Microbial diversity of a full-scale UASB reactor applied to poultry slaughterhouse wastewater treatment: integration of 16S rRNA gene amplicon and shotgun metagenomic sequencing

Tiago Palladino Delforno1 | Gileno Vieira Lacerda Júnior1 | Melline F. Noronha1 | Isabel K. Sakamoto2 | Maria Bernadete A. Varesche2 | Valéria M. Oliveira1

1Microbial Resources Division, Research Center for Chemistry, Biology and Agriculture (CPQBA), Campinas University - UNICAMP, Campinas, São Paulo, Brazil
2Laboratory of Biological Processes, Department of Hydraulics and Sanitation, Engineering School of São Carlos - University of São Paulo (EESC - USP) Campus II, São Carlos, São Paulo, Brazil

Correspondence
Tiago Palladino Delforno, Microbial Resources Division, Research Center for Chemistry, Biology and Agriculture (CPQBA), Campinas University - UNICAMP, Campinas, São Paulo, Brazil.
Email: tiago.palladino@gmail.com

Funding information
Fundação de Amparo à Pesquisa do Estado de São Paulo. Grant/Award Number: 2014/16426-0 and 2013/09386-9.

Abstract
The 16S rRNA gene amplicon and whole-genome shotgun metagenomic (WGSM) sequencing approaches were used to investigate wide-spectrum profiles of microbial composition and metabolic diversity from a full-scale UASB reactor applied to poultry slaughterhouse wastewater treatment. The data were generated by using MiSeq 2 × 250 bp and HiSeq 2 × 150 bp Illumina sequencing platforms for 16S amplicon and WGSM sequencing, respectively. Each approach revealed a distinct microbial community profile, with *Pseudomonas* and *Psychrobacter* as predominant genus for the WGSM dataset and *Clostridium* and *Methanosaeta* for the 16S rRNA gene amplicon dataset. The virome characterization revealed the presence of two viral families with *Bacteriidae* and *Archaeidae* as host, *Myxococcidae*, and *Siphoviridae*. A wide functional diversity was found with predominance of genes involved in the metabolism of acetone, butanol, and ethanol synthesis; and one-carbon metabolism (e.g., methanogenesis). Genes related to the acetotrophic methanogenesis pathways were more abundant than methylotrophic and hydrogenotrophic, corroborating the taxonomic results that showed the prevalence of the acetotrophic genus *Methanosaeta*. Moreover, the dataset indicated a variety of metabolic genes involved in sulfur, nitrogen, iron, and phosphorus cycles, with many genera able to act in all cycles. BLAST analysis against Antibiotic Resistance Genes Database (ARDB) revealed that microbial community contained 43 different types of antibiotic resistance genes, some of them were associated with growth chicken promotion (e.g., bacitracin, tetracycline, and polymyxin).

Keywords
anaerobic microbial community, antibiotics resistance genes, genetic potential, granular sludge, virome
1 | INTRODUCTION

The anaerobic digestion (AD) corresponds to a consolidated technology that is successfully applied in different wastes and influents for the conversion of complex organic compounds to CH\textsubscript{4} and CO\textsubscript{2}, nutrient recovery (phosphorus and sulfur), pollutant removal and/or energy (CH\textsubscript{4} and H\textsubscript{2}), and fermentative compound (butyrate, butanol) production (Batstone & Virdis, 2014; Battistoni et al., 1997).

The large AD application is only possible due to the occurrence of complex interactions of microorganisms belonging to Bacteria and Archaea domains (Zinder, 1984). Previous studies have reported that microbial composition of anaerobic digestion sludge is highly diverse (Cabezas et al., 2015). However, information on the key players in anaerobic conditions and their interactions is not comprehensive. These results indicate a high genetic potential that has been poorly explored and open promising perspectives for biotechnological applications.

WGSM sequencing has been shown as an excellent approach to assess the genetic potential of a microbial community, unraveling the phylogenetic composition, metabolic capacity, and functional diversity (Streit & Schmitz, 2004) of community members. However, comprehensive information on the phylogenetic composition is highly dependent on the sequence number and in general the shotgun metagenomic sequencing is not deep enough. On the other hand, WGSM sequencing does not involve the biased amplification of 16S rRNA genes, and although more precise, the organism abundances are still dependent on the DNA extraction and sequencing protocols used (Kalyuzhnaya et al., 2008; Morgan, Darling, & Eisen, 2010).

Although the 16S rRNA gene amplicon sequencing has showed several artifacts due to PCR amplification and sequencing errors, this approach offers a deep sequencing for a large microbial community characterization allowing to detect rare species in complex communities (Quince et al., 2009). For this reason, the integration of 16S rRNA gene amplicon and WGSM sequencing approaches represent a robust strategy for the study of microbiomes. It is important to emphasize that the microbiological knowledge (phylogenetic composition, metabolic capacity, and functional diversity) in AD is fundamental to elect microbial indicators of optimal performance, to understand the microbial behavior in response to environmental disturbance and to optimize the process (Carballa, Regueiro, & Lema, 2015).

The term metagenomics was first used by Handelsman et al. (1998) in a soil microbiome study and the first report associating metagenomics with AD occurred only 10 years later in a deep microbial analysis of a German full-scale biogas plant treating farm waste (Schlüter et al., 2008). Thenceforward, several studies have examined the AD under a taxonomic and functional perspective, allowing the reconstruction of important metabolic pathways and genomes (Campanaro et al., 2016; Wang et al., 2013; Wirth et al., 2012; Wong et al., 2013) with focus on the understanding of biological interactions and optimization of bioprocesses. Despite the fact that the first report on the application of metagenomics in AD has been from real-scale reactors, most of the studies have focused on the evaluation of laboratory-scale reactors. Therefore, scarce investigation has been performed in full-scale reactors treating real wastewater and these reactors should not be neglected, especially those applied to wastewater treating of industrial activities of global impact like poultry production.

Brazil represents one of main poultry exporters worldwide, with more than 4.2 millions tons per year (Facta 2016). This intense poultry production results in wastewater with high polluting potential due mainly to the process and the washing of equipment with high chemical oxygen demand (COD; 1.790–4.760 mg L\textsuperscript{-1}), oil and grease (114–640 mg L\textsuperscript{-1}), nitrogen (90–196 mg N L\textsuperscript{-1}), and phosphorus (22–84 mg L\textsuperscript{-1}) (de Nardi, Fuzi, & Del Nery, 2008; Del Nery et al., 2013). The AD technology, especially the up-flow anaerobic sludge blanket (UASB) reactor, has been applied as the main alternative for poultry slaughterhouse wastewater treatment. Apart from the engineering aspects related to the optimization of UASB reactors for treatment of this type of wastewater, several studies have investigated the microbiological potential of the microbiome in granular sludge from UASB reactors, and the results are promising. The high microbial diversity observed (Delforno et al., 2015) represents a high genetic diversity and, consequently, a broad metabolic potential for several bioprocesses. Therefore, a wide application of granular sludge for nitrogen (Moura, Damianovic, & Foresti, 2012) and sulfide removal (Camiloti et al., 2014), hydrogen production (Penteado et al., 2013), anionic surfactant degradation (Okada et al., 2014), polychlorinated biphenyl (PCB) degradation (Gomes et al., 2014), and veterinary antimicrobial removal (Oliveira, Santos-Neto, & Zaiat, 2015) has been observed.

In this context, the objective of this study was to carry out a robust description of the microbiome in a full-scale UASB reactor applied to poultry slaughterhouse wastewater treatment. For the in-depth microbiome survey, an integrated approach combining 16S rRNA gene amplicon and WGSM sequencing was used. In addition to the description of the community composition (Bacteria, Archaea, and Virus), the metabolic processes related to biogeochemical cycles (e.g., sulfur, iron, phosphorus, nitrogen) and antibiotic resistance genes were explored (microbial metabolic potential). In this sense, the robust information on the microbiome composition, functioning, and interactions may guide novel biotechnological applications of the granular sludge present in the UASB reactor under study.

2 | MATERIAL AND METHODS

2.1 | Wastewater treatment system of the poultry slaughterhouse

The wastewater treatment system of poultry slaughterhouse is located in Tiete, Sao Paulo State (Brazil) and consists of four parts: (1) rotary screens to remove feathers and bowels; (2) static screens to remove the fine solids; (3) dissolved-air flotation system (DAF); and (4) full-scale UASB reactor. The circular UASB reactor had a flow rate of 975 m\textsuperscript{3}d\textsuperscript{-1}, an area of 170 m\textsuperscript{2} and operated under mesophilic conditions. The hydraulic retention time (HRT) was 1.02 days. The average influent COD of reactorwas 2,000 mg L\textsuperscript{-1} with 77–88% of removal. Average grease and fats were 308 and 469 mg L\textsuperscript{-1} for effluent and influent, respectively. The total solids (TSS) were 2,133 mg L\textsuperscript{-1} in
the effluent and 1,593 in the influent, whereas, the concentration of N-NH₃ and P-PO₄ in the influent reached 48 mg L⁻¹ and 38 mg L⁻¹, respectively.

2.2 | Sampling, DNA extraction, Sequencing, and Preprocessing data

Samples examined in this study were collected from UASB reactor in December 2013. Initially, granular sludge was collected in 20 l high-density polyurethane bottles from UASB reactor. In the laboratory, 25 ml of the granular sludge was transferred to 50 ml Falcon tubes and stored at −20°C until analyzed. The surplus, present in the high-density polyurethane bottles, was stored at 4°C to be used as an inoculum for other anaerobic bioreactor studies.

For the in-depth microbiome characterization, two strategies were adopted: (i) the sequencing of 16S rRNA gene amplicons for phylogenetic characterization and (ii) metagenomic sequencing for functional diversity characterization.

The sample for phylogenetic characterization was sequenced using the MiSeq 2 x 250 bp and DNA was extracted using a modified phenol–chloroform protocol described by Griffiths et al. (2000). The 16S rRNA genes were amplified using the primer set S-D-Bact-0341-b-S-17 (5′-CCTACGGGNGGCWGCAG-3′) and S-D-Bact-0785-a-A-21 (5′-GACTACHVGGGTATCTAATCC-3′), flanking the V3 and V4 hypervariable regions, as described by Klindworth et al. (2013). 16S amplicon sequencing was performed at the Animal Biotechnology Lab, Department of Animal Science (ESALQ/USP Piracicaba, Brazil) on a MiSeq platform following the manufacturer’s guidelines. On the other hand, whole metagenome was sequenced using the HiSeq 2 x 150 bp and DNA was extracted using the PowerSoil® DNA Isolation kit (MoBio Laboratories, Inc., Carlsbad, CA, USA). WGSM sequencing was performed at MR DNA (www.mrdnalab.com, Shallowater, TX, USA) on a HiSeq platform following the manufacturer’s guidelines.

The differences between the platforms used for each sample include the desired fragment size, quantity of data generated, and cost of sequencing. For the phylogenetic diversity, fragment size is highly important to increase the reliability of the evolutionary relatedness among sequences used as OTUs, whereas for the evaluation of the functional diversity the amount of generated data is important to increase the coverage. In both cases, the DNA quality was assessed by an ND-2000 spectrophotometer (Nanodrop Inc., Wilmington, DE), using a ratio of 260/280 nm > 1.8, and agarose gel electrophoresis.

All datasets in FastQ format were uploaded to MG-RAST server (http://metagenomics.nmpdr.org/) version 3.5 (Meyer et al., 2008) for preprocessing and annotation. The artificial replicate sequences were removed according to Gomez-Alvarez, Teal, and Schmidt (2009), whereas the low-quality sequences were removed using the modified Dynamic Trim (Cox, Peterson, & Biggs, 2010); the default parameters were adopted.

The sequence data are available at MG-RAST server access number 4633385.3(Amplicon_PS) and 4626733.3 (WGS_whole).

2.3 | Microbial community structure and composition (phylogenetic characterization)

The microbial community structure and composition of the UASB reactor sludge was evaluated using two approaches: (1) the 16S rRNA gene sequences derived from the amplicon sequencing were used against the M5RNA (Nonredundant multisource ribosomal RNA annotation) database (Meyer et al., 2008); (2) the 16S, 23S, ITS, and 18S rRNA sequences derived from the metagenome sequencing were extracted and analyzed against the M5RNA database (WGS_rRNA; Figure 1). The parameters adopted in both analyses were as follows: max. E-value cutoff = 1e⁻⁵, min. % identity cutoff = 60% and min. alignment length cutoff = 15. Moreover, taxonomic assignment of metagenomic sequence (WGS_whole) was performed against the SEED database with default parameters (Overbeek et al., 2005). In both cases (using the M5RNA and SEED database), the best hit classification was used to visualize the results.

2.3.1 | Diversity analyses

Coverage and alpha diversity indices (dominance, Simpson, Shannon, and Chao-1) were calculated using the PAST software version 3.07.

---

**FIGURE 1** Pipeline of methods and analyses. The sequence data are available at MG-RAST server under the access number 4633385.3 (Amplicon_PS) and 4626733.3 (WGS_whole).
(Hammer, Harper, & Ryan, 2001). Thus, the .biom files of each sample generated by the MG-RAST (QIIME report; Caporaso et al., 2010) were the input files for the PAST software.

2.4 Functional analysis

Functional classification of WGS_whole-sequence dataset was conducted by MG-RAST annotation pipeline using the SEED subsystems database. The data were compared using a maximum e-value of 1e⁻⁵, a minimum identity of 60%, and a minimum alignment length of 15 amino acids for proteins.

2.4.1 Metabolic mapping

Some specific metabolic mappings, such as carbon, nitrogen, phosphorus, sulfur, iron, protein, and aromatic compounds, were investigated in details. Sequences assigned to each of the metabolic mappings mentioned above were extracted from SEED dataset and affiliated to the taxonomic groups harboring the most related function (protein) to provide a picture of the microbial community likely involved in such specific metabolic pathways. The taxonomic assignment was performed using the SEED database with default parameters and the data were visualized using the best hit classification for each feature.

Moreover, putative antibiotic resistance genes (ARGs) were searched in the WGS_whole dataset against the Antibiotic Resistance Database (ARDB), which include nonredundant genes, using BLASTX at a cutoff E value ≤10⁻⁵. A read was identified as an ARG-like sequence according to its best BLASTX hit with amino acid identity of ≥90% and alignment length of ≥25 amino acids (Kristiansson et al., 2011; Yang et al., 2013).

3 RESULTS

3.1 Sequencing statistics

In total, 52,826,969 and 293,825 sequences, with average lengths of 233 ± 78 bp and 454 ± 14 bp, were obtained for WGSAM sequencing and 16S amplicon sequencing, respectively. After trimming, 66.4% and 34.0% of sequences were removed due to the low quality and the average lengths were reduced to 213 ± 91 bp and 336 ± 156 bp, for WGSAM sequencing and 16S amplicon sequencing, respectively. The values of GC-content were around 53–54% for both samples.

3.2 Taxonomic characterization of UASB reactor microbiome

The number of rRNA genes extracted from WGSAM dataset was 211,795 sequences (1.0% of total sequences; Table 1). On the other hand, for the 16S amplicon dataset, 191,804 sequences were used for taxonomic assignment.

The total number of OTUs (97% similarity cutoff) and singletons obtained were 845–880 and 246–203, for WGS_rRNA and 16S amplicon datasets, respectively. The highest number of OTUs and the lowest number of singletons were found in the 16S amplicon dataset. The values of Shannon_H and Chao-1 ranged from 3.32 to 3.67 and 1,204 to 1,124, for WGS_rRNA and 16S amplicon datasets, respectively. Slight differences were observed in the microbial structure of the UASB reactor when comparing the WGS_whole_SEED and WGS_rRNA_M5RNA datasets. For the domain level, the relative abundance obtained were 89.6–96.6% for Bacteria and 1.0–0.2% for Archaea, using the SEED and M5RNA databases. On the other hand, the microbial composition from 16S amplicon dataset showed 84.0% for Bacteria and 6.5% for Archaea domains (data not shown).

At the phylum level, 36, 29, and 37 distinct phyla were observed for Amplicon_PS_M5RNA, WGS_rRNA_M5RNA and WGS_whole_SEED, respectively. Proteobacteria and Firmicutes were the most representative phyla in WGS_rRNA M5RNA and WGS_whole_SEED datasets (Figure 2); whereas for the Amplicon_PS_M5RNA dataset, Bacteroidetes was the most abundant phylum, followed by Proteobacteria and Firmicutes. The Euryarchaeota phylum was the only representative (0.25–8.7% relative abundance) for the Archaea domain in all datasets.

At the genus level, a high abundance of Pseudomonas (55.8–56.2%), Psychrobacter (14.5–22.6%), Sporosarcina (3.2%), Clostridium (1.8–10.6%), and Bacillus (1.3–1.8%) were observed in WGS_rRNA M5RNA and WGS_whole_SEED datasets (Figure 2). On the other hand, many different genera with similar abundances were observed in the Amplicon_PS_M5RNA dataset, as follows: Clostridium (10.5%), Methanosaeta (9.5%), Tannerella (8.0%), Ornithobacterium (7.8%), Candidatus Cloacimonas (6.6%), Porphyromonas (3.7%), Deinococcus (3.0%), and Thermincola (3.0%).

The proportional Venn diagram showed taxonomic overlap percentages from 24% to genus level comparing WGS_whole_SEED and WGS_rRNA_M5RNA datasets. The comparison between datasets containing only rRNA sequences (WGS_rRNA_M5RNA and Amplicon_PS_M5RNA) revealed taxonomic overlap percentages ranging from 29% to 57%, with the lowest overlap value observed at the genus level (Figure 3).

| TABLE 1 | Diversity analyses indices and estimations. Amplicon_PS → sequencing of 16S rRNA gene amplicons. WGS_rRNA → rRNA (16S, 23S, ITS, and 18S) sequence extracted from metagenomic sequencing |
|---------------------------------|---------------------------------|
| | Numbers of sequences | Diversity index |
| | | Shannon_H | Chao-1 |
| | | 3.67 ± 0.01 | 1.124 ± 55 |
| OTU number (observed richness) | 880 | 1,204 ± 62 |
| Singletons | 203 | 246 |
| Good’s Coverage | 99.93% | 99.88% |

The values of Shannon_H and Chao-1 ranged from 3.32 to 3.67 and 1,204 to 1,124, for WGS_rRNA and 16S amplicon datasets, respectively. Slight differences were observed in the microbial structure of the UASB reactor when comparing the WGS_whole_SEED and WGS_rRNA_M5RNA datasets. For the domain level, the relative abundance obtained were 89.6–96.6% for Bacteria and 1.0–0.2% for Archaea, using the SEED and M5RNA databases. On the other hand, the microbial composition from 16S amplicon dataset showed 84.0% for Bacteria and 6.5% for Archaea domains (data not shown).

At the phylum level, 36, 29, and 37 distinct phyla were observed for Amplicon_PS_M5RNA, WGS_rRNA_M5RNA and WGS_whole_SEED, respectively. Proteobacteria and Firmicutes were the most representative phyla in WGS_rRNA M5RNA and WGS_whole_SEED datasets (Figure 2); whereas for the Amplicon_PS_M5RNA dataset, Bacteroidetes was the most abundant phylum, followed by Proteobacteria and Firmicutes. The Euryarchaeota phylum was the only representative (0.25–8.7% relative abundance) for the Archaea domain in all datasets.

At the genus level, a high abundance of Pseudomonas (55.8–56.2%), Psychrobacter (14.5–22.6%), Sporosarcina (3.2%), Clostridium (1.8–10.6%), and Bacillus (1.3–1.8%) were observed in WGS_rRNA M5RNA and WGS_whole_SEED datasets (Figure 2). On the other hand, many different genera with similar abundances were observed in the Amplicon_PS_M5RNA dataset, as follows: Clostridium (10.5%), Methanosaeta (9.5%), Tannerella (8.0%), Ornithobacterium (7.8%), Candidatus Cloacimonas (6.6%), Porphyromonas (3.7%), Deinococcus (3.0%), and Thermincola (3.0%).

The proportional Venn diagram showed taxonomic overlap percentages from 24% to genus level comparing WGS_whole_SEED and WGS_rRNA_M5RNA datasets. The comparison between datasets containing only rRNA sequences (WGS_rRNA_M5RNA and Amplicon_PS_M5RNA) revealed taxonomic overlap percentages ranging from 29% to 57%, with the lowest overlap value observed at the genus level (Figure 3).
FIGURE 2  Taxonomic affiliation of reads. WGS_rRNA_M5RNA = rRNA gene sequences against the M5RNA database. Amplicon_PS_M5RNA = 16S rRNA gene amplicons against the M5RNA database. WGS_whole_SEED = all sequences against the SEED database.
The virome was analyzed from the metagenomic datasets (WGS_whole) using the SEED database (Table 2). Only 0.13% of the total sequences were related to viral genotypes, and the most abundant taxonomies were Microviridae (67.4%), Myoviridae (20.7%), Siphoviridae (9.2%), and Podoviridae (0.5%). Among these, the families Microviridae and Podoviridae have the Bacteria domain as the unique host accounting over 68% of relative abundance; whereas only 30% may use microorganism belong the Archaea domain as host.

### 3.3 Functional profile of UASB reactor microbiome

The main Level 1 subsystems were clustering-based subsystems (13.5%), carbohydrates (12.3%) and amino acids and derivatives (11.3%; Table S1). Inside the carbohydrates category, the subsystems related to central carbohydrate metabolism (3.6%), such as dehydrogenase complexes, glycolysis and gluconeogenesis and pyruvate metabolism II: acetogenesis from pyruvate were the most represented in the dataset. Another important function inside the carbohydrates category was fermentation (1.8%), for example: acetyl-CoA fermentation to butyrate, acetone butanol ethanol synthesis, and butanol biosynthesis (Figure 4a–c and Table S2).

In the category of one-carbon metabolism (Level 2 SEED), the relative abundance reached 1.3% (Figure 4A; Table S2). The most abundant metabolic type in this category were serine-glyoxylate cycle (85.2%), followed by one-carbon metabolism by tetrahydropterines (10.2%), methanogenesis (3.3%), and methanogenesis from methylated compounds (0.5%; Figure 4d and Table S2). Moreover, functional enzyme-encoding genes for the methanogenesis pathways were identified and annotated based on the database extracted from KEGG modules. Based on the results, the most prevalent methanogenesis pathway was acetoclastic (43.7%) followed by methylo trophic (34.7%) and hydrogenotrophic (21.6%; Table S3).

In the metabolism of aromatic compounds, 0.14% of reads were related to the anaerobic degradation of aromatic compounds with emphasis on anaerobic benzoate metabolism (99.0%; Figure 4a–b), specifically the function Acetyl-CoA acetyltransferase.

The cell wall and capsule (4.4%) and quorum sensing and biofilm formation (0.12%) categories, that are likely involved in the formation and maintenance of flocs and biofilms (e.g., granular sludge present in UASB reactors), were found. The main functions in these categories were gram-negative cell wall components (1.4%), capsular and extracellular polysaccharides (1.2%) and biofilm adhesion biosynthesis (0.06%).

Proteins assigned to iron, nitrogen, sulfur, phosphorus, and potassium metabolisms were found with relative abundance ranging from 0.5% to 2.3% (Tables S1 and S2). In the sulfur metabolism, sequences were assigned to organic and inorganic sulfur assimilation, alkanesulfonate assimilation, sulfur oxidation, and sulfate reduction-associated complexes; whereas in the nitrogen metabolism, sequences were mainly related to ammonia assimilation, nitrate and nitrite ammonification, and denitrification.

Taxonomic inferences of the genes involved in biogeochemical cycles (sulfur, nitrogen, iron, phosphorus, and potassium) indicated that 248 different bacterial genera showed enzymatic machinery for all biogeochemical cycles (Figure 5a and b). For the domain Archaea, most of the genera (17) are involved in all biogeochemical cycles.

---

**Table 2** Virus families from WGS_whole_SEED dataset annotated by SEED database through MG-RAST server

| Viral Family    | Number of Sequence | Relative Abundance | host   |
|-----------------|--------------------|--------------------|--------|
| Baculoviridae   | 159                | 1.0%               | Eukarya|
| Microviridae    | 11,224             | 67.4%              | Bacteria|
| Mimiviridae     | 58                 | 0.3%               | Eukarya|
| Myoviridae      | 3,447              | 20.7%              | Bacteria, Archaea|
| Phycodnaviridae | 19                 | 0.1%               | Eukarya|
| Siphoviridae    | 1,535              | 9.2%               | Bacteria, Archaea|
| Podoviridae     | 88                 | 0.5%               | Bacteria|
| Others          | 112                | 0.7%               | –      |
However, only five genera (Metallosphaera; Thermofilum; Thermoproteus; Ferroplasma; and Picrophilus) showed enzymatic machinery for the sulfur cycle.

**3.4 Antibiotic Resistance Genes and their host organisms in the UASB reactor microbiome**

In total, 43 different ARGs were found in the UASB reactor microbiome, representing 0.03% of total metagenomics sequences (Figure 6 and Table S4). The top five dominant ARGs types corresponded to multidrug, bacitracin, tetracycline, lincomycin, and polymyxin genes, accounting for over 95% of the total ARGs sequences. The last four ARGs cited are related to nontherapeutic antibiotics, and it can be used as chicken growth promoters. On the other hand, vancomycin resistance (vanRG) genes, an important therapeutic antibiotic, were found accounting for 0.03% of all ARGs determined. The taxonomic assignment of ARG sequences showed high predominance of the phylum Proteobacteria (91.7%), followed by Firmicutes (3.5%) and...
Bacteroidetes (0.6%); whereas the main genera were Pseudomonas (75.7%), Burkholderia (7.1%), Psychrobacter (3.9%), and Moraxella (3.4%; Table 3 and Table S4).

4 | DISCUSSION

Each approach (16S rRNA gene amplicon and WGSM sequencing) revealed a distinct microbial community profile. Probably, these differences observed are a result of the combination of biases introduced during PCR amplification of 16S rRNA gene, different DNA extraction procedures and different high-throughput sequencing platforms. Differences in microbial composition comparing 16S rRNA amplicon and WGSM sequencing approaches were also obtained by Shah et al. (2010). The authors carried out a systematic comparison using 33 metagenomic datasets of human-associated bacterial communities and 71 datasets derived from 16S rRNA amplicon sequencing and observed significant differences between the community structures due to the distinct experimental approaches used, suggesting that sample replicates and dataset sizes had smaller influence in the microbial profile than the DNA extraction methods or sequencing protocols. Additionally, De bautista-los Santos et al. (2016) using high throughput sequencing of 16S rRNA amplicons for microbial characterization of drink water observed that bias and variability inherent to the PCR amplification and sequencing steps is significant enough to mask differences between bacterial communities from replicate samples. In addition, the average size of the reads obtained in this study (454 bp for 16S rRNA amplicon sequencing and 233 bp for WGSM sequencing) and the hypervariable regions analyzed differed between both strategies, potentially leading to different accuracy levels in the phylogenetic affiliation and distinct microbial community profiles (Wang et al., 2007). Poretsky et al. (2014), studying the planktonic microbial community, observed that at phylum level the taxonomic composition in the 16S rRNA library was similar to its corresponding metagenome. On the other hand, at the genus level a large amount of variation between the libraries was observed. This is probably due to the fact that many genera with distinct metabolisms are usually encompassed by the same phylum, and thus important information about variation in particular community members among different communities remains inaccessible. However, it is interesting to emphasize that with the combined use of different tools, more taxonomies could be observed.

The Venn diagram showed the proportion of shared and exclusive taxa found when using the different approaches for phylogenetic inference. Overlap values were similar when comparing datasets derived from the whole-genome shotgun metagenomic sequencing (WGS_whole_SEQUENCE and WGS_rRNA_M5RNA) and datasets containing only rRNA sequences (WGS_rRNA_M5RNA and Amplicon_PS_M5RNA).

![FIGURE 6](image_url) Distribution of reads among the classes of antibiotic resistance genes (ARGs) in the WGS_whole dataset annotated by Antibiotic Resistance Database (ARDB)

**TABLE 3** Taxonomic classification of sequences related to antibiotics resistance genes (ARGs) from WGS_whole dataset using the SEED database through MG-RAST server.

| Taxonomy               | Relative Abundance |
|------------------------|--------------------|
| Bacteroidetes          | 0.61%              |
| Bacteroides            | 0.37%              |
| Prevotella             | 0.24%              |
| Firmicutes             | 3.49%              |
| Bacillus               | 0.10%              |
| Clostridium            | 0.48%              |
| Enterococcus           | 0.91%              |
| Lactobacillus          | 0.23%              |
| Staphylococcus         | 0.46%              |
| Streptococcus          | 1.31%              |
| Proteobacteria         | 91.75%             |
| Acinetobacter          | 0.57%              |
| Burkholderia           | 7.12%              |
| Campylobacter          | 0.13%              |
| Escherichia            | 0.17%              |
| Klebsiella             | 0.10%              |
| Methyllobacillus       | 0.63%              |
| Moraxella              | 3.44%              |
| Pseudomonas            | 75.66%             |
| Psychrobacter          | 3.93%              |
| Others                 | 4.16%              |
The study of viral community in anaerobic wastewater treatment is scarce. However, the viral community may play an important role in methanogenic digestion, since the Archaea (e.g., Methanoseta) may be a favorable target to viral attack (Chien et al., 2013). Therefore, viral infection might explain the decrease in methane production, in upsets and process failures with no obvious explanation (Kroeker et al., 1979). Tamaki et al. (2012) evaluated the viral community in the different stages of wastewater treatment plants, named activated sludge and anaerobic digestion. In the activated sludge the predominant viral families were Podoviridae (45.2%), Siphoviridae (26.5%), and Myoviridae (22.1%). On the other hand, in the anaerobic digestion the order of predominance changed, as follows: Siphoviridae (42.4%), Myoviridae (36.0%), and Podoviridae (16.2%). It is worth to highlight that Siphoviridae and Myoviridae (high presence in the anaerobic digestion where methanogens are abundant) can have Archaea as host, whereas Podoviridae (high presence in the activated sludge) can use only Bacteria as host. In this study, similar results were obtained, with the predominance of Myoviridae (20.7%) and Siphoviridae (9.2%), the exception was the Microviridae family, which was not detected by Tamaki et al. (2012).

The functional profile observed for the UASB reactor microbiome was similar to the one obtained by Lv et al. (2015) when analyzing global functions in the anaerobic-anoxic reactor applied to denitrifying phosphorus removal. The main functions inside the carbohydrates category and clustering-based subsystems indicate the microbial genetic potential for fermentative processes. In the category of one-carbon metabolism, similar values (~1.77% and 2.33%) were observed by Yang et al. (2013), when analyzing the sludge metagenome from full-scale anaerobic digesters operated in municipal wastewater treatment plants, and by Guo et al. (2015), for the metagenome of anaerobic digester applied to the treatment of activated sludge waste and the production of methane for biofuel.

The abundance of sequences related to the methanogenesis process reached ~3.8% in the UASB reactor microbiome, whereas Guo et al. (2015) observed 9.0%. It is worth mentioning that the main goal of the wastewater treatment system under study is the treatment of organic load from poultry slaughterhouses and not the production of methane for biofuel. The current knowledge recognizes three pathways for methanogenesis named acetoclastic (acetate is reduced to CH₄), hydrogenotrophic (CO₂ is reduced to CH₄), and methylotrophic (methyl group is reduced to CH₄). Liu & Whitman, 2008. The prevalence of acetotrophic methanogenic pathways was also consistent with the taxonomic results, which indicated Methanoseta (~9.5% of relative abundance) as the most abundant methanogenic group from the Amplicon_M5RNA dataset. On the other hand, the hydrogenotrophic (Methanoculleus, Methanoliniea, and Methanococcus) and methylotrophic (Methanoscina, Methanomethylovorans) methanogenic genera showed 0.24% and 0.15% of relative abundance from the Amplicon_M5RNA dataset, respectively. The prevalence of acetoclastic pathway for methane production in AD has also been observed by Guo et al. (2015) and Yang et al. (2013).

The abundance of reads related to the metabolism of aromatic compounds and resistance to antibiotics and toxic compounds indicates the genetic potential of the microbial community to survive in environments with fluctuating levels of contaminants and stress, in addition to their potential for the application in bioremediation processes. Sanapareddy et al. (2009), evaluating industrial and medical wastewaters, observed 5.0% of aromatic compound metabolism compared to 1.8% observed in this study. These results suggest that the type of wastewater treatment employed may be a key factor in functional gene distribution.

The highest abundance observed was related to iron metabolism, probably due to the selection exerted by the iron present in the poultry blood. Differently, Guo et al. (2015) observed only 0.45% from full-scale anaerobic reactor treating the waste of activated sludge, suggesting that the presence of iron in poultry blood has probably favored this type of metabolism. The sequences assigned to the nitrogen and sulfur metabolism suggest the potential of the biomass from UASB reactor treating poultry slaughterhouse to be used as inoculum for different processes, such as: fermentative production, aromatic compound degradation, and sulfur, iron, nitrogen, and phosphorus removal. Moreover, the presence of proteins assigned to the resistance of cobalt-zinc-cadmium and arsenic might be biotechnologically explored. Except for the nitrogen metabolism, the relative abundance of reads related to sulfur, phosphorus, and potassium metabolisms were slightly higher than that obtained by Guo et al. (2015) and Yang et al. (2013) when analyzing the metagenome from full-scale anaerobic digesters.

Taxonomic inference of genes involved in biogeochemical cycles (sulfur, nitrogen, iron, phosphorus, and potassium) suggested that many microbial genera have enzymatic machinery for all biogeochemical cycles. According to Zumstein, Moletta, and Godon (2000), many microbes play the same role in the same biogeochemical cycles, strategy known as functional redundancy that makes the anaerobic digestion process robust and plastic. This strategy is considered as an insurance to maintain ecosystem functions under changing environmental conditions (McMahon, Martin, & Hugenholtz, 2007). The finding of a high number of microbial genera harboring genes involved in all biogeochemical cycles in the UASB reactor microbiome reinforces the importance of functional redundancy within the anaerobic digestion.

Similar conclusion was obtained by Langer et al. (2015) describing the functional redundancy and structural changes of microbial communities involved in laboratory-scale continuously stirred tank reactors treating maize silage in cofermentation with sugar beet silage. Although the authors observed structural changes according to the different mixtures of maize and sugar beet silage, similar biogas production rates were obtained for equal organic loading rates. Vanwonerghem et al. (2014), operating mesophilic reactors with HRT of 10 days and supplied with a model α-cellulose, observed similar results for the main steps in AD (hydrolysis, fermentation, and methanogenesis), with multiple phylogenetically diverse populations associated with each step.

Although ARGs have been described in chicken feces (Li et al., 2015), studies of such resistance genes in full anaerobic systems applied to the treatment of poultry slaughterhouse wastewater are scarce. In this study, 0.03% of total metagenomic sequences were
related to antibiotic resistance genes. Li et al. (2015) observed higher numbers of ARG sequences in wastewater from livestock farms (0.5%–3.0%). However, ARGs frequency found in this study is similar to the one in drinking water (0.01%–0.05%), soil (0.02%), sediment (0.004–0.03%), and river water (0.02–0.03%; Li et al., 2015).

The frequencies of ARGs are likely associated with the antibiotics used as veterinary medicine, including growth promoters. Previous authors have described that the use of nontherapeutic antibiotics as growth promoters enhanced the abundance of ARGs due to the selective pressure on the microbial community. Looft et al. (2012) observed an increase in ARGs abundance in the swine microbiome after 14 days under antibiotic feeding. Similar results were obtained by Li et al. (2015) when comparing feces from commercially grown chicks with those from 20-day- to 80-day-old ones. Thus, the presence of different ARGs types in the poultry slaughterhouse wastewater treatment is probably a consequence of the antibiotics added into the chicken food/water.

The presence of vancomycin ARGs deserves attention since it represents the last line of defense against some bacterial strains resistant to most antibiotics (e.g., Streptococcus, Enterococcus). Other authors have found vancomycin resistance genes in swine manure and human feces samples, but in low abundance (0.016%–0.18%). The presence of vancomycin resistance gene it might represent the uncontrolled use of the antibiotic and/or spread of the resistance gene by mobile elements along the microbial community (coselection and co-occurrence).

In summary, the integration of 16S rRNA gene and metagenome sequencing approaches allowed the in-depth access of the UASB reactor microbiome, revealing different microbial community profiles and reinforcing the need for integrating multiple tools for microbial community investigation. Analysis of the metagenomic dataset showed the prevalence of Pseudomonas, Psychrobacter, and Sporosarcina members. On the other hand, the 16S rRNA sequencing approach revealed a higher global diversity and did not show the prevalence of any specific genus. The viral community was mainly composed of the families Microviridae, Myoviridae, and Siphoviridae, with Bacteria and Archaea members as hosts. A wide functional diversity was found in the UASB reactor microbiome, representing a high genetic potential to be further explored in biotechnology such as anaerobic degradation of aromatic compounds, fermentation, protein degradation, and one-carbon metabolism. Regarding the methanogenic metabolism, genes related to the acetotrophic pathways were more abundant, corroborating the taxonomic results that showed the prevalence of the acetotrophic genus Methanoseta. Moreover, the data indicated the genetic potential of the inoculum for other bioprocesses, such as the removal of sulfur and nitrogen, with many genera able to play a role in all cycles. Different types of antibiotic resistances genes (ARGs) were found, which are likely associated with the antibiotics used as growth chicken promoters and suggesting their uncontrolled use. Combined data gathered in the present work unraveled a robust description of the microbiome (community structure and functional diversity) in a full-scale UASB reactor and provided fundamental information about the bioprocess in AD that can be further explored in biotechnology.

ACKNOWLEDGMENTS

This study was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Process no. 2014/16426-0 and 2013/09386-9).

CONFLICT OF INTEREST

None declared.

REFERENCES

Batstone, D.J., & Virdis, B. (2014). The role of anaerobic digestion in the emerging energy economy. Current Opinion in Biotechnology, 27, 142–149.

Battistoni, P., Fava, G., Pavan, P., Musacchio, A., & Cecchi, F. (1997). Phosphate removal in anaerobic liquid by struvite crystallization without addition of chemicals: Preliminary results. Water Research, 31, 2925–2929.

Cabezas, A., Araujo, J.C., Callejas, C., Gales, A., Hamelin, J., & Marone, A., ... Etchebehere, C. (2015). How to use molecular biology tools for the study of the anaerobic digestion process? Reviews in Environmental Science and Bio/Technology 14, 555–593.

Camiloti, P.R., Mockaitis, G., Domingues Rodrigues, J.A., Rissato Zanariolli Damianovic, M.H., Foresti, E., & Zaiat, M. (2014). Innovative anaerobic bioreactor with fixed-structured bed (ABFSB) for simultaneous sulfate reduction and organic matter removal. Journal of Chemical Technology and Biotechnology, 89, 1044–1050.

Campanaro, S., Treu, L., Kouigias, P.G., De Francisci, D., Valle, G., & Angelidaki, I. (2016). Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy. Biotechnology for Biofuels, 9, 26.

Caporaso, G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F., & Costello, E., ... Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. Nature Methods, 7, 335–336.

Carballa, M., Regueiro, L., & Lema, J. M.(2015). Microbial management of anaerobic digestion: Exploiting the microbiome-functionality nexus. Current Opinion in Biotechnology, 33, 103–111.

Chien, I. C., Meschke, J. S., Gough, H. L., & Ferguson, J. F. (2013). Characterization of persistent virus-like particles in two acetate-fed methanogenic reactors. PLoS ONE, 8, e81040. doi:10.1371/journal.pone.0081040

Cox, M. P., Peterson, D. A., & Biggs, P. J. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics, 11(1), pp.1–6. doi: 10.1186/1471-2105-11-485

De bautista los Santos, Q.M., Schroeder, J.L., Blakemore, O., Moses, J., Haffey, M., Sloan, W., & Pinto, A.J. (2016). The impact of sampling, PCR, and sequencing replication on discerning changes in drinking water bacterial community over diurnal time-scales. Water Resource, 90, 216–224.

de Nardi, I. R., Fuzzi, T. P., & Del Nery, V. (2008). Performance evaluation and operating strategies of dissolved-air flotation system treating poultry slaughterhouse wastewater. Resources, Conservation and Recycling, 52, 533–544.

Del Nery, V., Damianovic, M. H. Z., Pozzi, E., de Nardi, I. R., Caldas, V. E. A., & Pires, E. C. (2013). Long-term performance and operational strategies of a poultry slaughterhouse waste stabilization pond system in a tropical climate. Resources, Conservation and Recycling, 71, 7–14.

Delforno, T. P., Moura, A. G. L., Okada, D. Y., Sakamoto, I. K., & Varesche, M. B. A. (2015). Microbial diversity and the implications of sulfide levels in an anaerobic reactor used to remove an anionic surfactant from laundry wastewater. Bioresource Technology, 192, 37–45.
short-read next generation DNA sequencing. *Biotechnology for Biofuels*, 5. doi:10.1186/1754-6834-5-41

Wong, M.T., Zhang, D., Li, J., Hui, R.K.H., Tun, H.M., & Brar, M.S. (2013). Towards a metagenomic understanding on enhanced biomethane production from waste activated sludge after pH10 pretreatment. *Biotechnology for Biofuels*, 6, 38.

Yang, Y., Li, B., Ju, F., & Zhang, T. (2013). Exploring variation of antibiotic resistance genes in activated sludge over a four-year period through a metagenomic approach. *Environmental Science and Technology*, 47, 10197–10205.

Zinder, S. H. (1984). Microbiology of anaerobic conversion of organic wastes to methane: Recent developments. *ASM News*, 50, 294–298.

Zumstein, E., Moletta, R., & Godon, J.-J. (2000). Examination of two years of community dynamics in an anaerobic bioreactor using fluorescence polymerase chain reaction (PCR) single-strand conformation polymorphism analysis. *Environmental Microbiology*, 2, 69–78.

**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the supporting information tab for this article.

*How to cite this article:* Delforno TP, Lacerda GV, Noronha MF, Sakamoto IK, Varesche MBA, and Oliveira VM. Microbial diversity of a full-scale UASB reactor applied to poultry slaughterhouse wastewater treatment: Integration of 16S rRNA gene amplicon and shotgun metagenomic sequencing. *MicrobiologyOpen*. 2017;6:e443. https://doi.org/10.1002/mbo3.443