthipunya[1], Pichahpuk Uthaipaisanwong[2], Nuananong Sinwat[3], Korntip Kanjanavaikoon[4], Supapon Cheevadhanarak[2,5] and Kanthida Kusonmano[2,6,*]

[1] Bioinformatics and Systems Biology Program, School of Bioresources and Technology, and School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
[2] Systems Biology and Bioinformatics Research Group, Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
[3] Department of Farm Resources and Production Medicine, Faculty of Veterinary Medicine, Kasetsart University, Kamphangsaen Campus, Nakornpathom, Thailand
[4] Huvepharma Ltd., Bangkok, Thailand
[5] School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
[6] Bioinformatics and Systems Biology Program, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand.
[*]Corresponding author: kanthida.kus@mail.kmutt.ac.th

Received 15 October 2022
Revised 6 February 2023
Accepted 8 February 2023

Abstract

High-throughput sequencing is widely applied to explore microbial communities and detect multiple pathogens simultaneously. The 16S rRNA amplicon sequencing is gaining attention in the poultry industry to monitor animal gut health and pathogens, especially *Salmonella*. A bioinformatics approach for the accurate characterization of the microbial communities obtained via high-throughput genetic sequences is crucial. This study provides a comparison between the commonly used bioinformatics tools QIIME2 and Mothur for the amplicon-based microbiome analysis of commercial broilers in Thailand and a mock community. We conducted QIIME2 and Mothur with the implementation of amplicon sequence variants (ASVs) and operational taxonomic units (OTUs) for grouping amplicon sequences as a unit of an organism, respectively. The two pipelines provided similar microbial profiles at all taxonomic levels. Most bacteria in the ileum samples belonged to the genera *Lactobacillus* and *Romboutsia*. However, different results were found in *Salmonella* detection. Based on the SILVA138 reference database and the same samples, Mothur provided the taxonomic assigned OTUs of the genus *Salmonella*, whereas QIIME2 could assign the taxonomy for the identified ASVs only at the level of the family Enterobacteriaceae. We recommend further annotating the ASVs of the family Enterobacteriaceae using a phylogenetic tree and basic local alignment search tool (BLAST) to discriminate the genus *Salmonella*. This study demonstrates that caution is required for *Salmonella* detection when performing 16S rRNA amplicon sequencing analysis for the best interpretation output of the tools.

Keywords: Chicken gut microbiome, Mothur, QIIME2, *Salmonella* detection, Commercial broilers

1. Introduction

Studying multiple bacterial species or gastrointestinal tract (GI) microbiomes of animals has become a challenge because only a small proportion of the GI microbiota can be grown and studied by conventional culturing techniques [1, 2]. High-throughput sequencing serves as a solution to culturing limitations by sequencing all genetic materials of microbes, including pathogens [2, 3]. Sequencing-based approaches, including specific

amplicon and metagenomic sequencing, have been widely used to study microbiota compositions, their metabolisms, and interactions. Amplicon sequencing of the 16S ribosomal (16S rRNA) RNA gene is frequently applied to identify gut microbiota and is more cost-efficient than metagenomic sequencing. In addition, the continuous update of the public 16S rRNA gene databases facilitates targeted bacterial and archaeal identification [4, 5].

The amplicon sequencing technique has successfully been applied in livestock farms and industries, including the poultry industry, to study and monitor the gut microbiome, which is significantly correlated to animal health and productivity [6]. The gut microbiome comprises around 300-1,000 taxa of both beneficial and harmful microbes [7, 8]. The majority of the detected phylum are Firmicutes, Proteobacteria and Actinobacteria [7-9]. Revealing the community members facilitates the identification of beneficial gut microbiota, which help to metabolize host nutrients, induce immunity, defend pathogens, and stabilize chicken gut health [10, 11]. In this sense, understanding the gut microbiome of chickens could enable us to identify probiotics to improve chicken health and productivity, which could be used as an alternative to subtherapeutic antibiotics [12]. Moreover, the technology provides an alternative monitoring method for pathogens such as *Salmonella* and *Campylobacter*, which is crucial in poultry production and exportation [13]. The 16S rRNA sequencing detects multiple pathogens simultaneously and is more time-efficient than culturing methods which require specific media and conditions for each pathogen and could detect one species at a single time.

When conducting 16S rRNA amplicon sequencing, a classification rank of bacteria is based on the nine variable regions (hypervariable regions, V1–V9) in the 16S rRNA gene [4, 14]. Usually, one or two consecutive hypervariable regions are selected as a target and used via the short-read sequencing technology (often 250 or 300 base pairs long), which is sufficient to differentiate microbes at the genus level [15]. Bioinformatic pipelines have been developed to characterize the amplicon sequences into microbial taxonomy. One approach is to cluster amplicon sequences based on their similarity, with 97% similarity threshold for grouping microbial sequences at genus level. Each identified cluster is referred to as an operational taxonomic unit (OTU) [16]. Another approach is the denoising approach, which discriminates the amplicon sequences by one nucleotide differentiation, the so-called "amplicon sequence variants" (ASV) [17]. For either the OTU or the ASV approach, the sequence unit will be subjected to annotate its taxonomy by searching against a database of the amplicon gene. Currently, there are two commonly used bioinformatics tools for amplicon-based microbiome analysis, namely QIIME2 [18] and Mothur [19]. With the derived high-throughput sequences, a bioinformatic pipeline for the accurate characterization of microbial communities from these short genetic sequences is important. Different tools and analytical approaches could yield different resulting microbiome profiles [20, 21]. Nevertheless, there is no explicit investigation of bioinformatic pipelines for detecting microbiome including pathogens in commercial broilers in Thai industry.

This work compares bioinformatics pipelines for microbiome analysis using 16S rRNA amplicon sequencing to characterize the commercial broiler gut microbiome and pathogens such as *Salmonella* and *Campylobacter* in the context of industrial farming in Thailand. We investigated the commonly used bioinformatic tools for amplicon-based microbiome analysis, QIIME2 and Mothur, with the implementation of ASV and OTU approaches, respectively. This paper provides information about the use of an amplicon-based sequencing data analysis pipeline for chicken gut microbiome studies.

2. Materials and methods

2.1 Chicken ileum sample

The intestinal contents of ileum samples were aseptically collected from three healthy commercial broilers from the same farm. The chickens had been fed with basal chicken diet addition containing an equal dose of *Bacillus subtilis* (*B. subtilis*) probiotics, administered daily, and were sacrificed at day 45. Total genomic DNA was extracted using the QIAamp Fast DNA Stool Mini Kit (QIAGEN, Germany) according to the manufacturer's procedures. Library preparation and 16S rRNA gene sequencing for the extracted genomic DNA of chicken ileum were performed by an external laboratory (NovogeneAIT Genomics, Singapore). The V4 region of the 16S rRNA was amplified using the primer pair 515 F (5′-CTAGTGCCAGCMGCCGCGGTAA-3′) and 806 R (5′-CTAGGACTACHVGGGTWTCTAAT-3′). Finally, the 16S rRNA amplicons were sequenced using the Illumina MiSeq platform. The chicken ileum 16S rRNA gene sequencing data have been submitted to the NCBI Sequence Read Archive (SRA) database under the accession number PRJNA889347.

2.2 Mock sample

A mock microbial community, consisting of eight bacterial and two yeast species (ZymoBIOMICS™ Microbial Community standard D6300, Zymo Research, USA), was included in the analysis. The microbial theoretical composition of only 16S rRNA gene of bacterial species includes 4.2% *Pseudomonas aeruginosa*,

10.1% *Escherichia coli*, 10.4% *Salmonella enterica*, 18.4% *Lactobacillus fermentum*, 9.9% *Enterococcus faecalis*, 15.5% *Staphylococcus aureus*, 14.1% *Listeria monocytogenes*, and 17.4% *Bacillus subtilis*. Total genomic DNA was extracted using the QIAamp Fast DNA Stool Mini Kit (QIAGEN, Germany) according to the manufacturer's procedures. The genomic DNA of the mock sample was concentrated using the Genomic DNA Clean & Concentrator-10 Kit (Zymo Research, USA). Library preparation and 16S rRNA gene sequencing for the extracted genomic DNA was performed by an external laboratory (Génome Québec, Canada). The V4 region of 16S rRNA was amplified using the primer pair 515 F (5′-CTAGTGCCAGCMGCCGCGGTAA-3′) and 806 R (5′-CTAGGACTACHVGGGTWTCTAAT-3′). Finally, the 16S rRNA amplicons were sequenced using the Illumina MiSeq platform.

2.3 Data assessment and preparation

To obtain the microbial profiles from the amplicon sequencing data, 250-bp paired-end raw read sequences from the V4 hypervariable region of the 16S rRNA gene were generated using QIIME2 v. 2021.4 (https://qiime2.org) [18] and Mothur v. 1.43.0 (https://mothur.org) [19]. The quality of the raw read sequences was assessed using FastQC v. 0.11.9 [22]. The software determined the sequence quality, including base quality score, length of sequence read, and GC content. Subsequently, the Illumina adapter and primer sequences of V4 were removed using Cutadapt v. 4.0 [23].

2.4 QIIME2 analysis

Data preprocessing was done using DADA2 [24], divisive amplicon denoising algorithm by filtering sequence error and denoising with default parameters. Chimeric sequences were also removed in this step. The forward and reverse reads were truncated (180 and 210 bps, respectively, for the chicken ileum samples and 100 and 200 bps, respectively, for the mock samples) and merged to obtain the maximum read counts for downstream analysis. The derived ASVs were used for diversity analysis and taxonomic assignment. The 16S rRNA genes of the SILVA 138 database (https://www.arb-silva.de) [5] were used for training the classifier for the V4 region, and subsequently, taxonomic classification of the representative sequence was performed by the sklearn classifier. Normalization was performed before the taxonomic assignment, based on the rarefying method. The workflow of the 16S rRNA amplicon analysis using QIIME2 is shown in Figure 1.

2.5 Mothur analysis

The *de novo* clustering method of Mothur was performed. For data preprocessing, paired-end reads were merged into contigs, and low-quality reads were filtered and removed (Avg. > Q30 after removal). Additionally, non-target sequence removal was performed based on the length between 248 and 302 bps (min-max) at the alignment position between 13,875 and 23,440. The aligned sequences were based on the 16S rRNA genes of the SILVA 138 database (https://www.arb-silva.de) [5]. Singleton, doubleton and chimeral sequences were removed before clustering to OTUs, based on their genetic distance. Normalization was performed before the taxonomic assignment, based on the rarefying method. The process of 16S rRNA amplicon analysis using Mothur is shown in Figure 1.

2.6 Taxonomic assignment

The representative sequences of ASVs from QIIME2 and OTUs from Mothur were assigned taxonomy based on the SILVA 138 database (https://www.arb-silva.de) [5]. We conducted the formatted database that was compatible with each tool. For Mothur, the formatted database was available at https://mothur.org/wiki/silva_reference_files/.

2.7 Diversity analysis

Both QIIME2 and Mothur were used to perform alpha diversity analysis measuring microbial community diversity within a sample. Chao1 and Shannon diversity indices were calculated to estimate microbial richness and evenness, respectively. Differential abundance was measured using Wilcoxon rank-sum test. For beta diversity analysis, Bray-Curtis dissimilarity matrices between the studied microbial abundance profiles were calculated and visualized using principal coordinates analysis (PCoA). The microbiome profiles between QIIME2 and Mothur were integrated at the genus level. Permutational multivariate analysis of variance was used to measure significant differences between microbial profiles derived from the different analysis methods.

2.8 Phylogenetic tree analysis

The phylogenetic tree of ASVs was analyzed using the MEGA X software [25]. Alignment was performed using ClustalW [26], and the sequence relationship was constructed based on the Neighbor-Joining Tree algorithm with 10,000 bootstraps. The 16S rRNA reference sequence of three *Escherichia coli* (NCBI Sequence ID: NR 024570.1, NR112558.1, and NR 114042.1), eight *Salmonella enterica* (NCBI Sequence ID: NR 044372.1, NR 044373.1, NR 074910.1, NR 074799.1, NR 104709.1, NR116125, NR119108.1, and NR 116126.1), and two *Salmonella bongori* (NCBI Sequence ID: NR 074888.1 and NR 116124.1) were selected as representatives of the family Enterobacteriaceae. *Lactobacillus* sp. (NCBI Sequence ID: E 10214.1) and *Bacillus subtilis* (NCBI Sequence ID: NR 112116.2) were selected as representatives of family Lactobacillaceae and Bacillaceae, respectively. They were used as the outgroups in the phylogenetic tree analysis.

2.9 BLAST (Basic Local Alignment Search Tool) analysis

The ASV sequences were annotated based on sequence similarity, using the nucleotide BLAST against the NCBI non-redundant nucleotide database (https://blast.ncbi.nlm.nih.gov/). The best hit results of each query ASV are shown with percent similarity and E-value (expect-value).

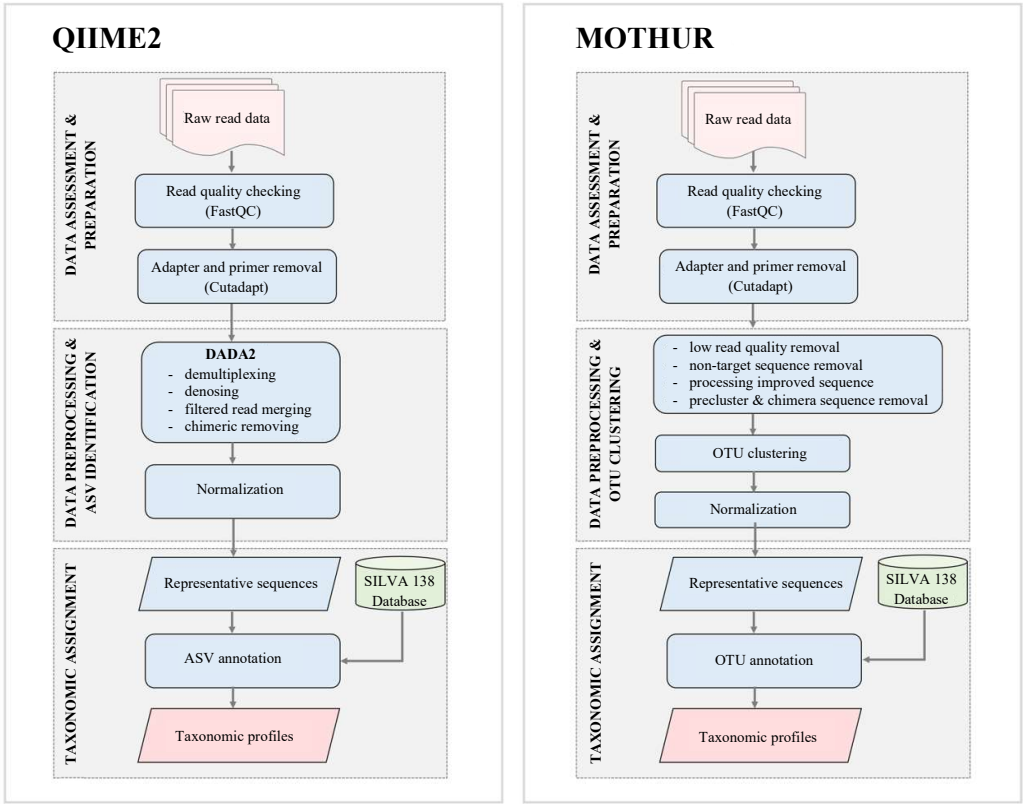

Figure 1 Workflow of the 16S rRNA amplicon analysis pipeline using QIIME2 (left) and Mothur (right). ASV, amplicon sequence variants; DADA2, divisive amplicon denoising algorithm 2; OTU, operational taxonomic unit.

3. Results

3.1 Chicken ileum microbiome

The gut microbiome of commercial broilers was analyzed based on the V4 hypervariable region of 16S rRNA sequencing data. The maximum, minimum, and average values of raw sequence reads of the three chicken ileum samples were 210,493, 184,495, and 194,031, respectively. The QIIME2 analysis resulted in a total of 1,078 ASVs, whereas Mothur analysis gave a total of 1,957 OTUs. The average estimated richness values based on the Chao1 index from QIIME2 and Mothur are 523.00 and 1422.57, and the average estimated evenness values based

on the Shannon index are 3.96 and 2.46, respectively. Although the diversity indices were different, the microbiome profiles from both analysis pipelines showed similar patterns for the resulting taxonomic profiles (Figure 2-4). The detected microbiome profiles from QIIME2 and Mothur are not significantly different (p-value=0.7, R^2=0 0.01535). Figure 4 shows that microbiome profiles of the same samples detected from different pipelines are more similar than among individual samples. Firmicutes was the most prevalent phylum and found in more than 80% of all phyla in the chicken ileum samples, followed by Bacteroidota (~0.5%) and Proteobacteria (~1%). At the family level, Lactobacillaceae was the most representative family, accounting for more than 90% of the families in the Chicken_3 sample. In addition, common microbial families found in chicken ileum were Peptostreptococcaceae, Clostridiaceae, Streptococcaceae, Lachnospiraceae, and Ruminococcaceae. The most common genus was *Lactobacillus* (a member of Lactobacillaceae), followed by *Romboutsia* and *Terrissporobacter* (members of Peptostreptococcaceae) (Figure 2). Clostridiaceae showed slightly different proportions of 0.015% and 0.016% by QIIME2 and Mothur, respectively. The top 10 families and genera characterized by QIIME2 and Mothur are shown in Figures 2 and 3, respectively. There were no differential abundances of any genera detected by the two pipelines (p-value > 0.05). The probiotic genus of *Bacillus* in the feed was detected with low relative abundance (<0.01%). Moreover, we investigated the pathogenic bacteria *Campylobacter* and *Escherichia-Shigella*. Average relative abundances of 0.085% and 0.075% of *Campylobacter* were detected by QIIME2 and Mothur, respectively. *Escherichia-Shigella* showed a relative abundance of 0.28% by QIIME2 and 0.27% by Mothur. For *Salmonella*, relative abundance was only reported by Mothur, at 0.04%.

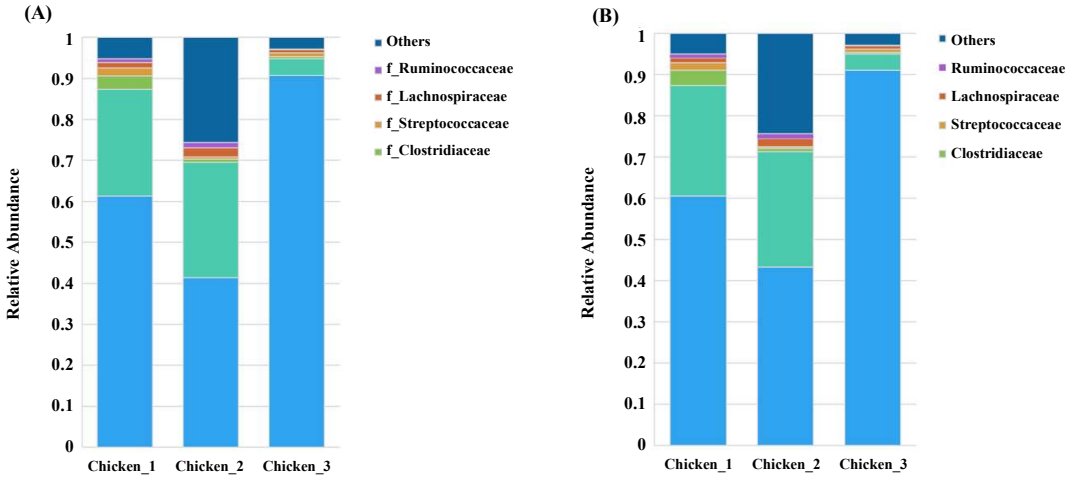

Figure 2 Relative abundances of the top 10 families of the broiler ileum microbiome based on QIIME2 (A) and Mothur (B) analyses.

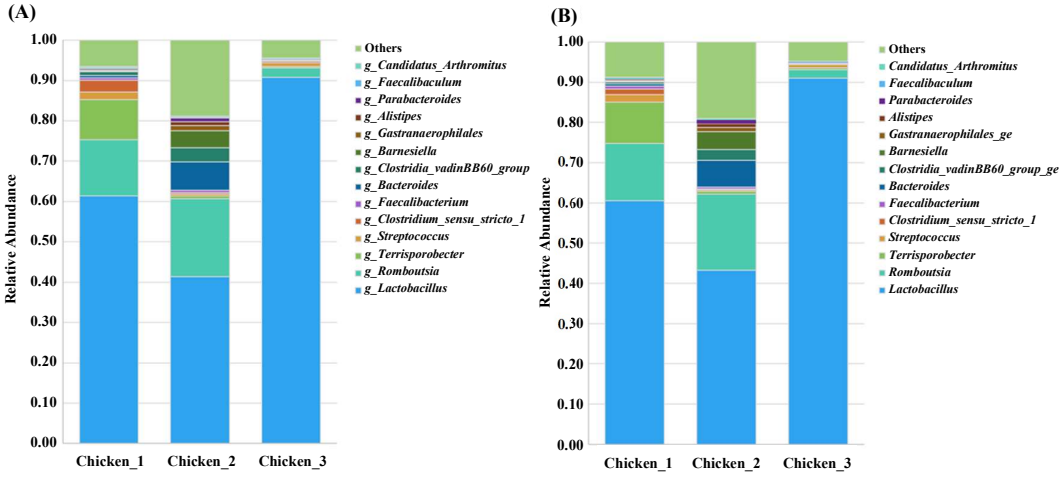

Figure 3 Relative abundances of the top 10 genera of the broiler ileum microbiome based on QIIME2 (A) and Mothur (B) analyses.

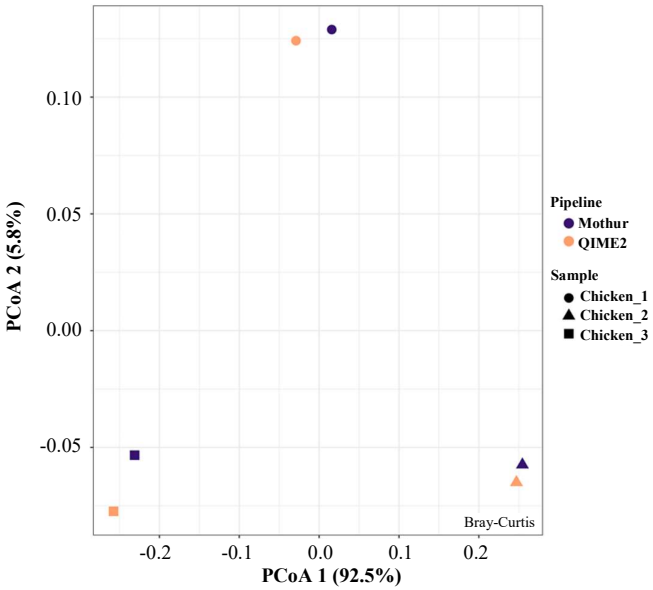

Figure 4 Principal coordinates analysis (PCoA) visualizing microbiome profiles of commercial broilers detected by QIIME2 (Orange color) and Mothur (Purple color) analyses. Each microbiome sample of studied broilers is represented by a different symbol (Circle for Chicken_1, triangle for Chicken_2, and square for Chicken_3).

3.2 Mock sample community

The mixture of eight bacterial species from the mock community was subjected to 16S rRNA amplicon sequencing, and the raw read sequences were analyzed in the amplicon analysis pipelines, QIIME2 and Mothur. The theoretical microbial composition of the 16S rRNA gene of mock bacterial species is shown in Figure 5. The 16S rRNA analysis showed the microbial composition of the eight genera in the two mock samples (replicates). All eight microbial units (ASV or OTU) were detected; however, the relative abundance of each genus was shifted from the theoretical composition. Nevertheless, the resulting profiles were similar for the two analytic pipelines QIIME2 and Mothur. There is no significant difference between the profiles detected by the two pipelines (p-value=0.67, R^2=0.3234). The genera *Lactobacillus* (a member of Lactobacillaceae) and *Pseudomonas* (a member of Pseudomonadaceae) were found at more than 50% of the abundance proportion of the whole community. A slightly higher abundance of *Lactobacillus* was reported by QIIME2 compared to Mothur. About 30% of the abundance proportion was from the family Enterobacteriaceae. Mothur detected three genera of the family Enterobacteriaceae, namely *Salmonella*, *Escherichia-Shigella*, and *Enterococcus*. In contrast, QIIME2 detected only *Escherichia-Shigella* and *Enterococcus*. The genera *Bacillus* (a member of Bacillaceae), *Staphylococcus* (a member of Staphylococcaceae), and *Listeria* (a member of Listeriacaea) accounted for approximately 20% of the whole community. The relative abundances of the mock microbial profiles at genus level, characterized by QIIME2 and Mothur, are shown in Figure 5.

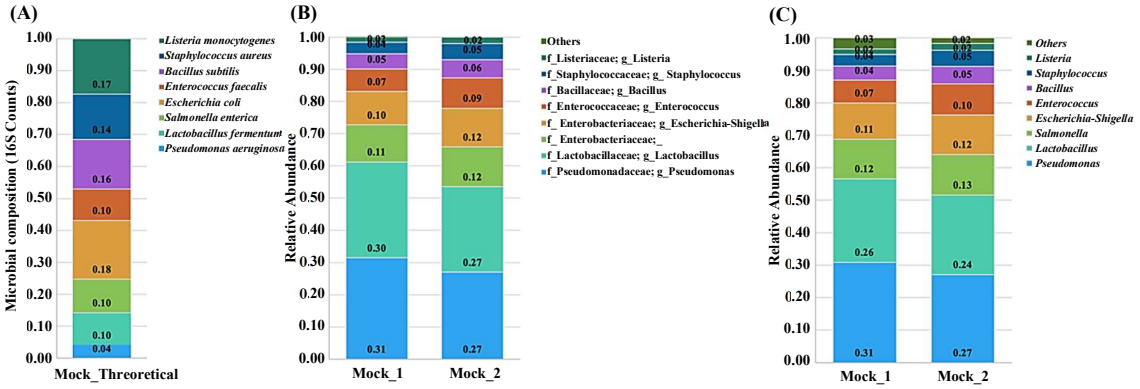

Figure 5 Theoretical mock community composition (A) and the detected mock community compositions at genus level (two replicates) based on QIIME2 (B) and Mothur (C) analyses.

3.3 Salmonella identification

Based on the QIIME2 and Mothur analyses, we considered the ASVs or OTUs that were taxonomically assigned to the family Enterobacteriaceae of both chicken ileum and mock samples (Table 1). For both ileum and mock microbiome datasets, the pathogenic amplicon units (ASV or OTU) of the genus *Escherichia-Shigella* could be annotated. However, the units of the genus *Salmonella* could be assigned by only the Mothur pipeline. Some ASVs from the QIIME2 analysis could not be taxonomically assigned at genus level, as shown in Table 1.

Table 1 The abundance of each genus of the family Enterobacteriaceae in the chicken ileum and mock samples identified by QIIME2 and Mothur pipelines, respectively.

Pipelines	Taxonomy assignment ranking		Abundance of genus				
			Chicken 1	Chicken 2	Chicken 3	Mock 1	Mock 2
	Family	Genus					
QIIME2	Enterobacteriaceae	*Escherichia-Shigella*	93	945	64	6,993	7,513
	Enterobacteriaceae	unassigned	68	77	40	6,311	7,321
Total normalized abundance in a sample			127,687	130,546	128,829	60,917	60,917
Mothur	Enterobacteriaceae	*Escherichia-Shigella*	67	807	44	6,529	7,137
	Enterobacteriaceae	*Salmonella*	52	53	30	7,216	7,382
Total normalized abundance in a sample			114,887	114,770	114,871	59,055	59,056

From Table 1, the unassigned genus comprised one ASV (Chicken ASV 1) from the chicken ileum dataset and two ASVs (Mock ASV 1 and Mock ASV 2) from the mock dataset. On the other hand, two ASVs (Chicken ASV 2 and Chicken ASV 3) from the chicken ileum dataset and one ASV (Mock ASV 3) from the mock dataset were assigned as *Escherichia-Shigella*.

We further investigated the ASVs that could not be assigned at genus level. Representative sequences of ASVs from the microbiome analysis using QIIME2 were retrieved for the phylogenetic tree analysis. The phylogenetic tree was constructed with 10,000 bootstraps (Figure 6). The ASV sequences of Chicken ASV 1, Mock ASV 1, and Mock ASV 2 were closely related to a group of *Salmonella enterica*, and these sequences came from the unassigned genus of Enterobacteriaceae. In contrast, the ASV sequences of Chicken ASV 2, Chicken ASV 3, and Mock ASV 3 were closely related to a group of *Escherichia* species and came from the assigned genus *Escherichia-Shigella* of Enterobacteriaceae. Likewise, the unassigned ASVs were searched by conducting an online NCBI BLAST against the non-redundant nucleotide database (nr/nt) database. The BLAST analysis revealed that all unassigned ASVs were 99.0%–100% identical to *Salmonella enterica*. The best BLAST hits of the unassigned ASVs are shown in Table 2.

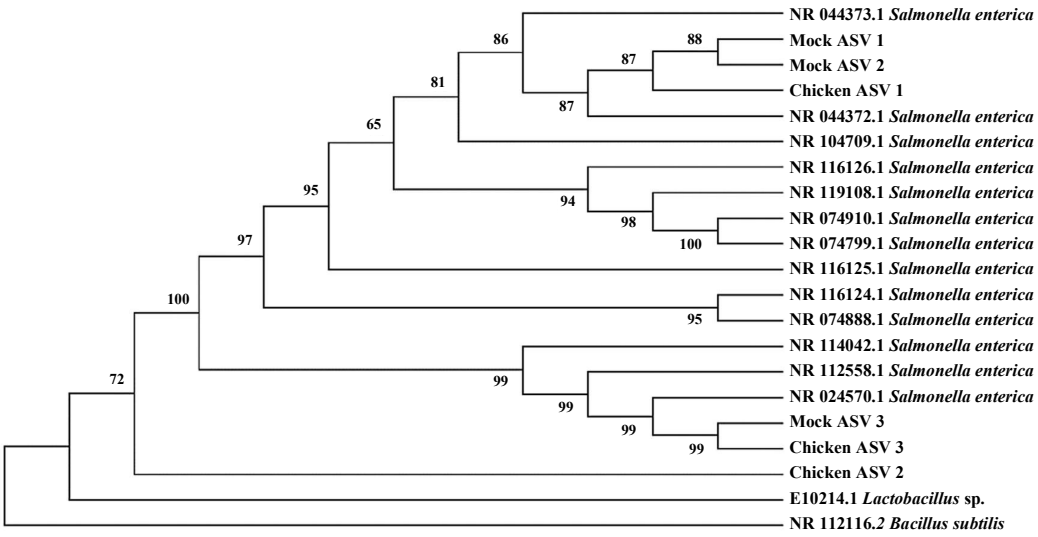

Figure 6 Phylogenetic tree of 15 16S rRNA reference sequences and 3 ASVs from the chicken ileum samples (Chicken ASV 1, Chicken ASV 2, and Chicken ASV 3) and 3 ASVs from the mock samples (Mock ASV 1, Mock ASV 2, and Mock ASV 3). All ASVs belong to the family Enterobacteriaceae. Chicken ASV 1, Mock ASV 1, and Mock ASV 2 were unassigned ASVs at the genus level. Chicken ASV 2, Chicken ASV 3, and Mock ASV 3 were annotated as *Escherichia-Shigella*. The number in each branch represents the percentage of consensus in the tree, based on bootstrapping.

Table 2 Best BLAST hits of unassigned ASVs belonging to the family Enterobacteriaceae.

Representative Sequence	Percent identity	E-value	Best blast hit
Chicken ASV 1	100.00	2e-127	*Salmonella enterica* subsp. enterica (Accession ON365723.1)
Mock ASV 1	100.00	8e-126	*Salmonella enterica* subsp. enterica (Accession CP012344.2)
Mock ASV 2	100.00	8e-126	*Salmonella enterica* subsp. enterica (Accession ON365723.1)

Chicken ASV = Amplicon sequence variant from the chicken samples; Mock ASV = Amplicon sequence variant from the mock samples; E-value = expected value.

4. Discussion

Sequencing of the 16S rRNA gene amplicon has been applied for the study of the chicken gut microbiome, which facilitates the improvement of chicken feed and performance [27, 28]. The technology overcomes the limitation of the culturing method and allows the simultaneous detection of the microbial community and its proportion [29]. This provides an opportunity to detect and monitor beneficial gut microbes together with pathogens for improving poultry production. Cultivation methods are generally time and labor-consuming, and the profiling of ileum and cecum samples from broiler chickens by conventional culturing analysis results in large numbers of undetected genera [29]. Recently, high-throughput sequencing has been used to detect and track microbiome communities and the infectious pathogens *Salmonella, Campylobacter*, and *Escherichia coli* in poultry [30, 31]. However, it is noted that the use of a sub-region of the 16S rRNA gene provides reliable microbial classification only at a taxonomic genus level [15]. The technology could facilitate the screening for potential risks of pathogens. A higher resolution classification could be complementary utilized by other techniques considering more distinguishable genomic regions and/or biochemical properties [32].

When conducting high-throughput sequencing, a bioinformatics pipeline for accurate microbial detection is crucial. Nowadays, there are two commonly used bioinformatics tools for 16S rRNA microbiome analysis, namely QIIME2 and Mothur. These two tools have already been compared and discussed as a choice of analysis in different datasets [20, 21]. In our study, QIIME2 and Mothur analyses provided highly similar profiles of the commercial broiler ileum microbiome from a farm in Thailand. In previous studies, the genus *Lactobacillus* (of the family Lactobacillaceae) was the most dominant genus in adult chicken ileum [8, 33]. Positive effects of some *Lactobacillus* species have been reported on broiler growth performance, such as for *Lactobacillus paracaseis sparacasei* and *Lactobacillus rhamnosus* [34]. Similar to our findings, the families Peptostreptococcaceae, Streptococcaceae, and Clostridiaceae have been found in the ileum content of poultry [35]. QIIME2 and Mothur use different preprocessing steps and different concepts of grouping amplicon sequences as a unit of organisms (ASV and OTU for QIIME2 and Mothur, respectively). This results in different numbers of detected units and alpha diversity. Mothur detected a larger number of microbial taxa than QIIME2. This could be because of the denoising algorithm of QIIME2 that remove low-abundance microbes as false positives [24]. Nevertheless, dominant bacteria in the broiler ileum community showed highly similar profiles (both types and proportions), suggesting that both methods could be used interchangeably for detecting overall profiles.

We also investigated the ability of microbial profile detection of the pipelines using a mock community. Both QIIME2 and Mothur provided different proportions of microbes compared to the theoretical propositions, most likely because of the efficiency of the extraction kit used for the mock community. All expected taxa in the mock community were detected; however, some other taxa could also be found, albeit at low proportions. False positives could be detected by both pipelines. In addition, the genus *Salmonella* was not reported using the QIIME2 with SILVA 138 database. The results of *Salmonella* detection for the mock dataset aligned with the broiler ileum dataset.

Pathogen detection is usually a concern in chicken production, and *Salmonella*, as a foodborne pathogen, is subject to regulations for chicken exportation worldwide [13], along with other pathogens such as *Campylobacter* and *Escherichia coli*. Recently, high-throughput sequencing was used for the detection of infectious pathogens such as *Campylobacter* in poultry carcasses, but without reports of *Salmonella* [30]. Interestingly, in our study, when using the SILVA138 database, Mothur detected both *Salmonella* and *Escherichia-Shigella*, whereas QIIME2 could annotate only the genus *Escherichia-Shigella*. However, there were unassigned ASVs of the family Enterobacteriaceae, which we further analyzed and assumed to be ASVs of *Salmonella*. This finding was obtained for both chicken ileum content and mock communities and is in agreement with previous studies [36, 37]. Our results suggest both ASV and OTU methods could classify *Salmonella* sequences; however, there are challenges at the taxonomic assignment step. In our study, the unassigned ASVs were closely related to *Salmonella* and separated from *Escherichia-Shigella*, with a high similarity with *Salmonella* species through BLAST analysis. We suggest the further investigation of ASVs in the family Enterobacteriaceae for *Salmonella* detection when using the QIIME2 and SILVA 138 database for 16S rRNA microbiome analysis. Note that the studied samples show naturally infected pathogens which were detected at low abundances similar to other studies [38, 39].

5. Conclusion

We conducted and compared the microbiome analysis approaches based on 16S rRNA amplicon sequences in the commercial broiler gut microbiome and a mock community using QIIME2 and Mothur, with the implementation of amplicon sequence variants (ASVs) and operational taxonomic units (OTUs), respectively. The two pipelines provided similar microbial profiles at all taxonomic levels. The majority of the microbes belonged to the genera *Lactobacillus* and *Romboutsia*. However, different results were found for *Salmonella* detection using SILVA138 as a reference database. By conducting Mothur, the OTUs of the genus *Salmonella* were annotated, whereas QIIME2 could only assign the identified ASVs at the family level of Enterobacteriaceae. We recommend further ASV assignment for *Salmonella* identification using BLAST. This study demonstrates that caution is required for *Salmonella* detection when based on 16S rRNA amplicon sequencing analysis.

6. Ethical approval

The animal use protocol was approved by the Kasetsart University's Institutional Animal Care and Use Committee and was in accordance with the guidelines of animal care and use under the Ethical Board of the Office of the National Research Council of Thailand (NRCT) (ID: ACKU64-VET-049 received on 9 August 2021).

7. Acknowledgments

The authors would like to acknowledge Asst. Prof. Dr. Taradon Luangtongkum, Faculty of Veterinary Medicine, Chulalongkorn University, for scientific discussions. We also acknowledge Mr. Sirawich Danwanichwong for technical support. This study was supported by Huvepharma (Thailand) Ltd.

8. References

[1] Suau A, Bonnet R, Sutren M, Godon JJ, Gibson GR, Collins MD, et al. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl Environ Microbiol. 1999;65(11):4799-4807.

[2] Zhu XY, Zhong T, Pandya Y, Joerger RD. 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. Appl Environ Microbiol. 2002;68(1):124-137.

[3] Bell RL, Jarvis KG, Ottesen AR, McFarland MA, Brown EW. Recent and emerging innovations in *Salmonella* detection: a food and environmental perspective. Microb Biotechnol. 2016;9(3):279-292.

[4] Srinivasan R, Karaoz U, Volegova M, MacKichan J, Kato-Maeda M, Miller S, et al. Use of 16S rRNA gene for identification of a broad range of clinically relevant bacterial pathogens. PLoS One. 2015;10(2):e0117617.

[5] Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590-596.

[6] Rychlik I. Composition and function of chicken gut microbiota. Animals (Basel). 2020;10(1):103.

[7] Zhou X, Jiang X, Yang C, Ma B, Lei C, Xu C, et al. Cecal microbiota of Tibetan Chickens from five geographic regions were determined by 16S rRNA sequencing. Microbiologyopen. 2016;5(5):753-762.

[8] Glendinning L, Watson KA, Watson M. Development of the duodenal, ileal, jejunal and caecal microbiota in chickens. Anim Microbiome. 2019;1(1):17.

[9] Kollarcikova M, Kubasova T, Karasova D, Crhanova M, Cejkova D, Sisak F, et al. Use of 16S rRNA gene sequencing for prediction of new opportunistic pathogens in chicken ileal and cecal microbiota. Poult Sci. 2019;98(6):2347-2353.

[10] Yadav S, Jha R. Strategies to modulate the intestinal microbiota and their effects on nutrient utilization, performance, and health of poultry. J Anim Sci Biotechnol. 2019;10:2.

[11] Clavijo V, Flórez MJV. The gastrointestinal microbiome and its association with the control of pathogens in broiler chicken production: a review. Poult Sci. 2018;97(3):1006-1021.

[12] Jeni RE, Dittoe DK, Olson EG, Lourenco J, Corcionivoschi N, Ricke SC, et al. Probiotics and potential applications for alternative poultry production systems. Poult Sci. 2021;100(7):101156.

[13] Thames HT, Sukumaran AT. A review of *Salmonella* and *Campylobacter* in broiler meat: emerging challenges and food safety measures. Foods. 2020;9(6):776.

[14] Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A. 2011;108(Suppl 1):4516-4522.

[15] Bukin YS, Galachyants YP, Morozov IV, Bukin SV, Zakharenko AS, Zemskaya TI. The effect of 16S rRNA region choice on bacterial community metabarcoding results. Sci Data. 2019;6(1):190007.

[16] Větrovský T, Baldrian P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PLoS One. 2013;8(2):e57923.

[17] Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 2017;11(12):2639-2643.

[18] Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852-857.

[19] Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;75(23):7537-7541.

[20] López-García A, Pineda-Quiroga C, Atxaerandio R, Pérez A, Hernández I, García-Rodríguez A, et al. Comparison of Mothur and QIIME for the analysis of rumen microbiota composition based on 16S rRNA amplicon sequences. Front Microbiol. 2018;9:3010.

[21] Almeida A, Mitchell AL, Tarkowska A, Finn RD. Benchmarking taxonomic assignments based on 16S rRNA gene profiling of the microbiota from commonly sampled environments. Gigascience. 2018;7(5):1-10.

[22] Andrews S. FastQC: a quality control tool for high throughput sequence data 2010 [Internet]. 2010 [cited 2022 Sep 12]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

[23] Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17(1):3.

[24] Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581-583.

[25] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547-1549.

[26] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673-4680.

[27] Stanley D, Hughes RJ, Moore RJ. Microbiota of the chicken gastrointestinal tract: influence on health, productivity and disease. Appl Microbiol Biotechnol. 2014;98(10):4301-4310.

[28] Xiao Y, Xiang Y, Zhou W, Chen J, Li K, Yang H. Microbial community mapping in intestinal tract of broiler chicken. Poult Sci. 2017;96(5):1387-1393.

[29] Bjerrum L, Engberg RM, Leser TD, Jensen BB, Finster K, Pedersen K. Microbial community composition of the ileum and cecum of broiler chickens as revealed by molecular and culture-based techniques. Poult Sci. 2006;85(7):1151-1164.

[30] Kunert-Filho HC, Furian TQ, Sesterhenn R, Chitolina GZ, Willsmann DE, Borges KA, et al. Bacterial community identification in poultry carcasses using high-throughput next generation sequencing. Int J Food Microbiol. 2022;364:109533.

[31] Suminda GGD, Bhandari S, Won Y, Goutam U, Kanth Pulicherla K, Son YO, et al. High-throughput sequencing technologies in the detection of livestock pathogens, diagnosis, and zoonotic surveillance. Comput Struct Biotechnol J. 2022;20:5378-5392.

[32] Eng SK, Pusparajah P, Ab Mutalib NS, Ser HL, Chan KG, Lee LH. *Salmonella*: A review on pathogenesis, epidemiology and antibiotic resistance. Front. Life Sci. 2015;8(3):284-293.

[33] Videnska P, Rahman MM, Faldynova M, Babak V, Matulova ME, Prukner-Radovcic E, et al. Characterization of egg laying hen and broiler fecal microbiota in poultry farms in Croatia, Czech Republic, Hungary and Slovenia. PLoS One. 2014;9(10):e110076.

[34] Fesseha H, Demlie T, Mathewos M, Eshetu E. Effect of *Lactobacillus* species probiotics on growth performance of dual-purpose chicken. Vet Med (Auckl). 2021;12:75-83.

[35] van der Hoeven-Hangoor E, van der Vossen JM, Schuren FH, Verstegen MW, de Oliveira JE, Montijn RC, et al. Ileal microbiota composition of broilers fed various commercial diet compositions. Poult Sci. 2013;92(10):2713-2723.

[36] Clavijo V, Morales T, Vives-Flores MJ, Reyes Muñoz A. The gut microbiota of chickens in a commercial farm treated with a *Salmonella* phage cocktail. Sci Rep. 2022;12(1):991.

[37] Wilson A, Chandry PS, Turner MS, Courtice JM, Fegan N. Comparison between cage and free-range egg production on microbial composition, diversity and the presence of *Salmonella enterica*. Food Microbiol. 2021;97:103754.

[38] Huang P, Zhang Y, Xiao K, Jiang F, Wang H, Tang D, et al. The chicken gut metagenome and the modulatory effects of plant-derived benzylisoquinoline alkaloids. Microbiome. 2018;6(1):211.

[39] Pedroso AA, Lee MD, Maurer JJ. Strength lies in diversity: how community diversity limits *Salmonella* abundance in the chicken intestine. Front Microbiol. 2021;12:694215.