Wastewater genomic surveillance for the detection of emergence and transmissible diseases in farms using Nanopore sequencing

Presented by: Gabriela Maria Melgarejo Anamaria

Directed & tutorized by: Olga Francino Martí & Joan Martí Carreras

Digital signature

GABRIELA MARIA MELGAREJO ANAMARIA - DNI 122020329

Firmado digitalmente por GABRIELA MARIA MELGAREJO ANAMARIA - DNI 122020329
Fecha: 2023.06.28 12:20:49 +02'00'

Gabriela Melgarejo Anamaría

Digital signature

FRANCINO MARTI OLGA - 46222870 T

Digitally signed by FRANCINO MARTI OLGA - 46222870T
Date: 2023.06.28 12:25:17 +02'00'

Olga Francino Martí

Digital signature

MARTI CARRERAS JOAN - 41563671 A

Digitally signed by MARTI CARRERAS JOAN - 41563671A
Date: 2023.06.28 12:23:11 +02'00'

Joan Martí Carreras

AKNOWLEDGEMENTS

I want to thank my parents for their unconditional support and encourage me to take this master's degree abroad. I also want to thank dr. Olga Francino and dr. Joan Martí for advising me on this project. Likewise, to Marcel, Marina and Daniel for their patience and teaching. Finally, I want to thank the staff of Servei de Granges i Camps Experimentals for helping during the sampling.

INDEX

RESUME

Wastewater surveillance is based on monitoring pathogens in an aquatic environment. This technique can be useful to track potentially zoonotic diseases harbored in farm production through the slurry, in which methods like passive and grab samplings are used to collect and trace this residual waste. Therefore, the aim of this research was to conduct wastewater genomic surveillance in a swine farm using nanopore sequencing to identify possible pathogens and antibiotic resistance gens. In addition, we want to assess the effect of different sampling techniques for wastewater microorganism monitoring. The study was performed in an experimental swine farm at the Faculty of Veterinary Sciences in the Universitat Autònoma de Barcelona. In the study set-up, 26 piglets (2 pens of 13 piglets each) were sampled (rectal swabs) and the slurry was manually collected using a Moore swab (passive sampling) and a sterile immersion bottle (grab sampling) from the collection basin. Sampling was performed on four different dates. DNA was extracted and pooled by type of sampling (rectal swab, Moore swab and immersion bottle), biological replicate (pen A vs pen B) and by date. The sequencing was performed using nanopore sequencing, from Oxford Nanopore Technology. Sequencing reads were analyzed using the platform EPI2ME to classify them based on taxonomy and to detect antimicrobial resistance genes (AMR). The reads obtained were rarefied and data was analyzed with the Vegan R package. Alpha diversity was estimated in each sample with two different metrics: Richness and Shannon index. No statistical differences were found on Richness ($p = 0.243$, $p = 0.126$, $p = 0.741$) or Shannon index ($p = 0.3135$, $p = 0.1002$, $p = 0.5140$) between sampling techniques (rectal swab, Moore swab and immersion bottle), implying that each microbial community is similarly diverse. Regarding beta diversity, a pair-wise comparison of Bray-Curtis dissimilarities was conducted between types of sampling. No statistical difference in Bray-Curtis dissimilarities was observed between Moore swabs and immersion bottles ($p = 0.65$), indicating that both methods retrieve similar microbial community structures. In contraposition, the microbiome community structure of rectal swabs and Moore swab communities, and rectal swabs and immersion bottles were statistically different ($p = 0.043$ and $p = 0.043$), showing that the microbial community structure is different in swine rectum/feces and farm slurry. Regarding taxonomy, the most abundant genus found in rectal swab samples was *Prevotella* spp (10 – 11%). Contrary, the most abundant genus in Moore swab and immersion bottle samples was *Arcobacter* spp (25 – 54%). Even though the microbiome community structure of rectal swab and the slurry samples are statistically

different, it is worth noting that several genera, relevant for farm health management, can be identified in all sampling methods. Per example, the genus *Clostridium* spp can be traced in the three communities with an abundance ranged between 2% – 7%, in which species like *C. difficile*, *C. butyricum* and *C. perfringes* were the most frequent. Regarding AMR genes, no significant difference was found between the AMR genes detected in the three sampling methods ($p = 0.6592$, $p = 0.5637$ and $p = 0.8845$), implying that all methods recover AMR genes similarly. The most frequent gene groups detected in the three types of samplings were *ANT(6)s* and the groups of tetracycline resistant genes: *tet40, tet44, tetQ, tetW and tetO*. To conclude by using both methods of sampling it has been able to detect and trace ARM genes in farming slurry. Although, microbiome structure between rectal swab and Moore swab/immersion bottle has been statistically different, it is important to highlight the potential of long read DNA sequencing to detect several potential pathogenic organisms as *C. butyricum, C. perfringes, Arcobacter cryaerophilus* or *Prevotella dentalis* in the three types of samplings.

1. INTRODUCTION

Wastewater surveillance is based on monitoring pathogens as well as toxic chemicals or pharmaceutical residuals in sewage or aquatic environments (Gibbs *et al*., 2017; Sikorski *et al*., 2020). This method might be used to track potentially zoonotic diseases which could affect farm production (UddinKhan, *et al.,* 2013; Bonardi, 2017). Passive and grab samplings are methods used to collect and trace residual wastes, monitor bacterial or viral pathogens in aquatic environments (Gibbs *et al*., 2017).

Passive sampling is also referred to as "trap sampling", in which a sampler is immersed in the sampled medium without an active action for several hours (Bivins *et al*., 2022). These samplers can be designed with a single material, such as cotton gauze (capable of absorption), which can be hung with a fishing line, commonly referred as a "Moore swab" (Liu *et al*., 2021; Bivins *et al.,* 2022). This technique has been frequently used to monitor organic and inorganic chemicals pollutants in the water (Bivins *et al*., 2022). Moreover, it has been used in environmental surveillance to monitor etiological agents like *Salmonella enterica* in raw water, as well as to trace poliovirus in sewage samples near cities (Matrait *et al*., 2018; Liu *et al*., 2021).

The other sampling method, grab sampling, consist of a sample container, preferably made of sterile fluorocarbon polymers, which is filled at a single time point (Darryl *et al*., 2020). This technique is suitable to trace pollutants in a waterhole with sufficient volume collected (0.5 – 1 L) to ensure pollutant quantification during analysis (Valenzuela *et al.,* 2020). Regarding farming animals, this method has been also used for monitoring swine pathogens such as *betaarterivirus suid 1* (PRRSV), porcine circovirus 2 (PCV2) or Influenza A, B, C or D in slurry collected in farms from North Carolina in USA (Bailey *et al.,* 2020).

1.1. Pathogens present in animal farming

Wastewater surveillance can help to identify pathogenic species that might be potentially zoonotic, or which could affect farm production (Bonardi, 2017). Food-producing animals are considered the major reservoirs of zoonotic and food-borne pathogens (Thapaliya *et al*., 2015). For example, in one study, a human norovirus (NoV GII4) was detected in fecal samples of swine and cattle farms in Canada (Mattison *et al*., 2007). Additionally, according to Uddin Khan et al (2013), swine associated emerging zoonosis are widely distributed globally, which could entail a particular concern to human health (Figure 1). For instance, diseases like hepatitis

E virus (HEV), swine influenza or *Clostridium difficile* are categorized as worldwide swine-associated emerging zoonoses (UddinKhan, *et al.*, 2013). Similarly, swine is also an important carrier of *Salmonella* spp, which may cause fever, diarrhea, or eventually it can be mortal in some cases (UddinKhan, *et al.*, 2013; Thapaliya *et al*., 2015). Moreover, healthy swine might be carriers of disease which may pose a threat to human health (Bonardi, 2017). For instance, in one study the proportion of swine infected with *Salmonella* spp, based on fecal analysis, was 5.58%, presenting a risk of zoonotic infection to human populations that should not be neglected (Visscher *et al.,* 2011). Also, diseases like *Salmonella* spp might cause nonthyphoidal salmonellosis, a worldwide human affliction caused by the exposure to contaminated food (Thapaliya *et al*., 2015).

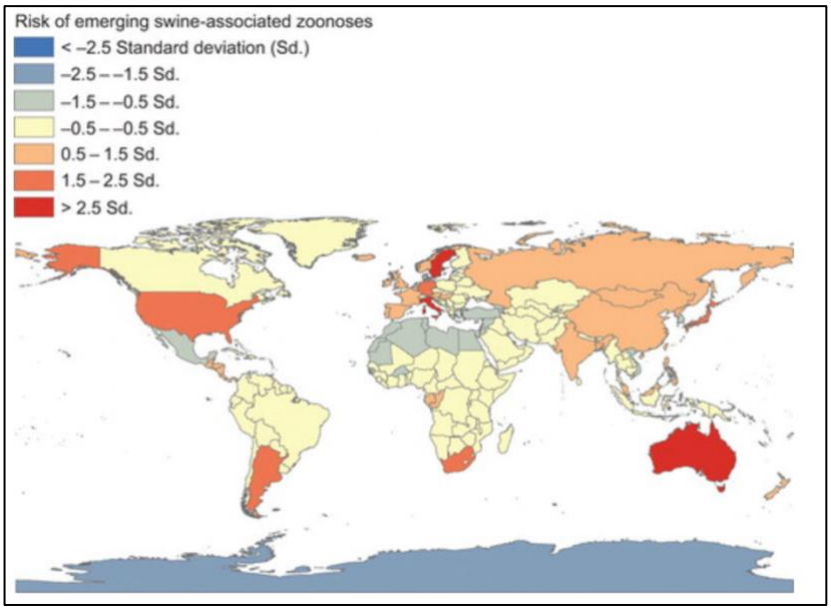


Figure 1. Global distribution of swine-associated emerging zoonoses: 1970 – 2010 (UddinKhan, *et al.*, 2013)

Swine-associated zoonosis can also be present in the food chain supply as food-born outbreaks. In the province of Castellon (Spain), a food-borne outbreak of *Salmonella typhimurium* was associated with the consumption of dried pork sausages (Arnedo *et al.,* 2016). Besides, diseases like Yersiniosis, cause by *Yersinia enterocolitica* serotypes O:3, O:9 and O:5, which can be transmitted by the ingestion of raw or undercooked pork meat (Thapaliya *et al*., 2015). In addition, pathogens like *Campylobacter jejuni or Campylobacter. coli* might be transmitted by the ingestion of undercooked farming meat (Thapaliya, D.*et al.* 2015). For example, in one study it was possible to genotype *C. coli* and *C. jejuni* from retail poultry meat in the district of Goriška (Slovenia) and Central Bosnia Canton (Bosnia) (Zorman *et al.,* 2006).

Furthermore, swine-associated zoonosis can concur with drug-resistant strains. According to Crespo – Piazuelo & Lawlor (2021) people who are in close contact with animals might carry resistance genes (Figure 2). Per example, methicillin–resistant *Staphylococcus aureus* (MRSA) which colonizes nose and throats in swine, may infect humans in close contact with them (Cui *et al.,* 2009; Crespo – Piazuelo & Lawlor, 2021). According to Cu *et al* (2009), MRSA isolates with genes *nuc* and *mecA* were detected in 2/13 swine workers from the province of Shanxi, China (Cui *et al.,* 2009). In Europe, 14.7% of swine workers in Poland were carriers of livestock–associated methicillin–resistant *S. aureus* (LA-MRSA) (Mroczkowska *et al.,* 2017).

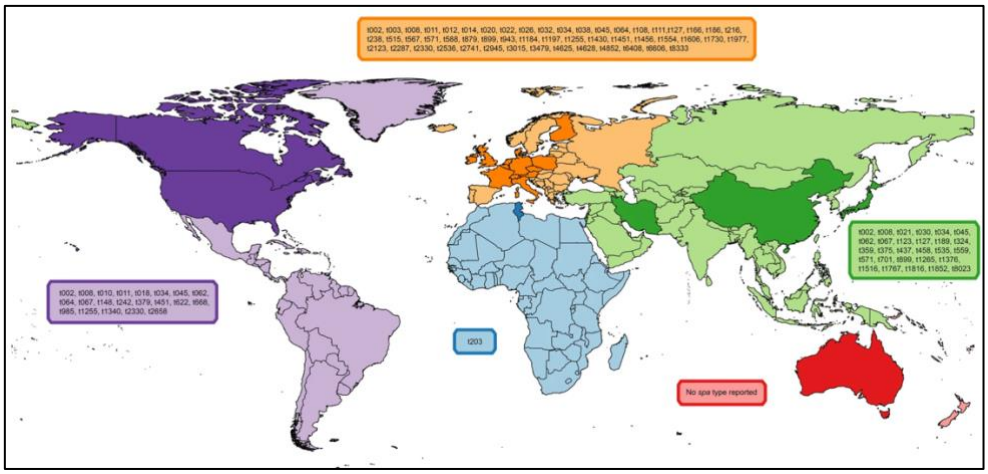


Figure 2. MRSA spa types carried by humans in close contact with animals' groups by continent (Crespo – Piazuelo & Lawlor, 2021)

1.2.Metagenomic surveillance

Environmental metagenomics permits the assembly of whole genomes from soils, waters, surfaces, etc to discover and investigate microbial taxa from different communities (taxonomy annotation) and their biochemical capabilities (functional annotation) (Brown *et al*., 2017; Ciuffreda *et al*., 2021). The microbial sample is sequenced and analyzed by taxonomy and function, helping us to identify an association between phylogeny and function, such as the discovery of antibiotic resistance genes, to identify unculture bacteria hosted in the environment or to find the ecological role of certain taxa in a community (Riesenfeld *et al*., 2004; Ciuffreda *et al*., 2021). Moreover, these methods can help to identify pathogenic species that might be potentially zoonotic or that could affect farm production (UddinKhan, *et al.*, 2013; Bonardi, 2017). It can also be used as a powerful routine tool for diagnostic and epidemiological investigation (Leekitcharoenphon *et al.,* 2014). There has been different techniques to characterize environmental metagenomes; such as short-read sequencing with Illumina (MiSeq)

or ThermoFisher (Ion Torrent) platforms to detect animal diseases, microbiomes, viromes in dromedary fecal samples (Woo *et al.*, 2014), microorganisms of zoonotic concern like *Listeria monocytogenes* or *C. jejuni* or to trace human-associated viruses pathogens in wastewater farming (Bibby *et al*., 2019; Uelze *et al*., 2019; Kwok *et al*., 2022). However, one problem for these methods is the limit of resolution during taxonomy profiling. Hence, one approach to overcome this limitation is the use of third generation long-read sequencing, such as nanopore sequencing (Oxford Nanopore Technologies), which with its higher yield and longer read size contributes to higher taxonomic specificity and functional characterization (Brown *et al*., 2017; Urban *et al*., 2021).

1.3. Third-generation sequencing: Nanopore sequencing in animal health

Nanopore sequencing relies on a nanoscale protein pore that is embedded in an electrically resistant polymer membrane (Lu *et al.,* 2016). As a constant voltage is applied the double-stranded DNA is unwounded and then the single-stranded DNA is driven through the nanopore from the negatively charged "cis" side to the positively charged "trans" side (Wang *et al*., 2021). Then the nucleic acid molecule is translocated and decoded using computational algorithms; allowing real-time sequencing (Figure 3) (Lu *et al.,* 2016; Wang *et al*., 2021). This technology allows to better assess and identify the taxonomy of several bacteria communities, due to the information contain in its long fragments of sequenced DNA (Cuscó *et al*., 2018).

Nanopore sequencing has been widely used as a diagnostic tool for the detection of infectious diseases in human and animal health. For instance, viral pathogens, such as Chikungunya, Ebola and Hepatitis C, were detected in less than 6 hours in clinical samples (Greninger *et al*., 2015), or more recently, rapid nanopore sequencing enabled the genomic characterization of SARS-CoV-2, which was performed from clinical samples from Hangzhou, China (Li *et al*., 2020). Moreover, nanopore sequencing was used as a detection and genomic surveillance tool during the Ebola virus in West Africa and Democratic Republic of Congo outbreaks, using a portable nanopore sequencer, the MinION (Hoenen *et al*., 2016; Quick *et al.,* 2016). Regarding veterinary sciences, this technology has been applied to recover a full-length poxvirus genome in skin lesions of a grey seal and to identify novel species of papillomavirus in giraffe lesions (Günther *et al.,* 2017; Vanmechelen *et al*., 2017). Similarly, nanopore sequencing has been used as a diagnostic tool for porcine viral enteric diseases from diarrheic fecal samples and to detect canine distemper virus through long-read sequencing (Theuns *et al.,* 2018; Peserico *et al*., 2019).

Beyond the identification of pathogens, this technology has been useful for the detection of antimicrobial - resistance (AMR) in sewage. For example, in the district of Shenzhen City (China), it was possible to detect beta-lactam and aminoglycoside resistance genes in treated wastewater using nanopore sequencing (Wu *et al.,* 2022). Even plasmid carriers as *aa*-families, *aad*-families and *sul1* from class 1 integron-integrase genes, could be revealed by nanopore long-reads (Wu *et al.,* 2022). Another example of AMR detected by long-read metagenomic sequencing was performed in five globally distributed wastewater sites from USA, India, Hong Kong, Switzerland and Czech Republic, in which beta-lactam and aminoglycosides genes were detected with a high relative abundance in all five countries (Dai *et al.,* 2022).

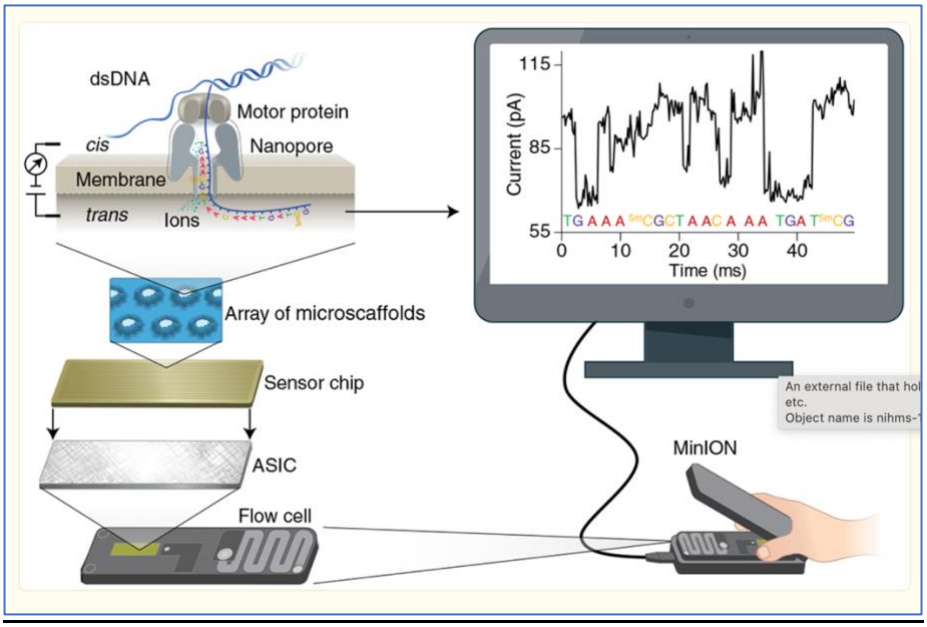


Figure 3. Principle of nanopore sequencing: A MinION flow cell containing 512 channels with 4 nanopores in each channel. The DNA translocates through a pore which applied constant voltage across the membrane where the "trans" side is positively charged. As nucleotides pass through the nanopore, the electric signal is used to determine the corresponding nucleotide type. (Wang *et al*., 2021).

Long read sequencing technologies has been used in wastewater surveillance to monitor and trace potentially pathogenic microorganisms in farm production with different sampling methodologies used to collect sewage in aquatic environments (UddinKhan, *et al.,* 2013; Bonardi, 2017; Matrait *et al*., 2018; Bivins *et al.,* 2022). Taking these into account, this study aims to perform a long read metagenome sequencing using nanopore sequencing, to perform genomic surveillance in an experimental farm and to develop methodologies for slurry

sampling. In addition, this study aims to detect possible pathogens and antibiotic resistant genes in farm slurry, as well as to assess the effect of different sampling.

2. OBJECTIVES

Conduct wastewater genomic surveillance in swine farms using nanopore sequencing to identify possible pathogens and antimicrobial resistance genes.

Specific objectives:

- Feasibility of nanopore sequencing for the identification of microorganisms and antibiotic resistance genes in farm wastewater.
- Ecological analysis and statistical association between wastewater microbiome and rectal samples.
- Assess the effect of different sampling techniques on the study of wastewater microbiome.

3. MATERIALS AND METHODS

3.1. Experimental design

Sampling was performed in an experimental swine farm from the Faculty of Veterinary Sciences in the Universitat Autònoma de Barcelona (Servei de Granges i Camps Experimentals). At the farm, 26 piglets (from approximately five weeks old) were kept in two pens (A and B) of 13 individuals each, with slatted floor sized 2 x 4 meters. The slurry collection basin was emptied before the study. The pens were cleaned and flushed daily. Wastewater was discharged directly to receiving channels under the slats to the collection basin. Pen A and pen B were used as biological replicates for the study.

The sampling consisted of rectal swabs from the individuals and slurry from the collection basin, if available. Sampling was performed on four dates with an approximately difference of 2 weeks, at days: 20th February (arrival), 7th March, 21st March and 28th March (few days before leaving), as detailed in Figure 4.

3.2. Sample collection

In each sampling date, feces were collected directly from the rectum with the DNA/RNA Shield Collection Tube w/Swab (1 ml fill, R1107, Zymo Research Corporation; Los Angeles, CA, USA). No slurry was available upon arrival. The mix of water, urine, manure, etc (slurry) was manually collected from the collection basin by two methods: Both, using a sterile plastic pot (SteriPlast, 5393-7011, Burkle, Bad Bellingen, Germany) as immersion bottle, and a passive composite method with Moore swabs (Liu *et al.,* 2021), which were manually created by pending a screw with a nylon thread, and a sterile gauze. From the second date of sampling, slurry was collected using a Moore swab, and from third date of sampling, slurry was also collected with immersion bottle.

The Moore swabs were installed 24 hours before collection time (composite sample over 24 hours) and the immersion bottle was used, immediately, after rectal collection. Within 10 minutes of collection, samples were stored at 5 °C and processed during the same day.

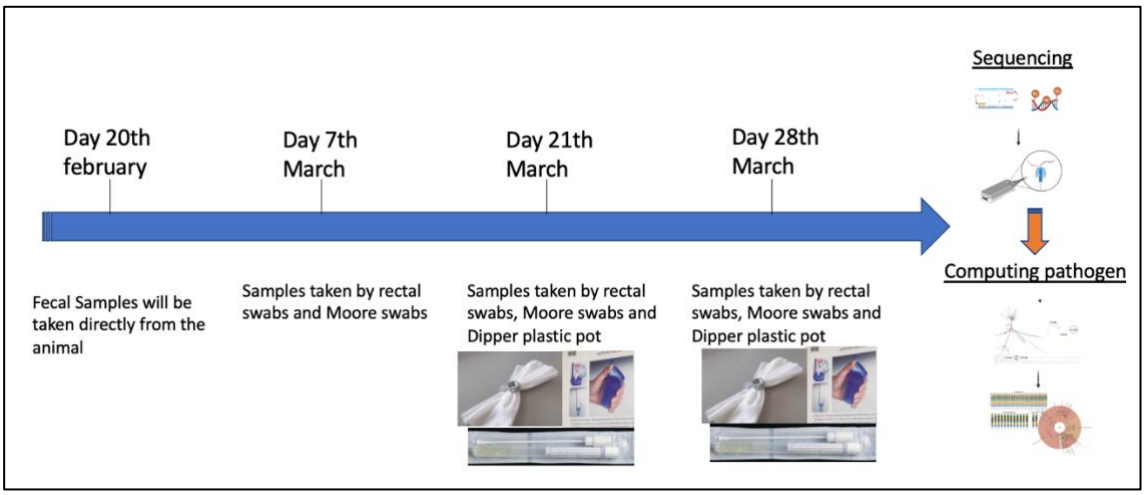


Figure 4. Work plan of this study

3.3.DNA Extraction

DNA was extracted from the rectal swabs using the ZymoBiomics DNA miniprep (D4300; Zymo Research Corporation; Los Angeles, CA, USA) extraction under manufacturer´s condition. Regarding the samples collected with the immersion bottle and the Moore swab (self-made), 200 ml of slurry was centrifuged first at 350 xg for 10 minutes at room temperature to separate large particles and possible host cells from the wastewater. The remaining supernatant (1500 µl) was transferred to a sterile Eppendorf tube and centrifugated at 16 000 xg for 5 minutes to pellet cells. After that, it was washed with 200 µl of phosphate-buffered saline (PBS) at 10% (w/v) twice to reduce sample acidity (Bayle *et al.,* 2020). Finally, 750 µl of lysis was added to the pellet and either used to extract DNA, immediately, according to ZymoBiomics DNA miniprep extraction protocols or stored at 5 °C for batch processing. Additionally, to assess contamination during the process, a sterile swab was processed together with the other samples, which was considered as a negative control. Also, a positive control with the code ZymoBIOMICS Microbial Community: DNA Standard (D6305, Zymo Research Corporation; Los Angeles, CA, USA) was used to assess bias during the sequencing process.

The extracted DNA was qualitatively analyzed using NanoDrop 2000 Spectrophotometer and quantified using a Qubit dsDNA BR Assay kit (Fisher Scientific S.L; Madrid, Spain).

Also, it was pooled by type of sampling, biological replicate (pen A vs pen B) and by date. Pool 19 was considered as a negative control and pool 20 was the positive control (Table 1).

Table 1. Samplings pooled by sampling type (rectal swab, Moore swab and immersion bottle), biological replicate (pen A vs pen B) and day (day 1 to day 4).

Pool	Type	Replicate	Day
Pool1	Rectal swab	A	20th February (Day 1)
Pool2	Rectal swab	B	20th February (Day1)
Pool3	Rectal swab	A	7th March (Day 2)
Pool4	Rectal swab	B	7th March (Day 2)
Pool5	Moore swab	A	7th March (Day 2)
Pool6	Moore swab	B	7th March (Day 2)
Pool7	Rectal swab	A	21th March (Day 3)
Pool8	Rectal swab	B	21th March (Day 3)
Pool9	Moore swab	A	21th March (Day 3)
Pool10	Moore swab	B	21th March (Day 3)
Pool11	Immersion bottle	A	21th March (Day 3)
Pool12	Immersion bottle	B	21th March (Day 3)
Pool13	Rectal swab	A	28th March (Day 4)
Pool14	Rectal swab	B	28th March (Day 4)
Pool15	Moore swab	A	28th March (Day 4)
Pool16	Moore swab	B	28th March (Day 4)
Pool17	Immersion bottle	A	28th March (Day 4)
Pool18	Immersion bottle	B	28th March (Day 4)
Pool19	Negative sample		
Pool20	Positive sample: code D6305 ZymoBIOMICS, DNA standard		

3.4.Library preparation and Sequencing

For sequencing, each pool was diluted to 33.33 ng/µl at equimolar ratios, according to the Native Barcoding Kit 24 V14 kit protocols (SQK – NBD 114.24, Oxford Nanopore Technologies, Oxford, UK). Hence, 20 pools were obtained and maintained in the same concentration of DNA (Table 1). Each pool was attached to a native barcoding of genomic DNA (gDNA) and two libraries, with 10 pools on each one, were loaded with 12 µl at 10 – 20 fmol for sequencing. Finally, two flowcells R10.4.1 (FLO-MIN114, Oxford Nanopore Technologies, Oxford, UK) were used to sequence the DNA extracted in each library.

3.5.Bioinformatic analysis

The sequenced fastq files were classified and analyzed using the platform EPI2ME client v3.6.2 ISO/IEC 27001 (Metrichor Ltd, Oxford, UK). First, the program What´s in my Pot (WIMP) v2023.04.21-1803611 was used to classify sequencing reads based on taxonomy profile. Then, on classified reads, the Antimicrobial Resistance pipeline (AMR) v2023.04.26-1808834 was used for antibiotic resistance gene detection. Regarding the WIMP program, included reads had a length greater than 200 bp and a mean quality score above 20 (phred score). The taxonomic classification was performed by an alignment-free classifier called Centrifuge v1.0.3-beta; which is a microbial classification method, based on Burrows-Wheeler transform and Ferragina-Manzini index, that enables sensitive labeling of reads and quantification of species

at very high speed with a small index database (Kim *et al.,* 2016; Ciuffreda *et al.,* 2021). Additionally, reads are classified against a data-structure based on the NCBI taxonomy (Ciuffreda *et al.,* 2021).

The AMR software is based on The Comprehensive Antibiotic Resistance (CARD) database version 1.1.3 (Alcock *et al*. 2023). This database includes 2357 proteins sequences and 3260 resistance ontology terms.

The WIMP and AMR data were downloaded as csv files. WIMP read annotations or counts were removed if belonged to the host or if ≤ 3 reads. In this study, a relative abundance cut-off of 0.5% was considered for data visualization and statistical analysis. AMR sequencing data with an accuracy ≥98% was analyzed, and AMR genes with less than three reads were removed.

3.6.Statistical analysis

The data was analyzed with R-base v3.6.3 with the taxonomic analysis program Vegan: Community Ecology version 2. 6-4 (Oksanen *et al.,* 2022) and the package Tidyverse v2.0.0 (Wickham *et al*., 2019). First, data normalization (rarefication) was performed by the "*rarefy()*" function in order to compare differences between taxa and compute diversity indexes, taking the number of counts in the smallest pool (19991 counts) and randomly subsampling the same amount of counts in each other sample (Figure 5).

To estimate the diversity inside a community, the alpha diversity was also calculated with the vegan package and two different metrics were use: Richness (number of species in each sampling) was estimated using the "*specnumber()*" function, and the Shannon index, that considers the number and abundance of taxa in the community, with the "*diversity()*" function (Magurran, 2004; Calle, 2019). The variables used for data analysis were biological replicates (A and B), day of sampling (3 and 4) and type of sampling (rectal swab, Moore swab and immersion bottle), as seen in Table 2.

	Samples	Replicates	Day	Type
1	Pool07	A	3	Rectal swab
2	Pool08	B	3	Rectal swab
3	Pool09	A	3	Moore swab
4	Pool10	B	3	Moore swab
5	Pool11	A	3	Immersion bottle
6	Pool12	B	3	Immersion bottle
7	Pool13	A	4	Rectal swab
8	Pool14	B	4	Rectal swab
9	Pool15	A	4	Moore swab
10	Pool16	B	4	Moore swab
11	Pool17	A	4	Immersion bottle
12	Pool18	B	4	Immersion bottle

Table 2. Classification of variables to use in the alpha diversity.

To assess the significance of Richness and Shannon index between groups, a Shapiro-Wilk test of normality was performed. Considering it followed a normal distribution, an independent t – student was performed to assess differences by day of collection and replicates. In addition, a one-way ANOVA test was used to assess the differences between types of sampling (Xia & Sun, 2017). The p-values were corrected for multiple testing using the false discovery test (FDR) method (Corcoll *et al*., 2017).

Beta diversity corresponds to the differences between communities: Bray-Curtis dissimilarity with the function *vegdist (method = "bray"),* which is based on number and abundance of taxa was calculated (Magurran, 2004; Calle, 2019). Those distances matrices were visualized with an NMDS plots using *metaMDS (k=11, trace=F, trymax=100).*

To test differences in beta diversity indexes between groups, the Permutational Multivariate Analysis of Variance (PERMANOVA) test was performed using the function *adonis2()*. The test compares the dissimilarity of samples within the same group against the dissimilarity of samples from a different group (Calle, 2019). Then, all samples are randomly mixed over the groups and the test is repeated many times (Anderson, 2001).

To assess the presence of AMR genes in the different types of samplings (rectal swab, Moore swab and immersion bottle), counts were discretized as a binary variable with 0, meaning no ARM genes detected and 1 meaning the presence of ARM genes. Significance testing was conducted using the Kruskal-Wallis test. The significance testing was performed by type of sampling, replicate and day. For all statistical test, significance was considered with a $p < 0.05$ (after adjustment, if needed) (Xia & Sun, 2017).

4. RESULTS

From the samples obtained, we analyzed 18 pools. Samples from day 1 (pools 01 and 02) were excluded due to the high proportion of host reads with respect to the rest, with a relative abundance of 66% and 64%, respectively (Annex 3). In addition, samples from the second day (pool 03 to pool 06) were also excluded as not enough slurry was present to be collected with the immersion bottle method. Thus, only samples from day 3 and 4 (pool 07 to 18) were analyzed. The csv files downloaded from EPI2ME gave a variable number of classified reads which ranged between 19 991 to 294 000. These taxonomically identified reads were summarized by taxa, here after referred as counts. Considering the large difference of counts between samples, the data was normalized (rarefied) by random sub-sampling the smallest count number between pools (19 991) in each other pool, considering its taxa diversity (Figure 5).

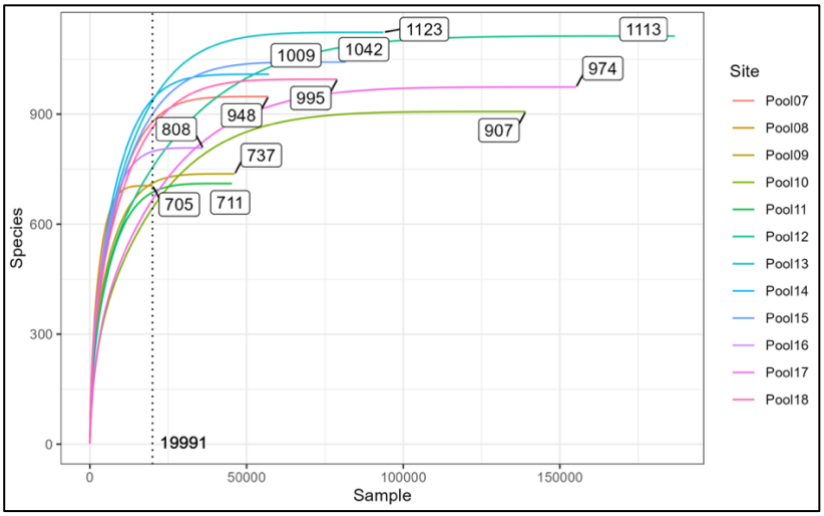


Figure 5: Rarefying curves of the number of species (*y* axes) by the number of counts (*x* axes). The lines are colored by pool from 07 to 18, showing the relationship between the sampled counts with the taxa.

4.1. Microbiome identified

In this research, we were able to identify the taxa of different microorganisms in which the majority of them were from *Bacteria* (98.5%), and the remaining were classified as *Archaea* (1%) and virus (0.5%) (Annex 1).

In this study, we found 30 different families and 50 different genera of *Bacteria*. The most abundant family differs between samples. The most abundant family in rectal swabs is *Lachnospiraceae* (11 – 12.8%), followed by *Prevotellaceae* (10 – 11.38%). Then, family

Campylobacteraceae was the most abundant in Moore swab (27 – 44%) and in the immersion bottle (33 – 55%) samples (Figure. 6). It is worth noting, that the most abundant families shared among the three types of sampling are *Clostridiaceae* (2 – 6%) and *Bacteroidaceae* (2 - 4.5%). Interestingly, the family *Campylobacteraceae* appears in the three types of samplings with a high variable abundance between them: Rectal swabs (0.92 – 1.03%), Moore swab (27.18 – 46.07%) and immersion bottle (33.6 – 55.87%) (Table. 3) (Annex 4).

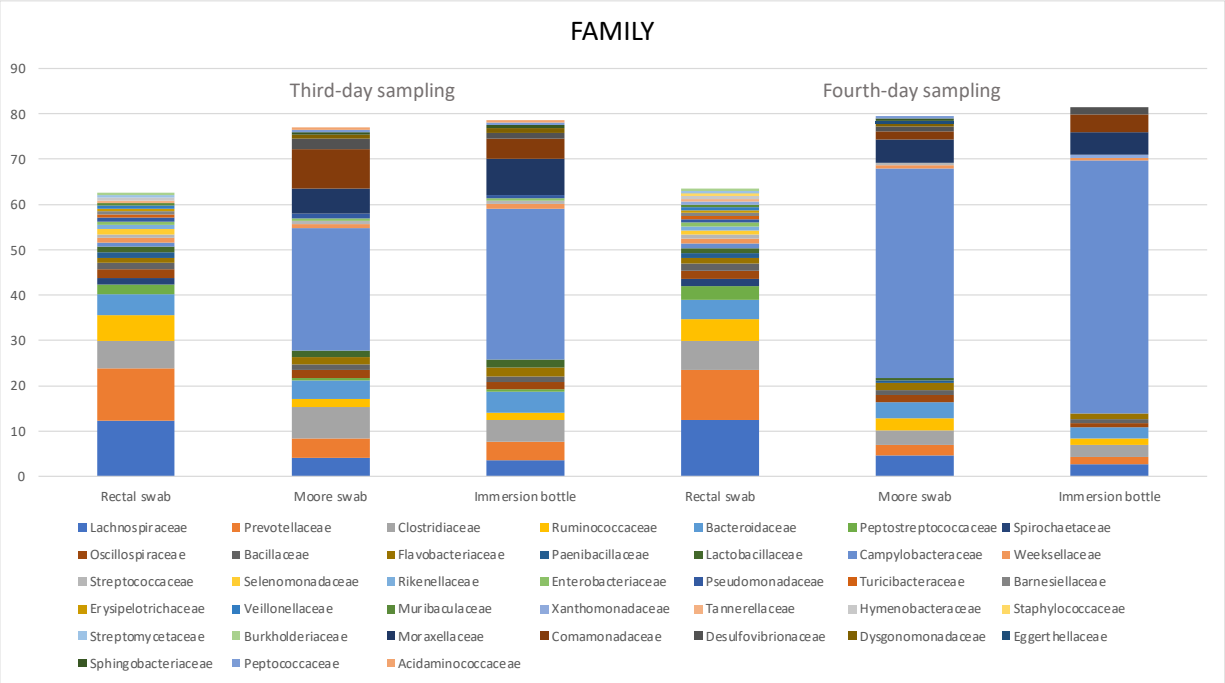


Figure 6. Relative abundance chart of the bacteria family from each type of sampling above 0.5% cut - off.

Table 3. Families shared among the three types of sampling (rectal swab, Moore swab and immersion bottle) by day (third-day and fourth-day sampling).

	Third-day sampling			Fourth-day sampling		
Family	Rectal swab	Moore swab	Immersion bottle	Rectal swab	Moore swab	Immersion bottle
Lachnospiraceae	12.38	4.08	3.65	12.44	4.64	2.65
Prevotellaceae	11.38	4.38	4.02	10.98	2.38	1.71
Clostridiaceae	6.14	6.88	4.81	6.4	3.19	2.64
Ruminococcaceae	5.73	1.85	1.64	4.80	2.61	1.34
Bacteroidaceae	4.50	4.00	4.59	4.42	3.54	2.45
Oscillospiraceae	1.93	1.80	1.67	1.65	1.59	1.05
Flavobacteriaceae	1.03	1.54	1.93	1.21	1.59	1.21
Bacillaceae	1.49	1.26	1.22	1.64	1.08	0.78
Weeksellaceae	0.99	0.83	0.96	0.99	0.73	0.57
Campylobacteraceae	0.92	27.18	33.6	1.03	46.07	55.87

Regarding the genera, the most abundant genus found in rectal swabs was *Prevotella* spp (10 – 11%). Contrary, the most abundant genus in Moore swabs and immersion bottles was *Arcobacte*r spp (25 – 54%) (Figure. 7). Interestingly, several genera with species that are relevant for health and disease management in farms were found in the three types of samplings (rectal swabs, Moore swabs and immersion bottles), such as *Clostridium* spp (2 – 7%) and *Bacteroides* spp (2 – 4.6%) and *Prevotella* spp (1.6 % - 10.6%) (Annex 5).

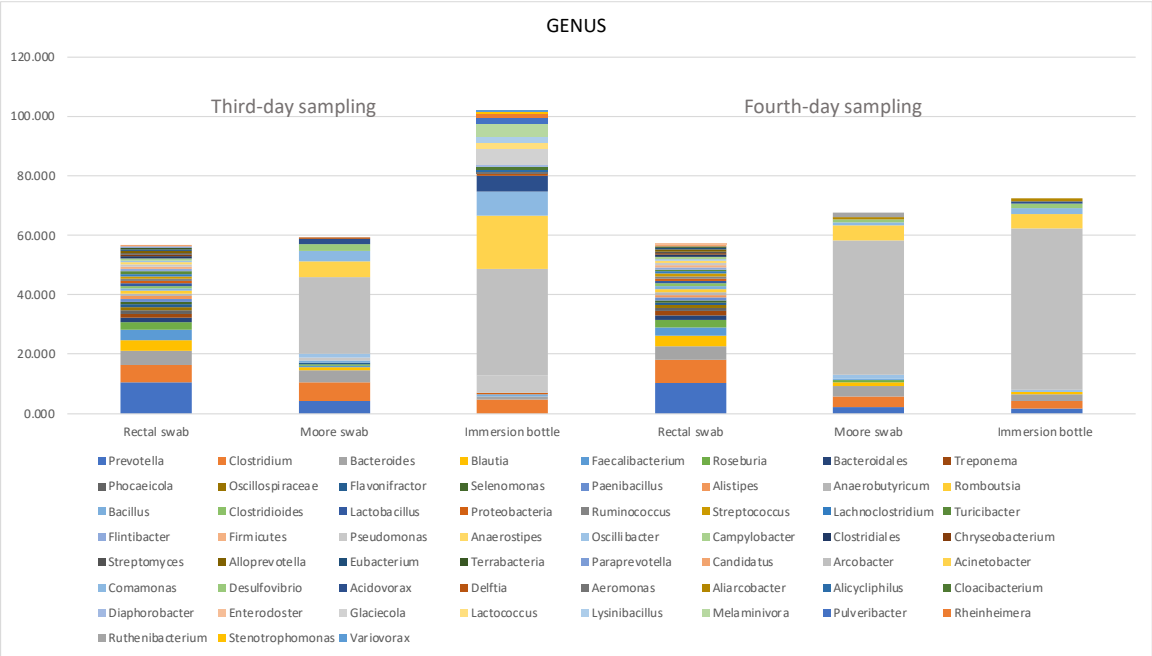


Figure 7: Relative abundance chart of bacterial genera from each sample above 0.5% abundance cut-off.

4.2.Diversity analysis

Regarding alpha diversity, two indexes were considered: Richness (observed taxa) and Shannon index (summary of observed taxa and their abundance). Richness was variable among samples with a minimum range of 705 to a maximum of 932 taxa (Table. 3).

To assess if the taxa diversity depends on the day of sampling or between biological replicates a t-student test was performed (biological replicates $p = 0.829$ and day $p = 0.048$), in which day 4 had a statistically higher mean of 848 genera, compared to day 3 with 727 number of genera. Regarding type of sampling, an ANOVA test was performed (1.648) $p < 0,05$, n^2 $p = 0.246$) (Figure. 8a). No statistical difference was observed during Richness pair-wise comparison between rectal swabs, Moore swabs and immersion bottles ($p = 0.243$, $p = 0.126$, $p = 0.741$).

Regarding Shannon index, Moore swab had the less diverse index 2.96 (Moore swab, day 3, replicate B); meanwhile the rectal swab had, the highest Shannon index5.12 (rectal swab, day 3, replicate B) (Table. 3). Moreover, rectal swabs and Moore swabs have their Shannon indexes in similar ranges (Figure. 8.b). No statistical difference between samplings groups was observed with a one-way ANOVA test ($p = 0.138$). Shannon index pair-wise comparison between types of sampling was also not significant ($p = 0.3135$, $p = 0.1002$, $p = 0.5140$). Additionally, no

statistical difference was found across the two days of sampling p = 0.35 and by biological replicate (A and B) p = 0.782 (Figure 8.b).

Table 3. Alpha diversity indexes (Richness and Shannon index) by day of sampling, type of sample and biological replicates (A and B).

Alpha Index	Third day sampling						Fourth-day sampling					
	Rectal swab(A)	Rectal swab(B)	Moore swab(A)	Moore swab(B)	Immersion bottle(A)	Immersion bottle(B)	Rectal swab(A)	Rectal swab(B)	Moore swab(A)	Moore swab(B)	Immersion bottle(A)	Immersion bottle(B)
Richness	**875**	**705**	**714**	**643**	**690**	**738**	**932**	**938**	**896**	**798**	**664**	**862**
Individuals	19991	19991	19991	19991	19991	19991	19991	19991	19991	19991	19991	19991
Shanon_H	**4.392119**	**5.128403**	**4.221882**	**2.966915**	**3.859489**	**3.501233**	**4.758778**	**4.632238**	**4.52762**	**4.773807**	**3.124609**	**4.593659**

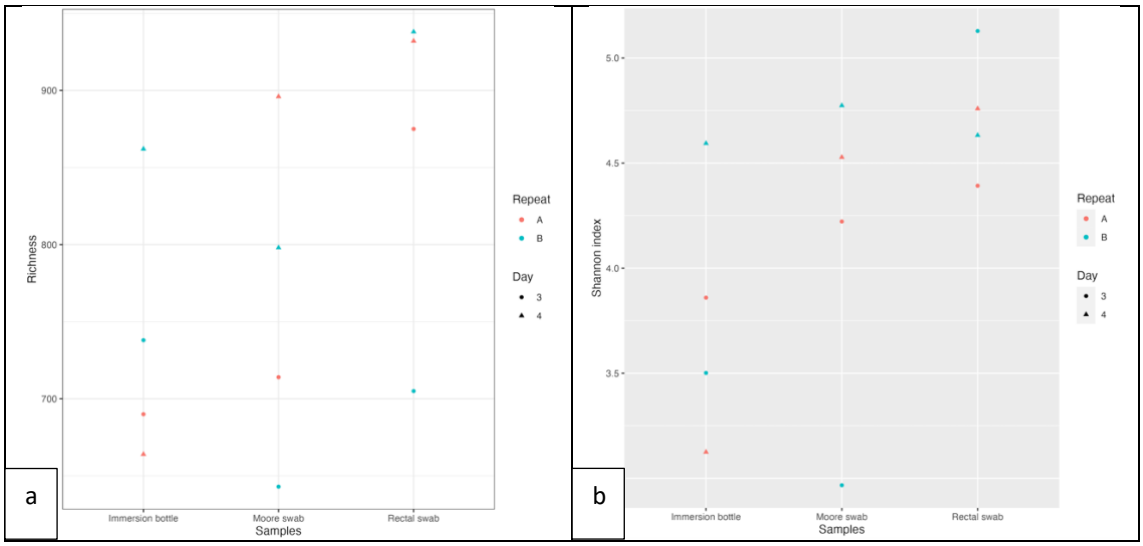


Figure 8: Genera alpha-diversity measured by Richness (a) and Shannon index (b) by type of sampling rectal swab, Moore swab and immersion bottle. Colors are done by biological replicates (A and B) and the shape by day (3 and 4).

The beta diversity analysis assesses the dissimilarities among different communities by the number of taxa and their relative abundance (Bray-Curtis dissimilarity). When the dissimilarities are projected by non-metric multidimensional scaling method (NMDS), rectal swab samples cluster well together with respect to the other samples (Figure 9). Moreover, Moore swab and the immersion bottle samples are scattered together and distanced from rectal swabs by projecting dimension NMDS2 (Figure 9). To assess significant differences of Bray-Curtis dissimilarities between groups of samples a PERMANOVA test was performed by day of sampling (p = 0.4426), biological replicates (p = 0.9564) and type of sampling (p = 0.0377). Pair-wise comparison of Bray-Curtis dissimilarities was conducted between types of sampling, showing no statistical differences between immersion bottle and Moore swabs communities (p

= 0.65). Regardless, significant differences between rectal swab and Moore swab communities (p = 0.043) and rectal swab and immersion bottle (p = 0.043) communities were observed.

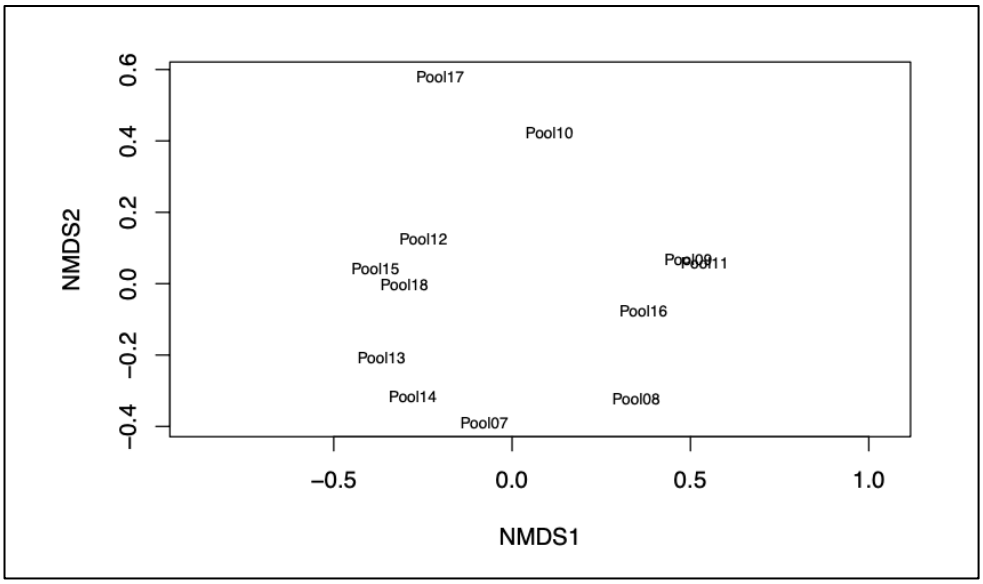


Figure 9: Non-metric multidimensional (NMDS1 and 2) scaling of Bray-Curtis distances done by genera of samples extracted with rectal swab (pool 07, 08, 13 and 14) Moore swab (pool: 09, 10, 15, 16) and immersion bottle (pool: 11, 12, 17, 18).

4.3.Microbial resistance genes

As we can see in Figure 10, 59 AMR genes above the 98% accuracy threshold were identified from the comprehensive Antibiotic Resistance Database (CARD) (Alcock *et al*. 2023). The reported genes were variable between samplings, where some genes were detected only in one type of sampling, as the *OXA-61*, *tetz* or *sat4* genes in pool 07 (rectal swab, day 3, repeat A), pool18 (Immersion bottle, day 4, repeat B and pool13 (rectal swab, day 4, repeat A), respectively (Table 1). The most frequent resistance genes in the three samplings were the, aminoglycoside antibiotic resistant gene family (*ANT(6)s*). Similarly, resistance genes belonging to the tetracycline group: *tet40, tet 44, tetQ, tetW* and *tetO* genes were also present in all sampling types (Table 4).

The Kruskal-Wallis test was performed in order to assess the statistical difference between the number of positives AMR genes by day (p = 0.8092) and by biological replicate (p = 0.7475) or between types of sampling (p = 0.6592, p = 0.5637 and p = 0.8845), where no significant difference was found between the AMR genes detected between rectal swab and Moore swab; rectal swab and immersion bottle; and immersion bottle and Moore swab.

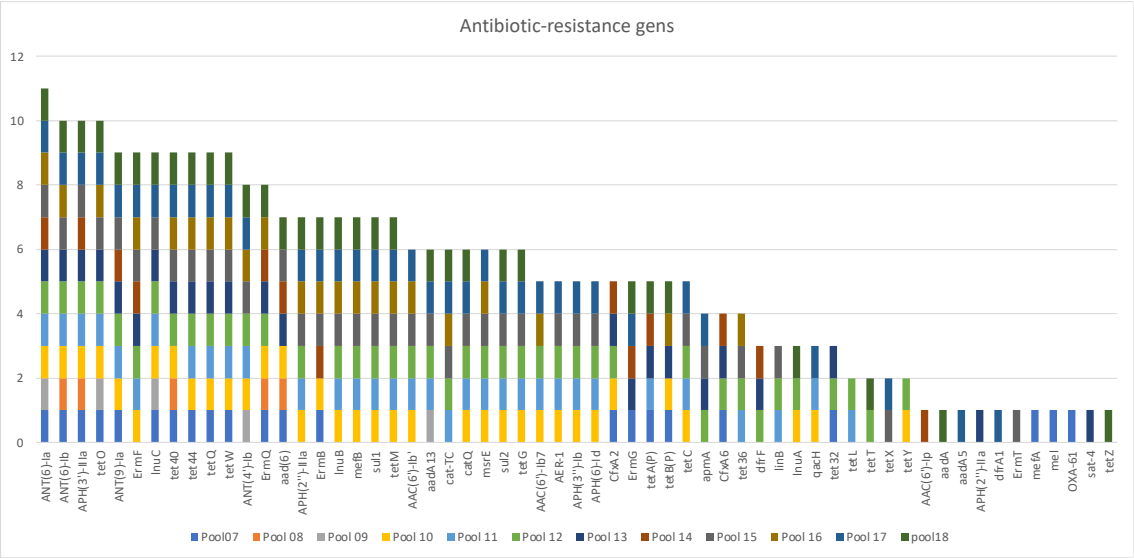


Figure 10: Antibiotic-resistance genes histogram above the 98% accuracy, the CARD genes are presented by sampling of rectal swab (pool07, pool08, pool13, pool14); Moore swab (pool 09, pool10, pool15, pool16) and immersion bottle (pool11, pool12, pool17, pool18).

Table 4. Antibiotic-resistance genes expressed in at least one replicate (A or B) and in the three types of samplings (rectal swab, Moore swab, immersion bottle).

	Third-day sampling						Fourth-day sampling						Positive
Resistance Genes	Rectal swab(A)	Rectal swab(B)	Moore swab(A)	Moore swab(B)	Immersion bottle(A)	Immersion bottle(B)	Rectal swab(A)	Rectal swab(B)	Moore swab(A)	Moore swab(B)	Immersion bottle(A)	Immersion bottle(B)	AMR counts by genes
ANT(9)-Ia	1	0	0	1	1	1	1	1	1	0	1	1	9
ErmF	0	0	0	1	1	1	1	1	1	1	1	1	9
lnuC	1	0	1	1	1	1	1	0	1	0	1	1	9
tet40	1	1	0	1	0	1	1	0	1	1	1	1	9
tet44	1	0	0	1	1	1	1	0	1	1	1	1	9
tetQ	1	0	0	1	1	1	1	0	1	1	1	1	9
tetW	1	0	0	1	1	1	1	0	1	1	1	1	9
ANT(6)-Ib	1	1	0	1	1	1	1	0	1	1	1	1	10
APH(3')-IIIa	1	1	0	1	1	1	1	1	1	0	1	1	10
tetO	1	0	1	1	1	1	1	0	1	1	1	1	10
ANT(6)-Ia	1	0	1	1	1	1	1	1	1	1	1	1	11

5. DISCUSSION

5.1. Microbiome identified

Metagenomic analysis identifies the taxa of different microorganisms in which the majority of them were from *Bacteria* (97%) (Annex 1). Regarding the taxonomy, an increase in the abundance of the family *Campylobacteraceae* was seen in swine wastewater: Moore swabs (27.18 – 46.07%) and immersion bottle (33.6 – 55.87%) compared to rectal swabs (0.9 – 1.03%). Additionally, from that family, the genus *Arcobacter* spp was the most abundant in wastewater (25.72% - 54.11%) in both slurry samplings, being the most abundant species *Arcobacter cryaerophilus* (Annex 6). Contrary, in rectal swabs this species was not detected. Similar results were obtained in other studies performed in wastewater close to an aquaculture center (Ríos – Castro *et al*., 2023), in a river close to Cambridge, UK (Urban *et al*., 2021), in which the most abundant genus was *Arcobacter* spp According to Kayman et al (2012), many species of this genus are implicated in acute gastrointestinal infections, such as *A*. *cryaerophilus* , which has been isolated from clinical cases of foodborne diseases.

For rectal swabs, the families *Lachnospiraceae* and *Prevotellaceae* were the most abundant with a relative abundance between 12.39 – 12.44% and 11.38 – 10.98%, respectively. Despite *Prevotellaceae* being the second most frequent family, the genus *Prevotella* spp was the most abundant genus in rectal swabs (10.18 – 10.61 %), as it is thought to be a significant member of the gut microbiome of swine (Flint & Duncan, 2014). In contrast, it is noticeable that this genus is presented in both slurry sampling methods with a lower abundance, ranged between 1.6% - 4.11%. Furthermore, species like *P. ruminicola, P. intermedia, P. melaninogenica* and *P. dentalis* were also identified in the three types of samplings. According to Ríos-Castro et al (2023), this genus is a potential pathogen present in wastewater, in which species, like *P. intermedia,* can be associated with periodontal diseases (Thomas et al., 2016) (Annex 6).

Additionally, another potential pathogenic organism detected in rectal swabs, Moore swabs and immersion bottles was the genus *Clostridum* spp with a relative abundance between 5 – 7% in rectal swabs but was also detected in the Moore swabs and immersion bottle with an abundance of 3.4 -6.4 % and 2.4 – 4.6%, respectively. Further, species like *C. difficile*, *C. butyricum* and *C. perfringes* were the most detected in the three samplings According to Rìos-Castro et al (2023), this genus is related to food-borne and water-borne diseases. Moreover, species like *C. difficile* are commonly detected in raw sewage and it can survive the wastewater treatment

process; categorizing it as a worldwide swine-associated emerging zoonoses (UddinKhan, *et al.*, 2013; Xu *et al*., 2014). It is important to highlight the potential of long read DNA sequencing to detect several potential pathogenic organisms in the active sampling and passive samplings.

5.2.Diversity analysis

Regarding alpha diversity, differences in Richness was not significant for biological repeats (p = 0.782) neither between type of sampling (rectal swab, Moore swab or immersion bottle) (p = 0.243, p = 0.126, p = 0.741). Similarly, no statistical difference was found in Shannon index between type of samplings (p = 0.3135, p = 0.1002, p = 0.5140). This implies that with both indexes the defined three communities are similarly diverse. Contrary to our results, in one study the Richness and Shannon index were higher in two different wastewater centers compared to an aquaculture center close by (Ríos – Castro *et al*., 2023). Regarding sampling days, the Richness (number of different species) is statistically different between day 3 (727 genera) and day 4 (848 genera) (p = 0.048), whereas Shannon index (number of different species by abundance) is statistically similar between the same time periods (p = 0.35). Implying that over time (approximately 15 days) the number of low abundant genera increases.

Additionally, beta diversity, Bray-Curtis dissimilarities were statistical similar between the communities in the Moore swabs and the immersion bottles (p=0.65). That suggest that it is possible to retrieve similar microbiome community structures with both Moore swabs (passive) and immersion bottle (grab samplings).

Conflicting results have been reported in literature between passive and grab sampling when used to track specific pathogens in wastewater. Rafiee et al (2021), where Moore swabs had equal results than grab sampling for SARS-CoV-2 monitoring. Nevertheless in another study passive sampling with Moore swabs had better sensitivity to detect virus of SARS-CoV-2 (Corchis-Scott *et al* 2021), Meanwhile, according to Bailey et al (2018), grab sampling has given good results as surveillance method for the detection of different microorganisms like *E. coli*, *Salmonella* spp, *Clostridium. perfringes*, *Norvovirus* GII, *Adenovirus* A-F or *Cryptosporidium* in stagnant wastewater.

It is worth noting that both type of samplings has been widely used to study raw wastewater or to monitor chemical pollutants (Schopfer *et al*., 2014; Rafiee *et al*., 2021). As well as to monitor bacterial and viral pathogens in aquatic environments or household drains (Gibbs *et al*., 2017).

Conversely, the microbiome community structure between rectal swabs and Moore swabs, and rectal swabs and immersion bottles have statistically different Bray-Curtis dissimilarities ($p = 0.043$ and $p = 0.043$). This indicates that the microbial community structure is different in swine rectum/feces and farm slurry.

Similar studies have been conducted, comparing feces microbiome and passive sampling methods (air particles or dust). In such studies, conflicting evidence has been found whether dust microbiome is a good sampling representation of fecal microbiome. In Luiken et al., (2020) poultry fecal microbiome and its dust microbiome was found to be statistically different, meanwhile in Kwok et al., (2022) both communities were statistically significant. The same study of Luiken et al., (2020), repeated the experiment in a swine farm, where the swine fecal microbiome and its dust microbiome was found to be statistically similar.

Santiago-Rodriguez & Hollister (2023) suggest that there are potential sources of bias which might influence the microbiome results. Such factors include sample processing, sequencing specificity or bioinformatic tools. In relation to the latter, it should be highly specific and sensitive (Shen *et al*., 2021; Santiago-Rodriguez & Hollister, 2023). For our study, the affiliation of microorganisms to a taxon was performed with Centrifuge, integrated in the ONT-platform EPI2ME, which enable a sensitive labeling of reads, as well as it has been used in other microbiome studies in wastewater (Kim *et al.,* 2016; Ciuffreda *et al.,* 2021). Although, Urban *et al* (2021) affirms that this bioinformatic tool yields less accurate classification than others such as Minimap2. Centrifuge provides a rapid and accurate metagenomic analysis which can reach a sensitivity > 90% as well as it is able to detect the most abundant species (Brown *et al*., 2017; Cuscó *et al*., 2018).

Furthermore, another source of bias would be the abundance of low taxa and a high percentage of unclassified reads (Annex 2), which might have an influence in the Bray-Curtis index. Additionally, in our results there was a high proportion of identified species below the 0.1% relative abundance threshold, as well as a dominant species in both slurry samplings with an abundance ranged between 25.72% - 54.11%.

5.3. Microbial resistance genes

Regarding AMR genes, no significant difference was found between the AMR genes detected between the three sampling methods ($p = 0.6592$, $p = 0.5637$ and $p = 0.8845$), suggesting that all methods recover AMR genes similarly. These results are in agreement with a study that reported a significant correlation between swine dust farm and feces antimicrobial resistomes (Luiken *et al*., 2020). This result supports the idea that slurry farm sampling can provide a qualitative detection of antibiotic resistance genes (presence/absence and type). Hence, this type of sampling can be incorporated as part of wastewater genomic surveillance to trace AMR genes in farm slurry.

The most frequent genes detected in the three types of sampling were the family *ANT(6)s* and five tetracycline resistant genes: *tet40, tet44, tetQ, tetW and tetO*. In agreement with a previous study, the most prevalent resistant genes detected in swine slurry were the ones for tetracycline (Ramesh *et al*., 2021). Similar results were found in another study performed in swine farm wastewater, where it was possible to detect ten resistance genes: *tetA, tetC, tetG, tetH, tetL, tetM, tetO, tetQ, tetW* and *tetX* as well as one aminoglycoside resistant gene (*strB*) (Yang *et al*., 2019). These techniques have also been used to detect other antibiotic resistance genes in wastewater and manure of swine, such as doxycycline, sulfonamides, fluoroquinolones and macrolides (Zhang *et al*., 2018; Chan *et al*., 2022). Chan et al (2022) elaborated that swine feed and drinking are generally mixed with antibiotics to prevent diseases, hence there is high concentration of antibiotics in the wastewater, pond, sediments and sludge. Therefore, the frequent detection of tetracycline resistant genes (tet) in swine wastewater from the present and previous studies suggest the extensive use of this drug in veterinary medicine as well as its intensive usage in feed (Zhou *et al*., 2013; Chan *et al*., 2022).

6. CONCLUSIONS

No statistical differences were found on Richness ($p = 0.243$, $p = 0.126$, $p = 0.741$) or Shannon index ($p = 0.3135$, $p = 0.1002$, $p = 0.5140$) between sampling techniques (rectal swab, Moore swab and immersion bottle), implying that each microbial community is similarly diverse.

No statistical difference in Bray-Curtis dissimilarities was observed between Moore swabs and immersion bottles ($p = 0.65$), indicating that both methods retrieve similar microbial community structures.

Microbiome community structure of rectal swabs and Moore swab communities, and rectal swabs and immersion bottles were statistically different ($p = 0.043$ and $p = 0.043$), showing that the microbial community structure is different in swine rectum/feces and farm slurry

Even though the microbiome community structure of rectal swab and the slurry samples are statistically different, it is worth noting that several genera, relevant for farm health management, can be identified in all sampling methods.

Metagenomic surveillance in swine fecal slurry has potential for antimicrobial- resistance genes detection.

This work has some limitations, which is the relationship of bacterial compositions with environmental parameters as pH, temperature or inorganic solutes. Also, it was limited by the seasonality when taking the samplings, which should be considered for future microbiome analysis of wastewater.

7. BIBLIOGRAPHY

Alcock et al. 2023. CARD 2023: Expanded Curation, Support for Machine Learning, and Resistome Prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Research, 51, D690-D699.

Anderson, M. 2001. A new method for non-parametric multivariate analysis of variance. Centre for Research on Ecological Impacts of Coastal Cities, Marine Ecology Laboratories, Australia. Austral Ecology (2001) 26, 32–46

Arnedo-Pena A, et al. 2016. An outbreak of monophasic and biphasic Salmonella Typhimurium, and Salmonella Derby associated with the consumption of dried pork sausage in Castellon (Spain). Enfermedades Infecciosas Microbiología Clínica; 34: 544–550. doi: 10.1016/j.eimc.2015.11.016

Bailey, E., Borkenhagen., L., Choi, J., Greer., A., Culhane, M., Gray, G. 2020. A feasibility study of conducting surveillance fro swine pathogens in slurry from North Carolina swine farms. Scientific Reports. V. 10:10059 | https://doi.org/10.1038/s41598-020-67313-x.

Bailey, E. Casanova, L. Simmons, O. Sobsey, M. 2018. Teritary treatment and dual disinfection to improve microbial quality of reclaimed water for potable and non-potable reuse: A case study of facilities in North Carolina. Science of the Total Environment 630; 379–388.

Bibby, K., Crank, K., Greaves, J. *et al.* Metagenomics and the development of viral water quality tools. *npj Clean Water* **2**, 9 (2019). https://doi.org/10.1038/s41545-019-0032-3

Bivins, A. Kaya, D. Ahmed, W. Brown, W. Butler, C. Greaves, J. Maas, K. Rao, G. Sherchan, S. Sills, D. Sinclair, R et all. 2022. Passive sampling to scale wastewater surveillance of infectious disease: Lessons learned from COVID-19. Science of The Total Environment, Volume 835, 2022, 155347, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2022.155347

Bonardi S. 2017. Salmonella in the pork production chain and its impact on human health in the European Union. Epidemiol Infect. Jun;145(8):1513-1526. doi: 10.1017/S095026881700036X. Epub 2017 Feb 28. PMID: 28241896; PMCID: PMC9203350.

Brown BL, Watson M, Minot SS, Rivera MC, Franklin RB. 2017. MinION™ nanopore sequencing of environmental metagenomes: a synthetic approach. Gigascience. Mar 1;6(3):1-10. doi: 10.1093/gigascience/gix007. PMID: 28327976; PMCID: PMC5467020.

Calle ML. 2019. Statistical Analysis of Metagenomics Data. Genomics Inform. 2019 Mar;17(1):e6. doi: 10.5808/GI.2019.17.1.e6. PMID: 30929407; PMCID: PMC6459172.

Chan, R. Chiemchaisri, C. Chiemchaisri, W. Boonsoongnern, A. Tulayakul, P. 2022. Occurrence of antibiotics in typical pig farming and its wastewater treatment in Thailand,
Emerging Contaminants, Volume 8, Pages 21-29, ISSN 2405-6650, https://doi.org/10.1016/j.emcon.2021.12.003.

Ciuffreda, L. L. Rodríguez-Pérez, H. Flores, C. 2021. Nanopore sequencing and its application to the study of microbial communities, Computational and Structural Biotechnology Journal, Volume 19, Pages 1497-1511, ISSN 2001-0370, https://doi.org/10.1016/j.csbj.2021.02.020.

Corcoll, N., Osterlund, T. Sinclair, L. Eiler, A., Kristiansson., E., Backhaus, T., Eriksson, M. 2017. Comparison of four DNA extraction methods for comprehensive assessment of 16S rRNA bacterial diversity in marine biofilms using high-throughput sequencing, FEMS Microbiology Letters, Volume 364. fnx139, https://doi.org/10.1093/femsle/fnx139

Crespo-Piazuelo, D., Lawlor, P.G. Livestock-associated methicillin-resistant Staphylococcus aureus (LA-MRSA) prevalence in humans in close contact with animals and measures to reduce on-farm colonisation. Ir Vet J 74, 21 (2021). https://doi.org/10.1186/s13620-021-00200-7

Cui S, Li J, Hu C, Jin S, Li F, Guo Y, Ran L, Ma Y. 2009. Isolation and characterization of methicillin-resistant Staphylococcus aureus from swine and workers in China. J Antimicrob Chemother. doi: 10.1093/jac/dkp275. Epub 2009 Aug 14. PMID: 19684078.

Cuscó, A. Catozzi, C. Viñes, J. Sanchez, A. Francino, O. 2018. Microbiota profiling with long amplicons using Nanopore sequencing: full-length 16S rRNA gene and the 16S-ITS-23S of the *rrn* operon. F1000Res. 6;7:1755. doi: 10.12688/f1000research.16817.2. PMID: 30815250; PMCID: PMC6372929

Darryl W. Hawker, Joseph Clokey, Sara G. Gorji, Rory Verhagen, Sarit L. Kaserzon, 2022. Chapter 3 - Monitoring techniques–Grab and passive sampling, Editor(s): Tatenda Dalu, Nikita T. Tavengwa, Emerging Freshwater Pollutants, Elsevier, Pages 25-48,
ISBN 9780128228500, https://doi.org/10.1016/B978-0-12-822850-0.00014-4.

Greninger, A. L. et al. 2015. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med 7, 99, https://doi.org/10.1186/s13073-015-0220-9.

Günther, T., Haas, L., Alawi, M. et al. (2017). Recovery of the first full-length genome sequence of a parapoxvirus directly from a clinical sample. Sci Rep 7, 3734 https://doi.org/10.1038/s41598-017-03997-y

Gibbs JL, Yost MG, Negrete M, Fenske RA. 2017. Passive Sampling for Indoor and Outdoor Exposures to Chlorpyrifos, Azinphos-Methyl, and Oxygen Analogs in a Rural Agricultural Community. Environ Health Perspect. Mar;125(3):333-341. doi: 10.1289/EHP425. Epub 2016 Aug 12. PMID: 27517732; PMCID: PMC5332193.

Hoenen, T. et al. Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool. Emerging infectious diseases 22, 331–334, https://doi.org/10.3201/eid2202.151796 (2016).

H.J. Flint, S.H. Duncan, Bacteroides and Prevotella, Editor(s): Carl A. Batt, Mary Lou Tortorello, 2014. Encyclopedia of Food Microbiology (Second Edition), Academic Press, , Pages 203-208, ISBN 9780123847331, https://doi.org/10.1016/B978-0-12-384730-0.00031-8.

Kayman T, Abay S, Hizlisoy H, Atabay HI˙, Diker KS, Aydin F. 2012. Emerging pathogen Arcobacter spp in acute gastroenteritis: molecular identification, antibiotic susceptibilities and genotyping of the isolated arcobacters. Journal of Medical Microbiology 61:1439–1444. DOI:

https://doi.org/10.1099/jmm.0.044594-0,
PMID: 22700547

Kim D, Song L, Breitwieser FP, Salzberg SL. 2016. Centrifuge: rapid and sensitive classification of metagenomic sequences. Genome Res. 26(12):1721-1729. doi: 10.1101/gr.210641.116. Epub 2016 Oct 17. PMID: 27852649; PMCID: PMC5131823.

Kwok KTT, de Rooij MMT, Messink AB, Wouters IM, Smit LAM, Cotten M, Heederik DJJ, Koopmans MPG, Phan MVT. 2022. Establishing farm dust as a useful viral metagenomic surveillance matrix. Sci Rep. Sep 29;12(1):16308. doi: 10.1038/s41598-022-20701-x. PMID: 36175536; PMCID: PMC9521564.

Quick, J. et al. 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature 530, 228–232, https://doi.org/10.1038/ nature16996 (2016).

Li, J., Wang, H., Mao, L. et al. 2020. Rapid genomic characterization of SARS-CoV-2 viruses from clinical specimens using nanopore sequencing. Sci Rep 10, 17492. https://doi.org/10.1038/s41598-020-74656-y

Liu P, Ibaraki M, Kapoor R, Amin N, Das A, Miah R, Mukhopadhyay AK, Rahman M, Dutta S and Moe CL (2021) Development of Moore Swab and Ultrafiltration Concentration and Detection Methods for Salmonella Typhi and Salmonella Paratyphi A in Wastewater and Application in Kolkata, India and Dhaka, Bangladesh. Front. Microbiol. 12:684094. doi: 10.3389/fmicb.2021.684094

Leekitcharoenphon P, et al. 2014. Evaluation of whole genome sequencing for outbreak detection of Salmonella enterica. PLoS ONE 2014; 9: e87991. Published on line: 4 February. doi.org/10.1371/journal.pone.0087991. [PMC free article] [PubMed] [Google Scholar]

Luiken REC, Van Gompel L, Bossers A, Munk P, Joosten P, Hansen RB, Knudsen BE, García-Cobos S, Dewulf J, Aarestrup FM, Wagenaar JA, Smit LAM, Mevius DJ, Heederik DJJ, Schmitt H; EFFORT-group. 2020. Farm dust resistomes and bacterial microbiomes in European poultry and pig farms. Environ Int. 2020 Oct;143:105971. doi: 10.1016/j.envint.2020.105971.

Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. Genomics, Proteomics & Bioinformatics, 14(5), 265–279. https://doi.org/10.1016/J.GPB.2016.05.004

Magurran, A. E. 2004. Measuring biological diversity. Blackwell, Oxford, United Kingdom.

Matrajt, G., Naughton, B., Bandyopadhyay, A.S., Meschke, J.S., 2018. A review of the most commonly used methods for sample collection in environmental surveillance of poliovirus. Clin. Infect. Dis. 67, S90–S97. https://doi.org/10.1093/cid/ciy638.

Mroczkowska A, Żmudzki J, Marszałek N, Orczykowska-Kotyna M, Komorowska I, Nowak A, et al. 2017. Livestock-associated Staphylococcus aureus on Polish pig farms. PLoS ONE.;12: e0170745.

Oksanen J, Gavin L. Simpson, F. Guillaume Blanchet, Roeland, Kindt, et al., 2022. Vegan: Community Ecology Package. R package version 2.6-4. https://CRAN.R-project.org/package=vegan

Peserico A, Marcacci M, Malatesta D, Di Domenico M, Pratelli A, Mangone I, D'Alterio N, Pizzurro F, Cirone F, Zaccaria G, Cammà C, Lorusso A. Diagnosis and characterization of canine distemper virus through sequencing by MinION nanopore technology. Sci Rep. 2019 Feb 8;9(1):1714. doi: 10.1038/s41598-018-37497-4. PMID: 30737428; PMCID: PMC6368598.

Rafiee, M., Isazadeh, S., Mohseni-Bandpei, A., Mohebbi, S.R., Jahangiri-rad, M., Eslami, A., Dabiri, H., Roostaei, K., Tanhaei, M., Amereh, F., 2021. Moore swab performs equal to composite and outperforms grab sampling for SARS-CoV-2 monitoring in wastewater. Sci. Total Environ. 790, 148205. https://doi.org/10.1016/j.scitotenv.2021.148205

Ramesh A, Bailey ES, Ahyong V, Langelier C, Phelps M, Neff N, Sit R, Tato C, DeRisi JL, Greer AG, Gray GC. Metagenomic characterization of swine slurry in a North American swine farm operation. Sci Rep. 2021 Aug 20;11(1):16994. doi: 10.1038/s41598-021-95804-y. PMID: 34417469; PMCID: PMC8379149.

Riesenfeld CS, Schloss PD, Handelsman J. 2004. Metagenomics: genomic analysis of microbial communities. Annu Rev Genet. 38:525-52. doi: 10.1146/annurev.genet.38.072902.091216. PMID: 15568985.

Ríos-Castro R, Cabo A, Teira E, Cameselle C, Gouveia S, Payo P, Novoa B, Figueras A. 2022. High-throughput sequencing as a tool for monitoring prokaryote communities in a wastewater treatment plant. Sci Total Environ. 2023 Feb 25;861:160531. doi: 10.1016/j.scitotenv.2022.160531.

Santiago-Rodriguez TM, Hollister EB. 2023. Viral Metagenomics as a Tool to Track Sources of Fecal Contamination: A One Health Approach. *Viruses*. 15(1):236. https://doi.org/10.3390/v15010236

Shen J, McFarland AG, Young VB, Hayden MK, Hartmann EM. Toward Accurate and Robust Environmental Surveillance Using Metagenomics. Front Genet. 2021 Mar 5;12:600111. doi: 10.3389/fgene.2021.600111. PMID: 33747038; PMCID: PMC7973286.

Schopfer, A. Estoppey, N. Omlin J. Udrisard, R. Esseiva, P. Alencastro, L. 2014. The use of passive samplers to reveal industrial and agricultural pollution trends in swiss. Environmetal chemistry in Switzerland CHIMIA,68, Nr. 1

Thapaliya, D. et al. (2015). Zoonotic Diseases of Swine: Food-borne and Occupational Aspects of Infection. In: Sing, A. (eds) Zoonoses - Infections Affecting Humans and Animals. Springer, Dordrecht. https://doi-org.are.uab.cat/10.1007/978-94-017-9457-2_2

Theuns, S. Vanmechlen, B. Bernaert, Q. Deboutte, Q. Vandenhole, M. et al. 2018. Nanopore sequencing as a revolutionary diagnostic tool for porcine viral enteric disease complexes identifies porcine kobuvirus as an important enteric virus. 8:9830 | DOI:10.1038/s41598-018-28180-9.

Thomas S. Murray, Todd Cassese. 2016. 2 - Bacteriology of the Head and Neck Regions, Editor(s): James R. Hupp, Elie M. Ferneini, Head, Neck, and Orofacial Infections, Elsevier, Pages 27-37. https://doi.org/10.1016/B978-0-323-28945-0.00002-8.

UddinKhan,S.,Atanasova,K.R.,Krueger,W.S.,Ramirez,A.&Gray,G.C.Epidemiology,geographicaldistribution,andeconomic consequences of swine zoonoses: A narrative review. Emerg. Microb. Infect. 2(12), e92. https://doi.org/10.1038/emi.2013.87 (2013).

Mattison, K., Shukla, A., Cook, A., Pollari, F., Friendship, R., Kelton, D....Farber, J. M. (2007). Human Noroviruses in Swine and Cattle. Emerging Infectious Diseases, 13(8), 1184. https://doi.org/10.3201/eid1308.070005.

Uelze L, Borowiak M, Bönn M, Brinks E, Deneke C, Hankeln T, Kleta S, Murr L, Stingl K, Szabo K, Tausch SH, Wöhlke A, Malorny B. German-Wide Interlaboratory Study Compares Consistency, Accuracy and Reproducibility of Whole-Genome Short Read Sequencing. Front Microbiol. 2020 Sep 11;11:573972. doi: 10.3389/fmicb.2020.573972. PMID: 33013811; PMCID: PMC7516015.

Urban L, Holzer A, Baronas JJ, Hall MB, Braeuninger-Weimer P, Scherm MJ, Kunz DJ, Perera SN, Martin-Herranz DE, Tipper ET, Salter SJ, Stammnitz MR. 2021. Freshwater monitoring by nanopore sequencing. Elife. Jan 19;10:e61504. doi: 10.7554/eLife.61504. PMID: 33461660; PMCID: PMC7815314.

Valenzuela, E.F., Menezes, H.C. & Cardeal, Z.L. Passive and grab sampling methods to assess pesticide residues in water. A review. Environ Chem Lett 18, 1019–1048 (2020). https://doi-org.are.uab.cat/10.1007/s10311-020-00998-8

Vanmechelen, B. Frost Bertelsen, M. Rector, A. Van den Oord, J. Laenen, L. Vergote, V. Maes, P. 2017. Identification of a novel species of papillomavirus in giraffe lesions using nanopore sequencing, Veterinary Microbiology, Volume 201. Pages 26-31, ISSN 0378-1135, https://doi.org/10.1016/j.vetmic.2016.12.035.

Visscher CF, et al. Serodiversity and serological as well as cultural distribution of Salmonella on farms and in abattoirs in Lower Saxony, Germany. International Journal of Food Microbiology 2011; 146: 44–51. doi: 10.1016/j.ijfoodmicro.2011.01.038

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). “Welcome to the tidyverse.” Journal of Open Source Software, 4(43), 1686. doi:10.21105/joss.01686

Woo PC, Lau SK, Teng JL, Tsang AK, Joseph M, Wong EY, Tang Y, Sivakumar S, Bai R, Wernery R, Wernery U, Yuen KY. 2014. Metagenomic analysis of viromes of dromedary camel fecal samples reveals large number and high diversity of circoviruses and picobirnaviruses. Virology. Dec;471-473:117-25. doi: 10.1016/j.virol.2014.09.020. Epub 2014 Oct 29. PMID: 25461537; PMCID: PMC7112128

Xia Y, Sun J. 2017. Hypothesis Testing and Statistical Analysis of Microbiome. Genes Dis. Sep;4(3):138-148. doi: 10.1016/j.gendis.2017.06.001. Epub 2017 Jun 23. PMID: 30197908; PMCID: PMC6128532.

Xu C, Weese JS, Flemming C, Odumeru J, Warriner K. Fate of Clostridium difficile during wastewater treatment and incidence in Southern Ontario watersheds. J Appl Microbiol. 2014 Sep;117(3):891-904. doi: 10.1111/jam.12575. Epub 2014 Jul 7. Erratum in: J Appl Microbiol. 2014 Oct:117(4):1219. PMID: 24930867.

Yang Y, Liu Z, Xing S, Liao X. 2019. The correlation between antibiotic resistance gene abundance and microbial community resistance in pig farm wastewater and surrounding rivers. Ecotoxicol Environ Saf. doi: 10.1016/j.ecoenv.2019.109452

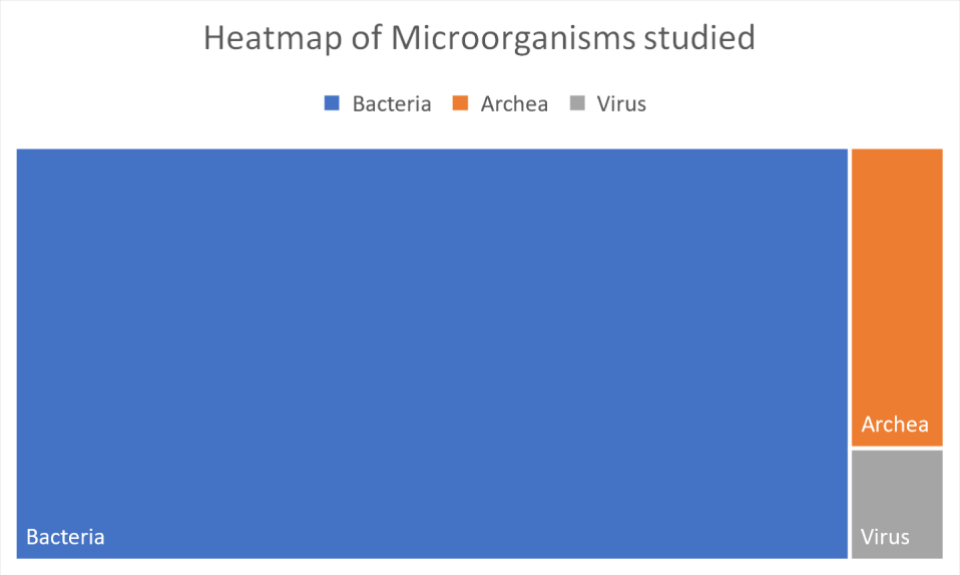
Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. 2021. Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol. Nov;39(11):1348-1365. doi: 10.1038/s41587-021-01108-x. Epub 2021 Nov 8. PMID: 34750572; PMCID: PMC8988251.

Zhang, M. Liu, YS. Zhao, JL. Liu, WR. He, LY. Zhang, JN. Chen, J. He, LK. Zhang, Q. Ying, GG. 2018. Occurrence, fate and mass loadings of antibiotics in two swine wastewater treatment systems, Science of The Total Environment, Volume 639, Pages 1421-1431, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2018.05.230.

Zorman T, Heyndrickx M, Uzunović-Kamberović S, Smole Mozina S. Genotyping of Campylobacter coli and C. jejuni from retail chicken meat and humans with campylobacteriosis in Slovenia and Bosnia and Herzegovina. Int J Food Microbiol. 2006 Jul 1;110(1):24-33. doi: 10.1016/j.ijfoodmicro.2006.03.001. Epub 2006 May 19. PMID: 16712997.

Zhou, L.J., Ying, G.G., Liu, S., Zhang, R.Q., Lai, H.J., Chen, Z.F., Pan, C.G., 2013b. Excretion masses and environmental occurrence of antibiotics in typical swine and dairy cattle farms in China. Sci. Total Environ. 444, 183–195.

Annex 1. Heatmap of the total microorganisms studied from the rectal swab, Moore swab and immersion bottle. Relative abundance was multiplied by 1000 for visualization purposes.



Annex 2. Table of total reads by pool proportioned by EPI2ME.

Pool	Type of sampling	Total reads	Reads analyzed	Reads classified	Reads unclassified	Proportion of unclassified reads (%)
pool 1	**Rectal swab**	176080	173708	103677	70031	40.315
pool2	**Rectal swab**	194754	185380	98481	86899	46.876
pool3	**Rectal swab**	115629	115333	64911	50422	43.719
pool4	**Rectal swab**	133704	131844	74060	57784	43.828
pool5	**Moore swab**	143998	143574	125304	18270	12.725
pool6	**Moore swab**	236155	235715	191367	44348	18.814
pool7	**Rectal swab**	198143	194441	104555	89886	46.228
pool8	**Rectal swab**	99613	98129	51751	46378	47.262
pool9	**Moore swab**	154723	154030	101213	52817	34.290
pool10	**Moore swab**	270022	253112	181286	71826	28.377
pool11	**Immersion bottle**	134590	134361	93359	41002	30.516
pool12	**Immersion bottle**	527606	522916	312469	210447	40.245
pool13	**Rectal swab**	452166	450632	270753	179879	39.917
pool14	**Rectal swab**	221879	221416	123901	97515	44.042
pool15	**Moore swab**	404267	403851	294508	109343	27.075
pool16	**Moore swab**	147236	147021	98318	48703	33.127
pool17	**Immersion bottle**	403828	403390	322251	81139	20.114
pool18	**Immersion bottle**	234412	234070	148089	85981	36.733
pool19	**Negative**	748	735	529	206	28.027
pool20	**Positive**	752903	748674	721577	27097	3.619

Annex 3. Table of the first and second most abundant species by type of sampling without the host depletion.

Pool	Type of sampling	Most abundant species	Cumulative reads	Cumulative reads proportion (%)	second most abundant species	Cumulative reads	Cumulative reads proportion %
pool 1	**Rectal swab**	*Homo sapiens*	68913	**66.469**	*E. Coli*	2361	2.277
pool2	**Rectal swab**	*Homo sapiens*	63131	**64.105**	*E. coli*	3367	3.419
pool3	**Rectal swab**	*Homo sapiens*	11275	**17.370**	*Faecalibacterium prausnitzii*	2260	3.482
pool4	**Rectal swab**	*Homo sapiens*	17845	**24.095**	*Lactobacillus amylovorus*	2495	3.369
pool5	**Moore swab**	*Comamonas thyoxidan*	22764	**18.167**	*Comamonas testosteroni*	8065	6.436
pool6	**Moore swab**	*Homosapiens*	42232	**22.069**	*Glaciesola amylolytica*	7941	4.150
pool7	**Rectal swab**	*Homo sapiens*	9300	**8.895**	*Faecalibacterium prausnitzii*	2741	2.622
pool8	**Rectal swab**	*Homo sapiens*	4731	**9.142**	*Faecalibacterium prausnitzii*	1164	2.249
pool9	**Moore swab**	*Acrobacter cryaerophilus*	25033	**24.733**	*Homo sapiens*	2026	2.002
pool10	**Moore swab**	*Acrobacter cryaerophilus*	64149	**35.386**	*Comamonas thyoxidans*	3740	2.063
pool11	**Immersion bottle**	*Arcobacter cryaerophilus*	28685	**30.725**	*Acinetobacter piscicola*	1741	1.865
pool12	**Immersion bottle**	*Arcobacter cryaerophilus*	81843	**26.192**	*Homo sapiens*	4589	1.469
pool13	**Rectal swab**	*Homo sapiens*	80088	**29.580**	*Faecalibacterium prausnitzii*	4081	1.507
pool14	**Rectal swab**	*Homo sapiens*	15858	**12.799**	*Faecalibacterium prausnitzii*	2419	1.952
pool15	**Moore swab**	*Arcobacter cryaerophilus*	116842	**39.674**	*Ruthinebacterium lactiforme*	3998	1.358
pool16	**Moore swab**	*Arcobacter cryaerophilus*	27353	**27.821**	*Homo sapiens*	1374	1.398
pool17	**Immersion bottle**	*Arcobacter cryaerophilus*	153357	**47.589**	*Aliarcobacter cibarius*	4146	1.287
pool18	**Immersion bottle**	*Arcobacter cryaerophilus*	22284	**15.048**	*Homo sapiens*	2413	1.629
pool19	**Negative**	*0*	0	**0.000**	*0*	0	0.000
pool20	**Positive**	*Bacillus*	1082	**0.150**	*0*	0	0.000

Annex 4. Distribution of the relative abundance (%) of taxa at the phylum level above 0.5% abundance cut-off.

Family	Third-day sampling			Fourth-day sampling		
	Rectal swab	Moore swab	Immersion bottle	Rectal swab	Moore swab	Immersion bottle
Lachnospiraceae	12.39	4.08	3.65	12.44	4.64	2.65
Prevotellaceae	11.38	4.38	4.02	10.98	2.38	1.71
Clostridiaceae	6.14	6.88	4.81	6.40	3.19	2.64
Ruminococcaceae	5.73	1.85	1.64	4.80	2.61	1.34
Bacteroidaceae	4.50	4.00	4.59	4.42	3.54	2.45
Peptostreptococcacea	2.28	0.54	0.53	2.90	0.00	0.00
Spirochaetaceae	1.39	0.00	0.00	1.71	0.00	0.00
Oscillospiraceae	1.93	1.80	1.67	1.65	1.59	1.05
Bacillaceae	1.49	1.26	1.22	1.64	1.08	0.78
Flavobacteriaceae	1.03	1.54	1.93	1.21	1.59	1.21
Paenibacillaceae	1.13	0.00	0.00	1.15	0.51	0.00
Lactobacillaceae	1.31	1.35	1.69	1.11	0.67	0.00
Campylobacteraceae	0.92	27.18	33.36	1.03	46.07	55.87
Weeksellaceae	0.99	0.83	0.96	0.99	0.73	0.57
Streptococcaceae	0.85	0.62	0.71	0.94	0.54	0.00
Selenomonadaceae	1.09	0.00	0.00	0.93	0.00	0.00
Rikenellaceae	0.88	0.00	0.00	0.86	0.00	0.00
Enterobacteriaceae	0.79	0.65	0.57	0.82	0.00	0.00
Pseudomonadaceae	0.81	1.08	0.78	0.79	0.00	0.00
Turicibacteraceae	0.75	0.00	0.00	0.66	0.00	0.00
Barnesiellaceae	0.69	0.00	0.00	0.65	0.00	0.00
Erysipelotrichaceae	0.61	0.00	0.00	0.65	0.00	0.00
Veillonellaceae	0.61	0.00	0.00	0.65	0.00	0.00
Muribaculaceae	0.60	0.00	0.00	0.62	0.00	0.00
Xanthomonadaceae	0.00	0.00	0.00	0.60	0.00	0.67
Tannerellaceae	0.61	0.00	0.00	0.60	0.00	0.00
Hymenobacteraceae	0.62	0.00	0.00	0.59	0.00	0.00
Staphylococcaceae	0.00	0.00	0.00	0.58	0.00	0.00
Streptomycetaceae	0.61	0.00	0.00	0.54	0.00	0.00
Burkholderiaceae	0.54	0.00	0.00	0.51	0.00	0.00
Moraxellaceae	0.00	5.48	7.89	0.00	5.21	4.93
Comamonadaceae	0.00	8.69	4.52	0.00	1.80	3.93
Desulfovibrionaceae	0.00	2.30	1.31	0.00	1.07	1.61
Dysgonomonadaceae	0.00	0.87	1.00	0.00	0.71	0.00
Eggerthellaceae	0.00	0.00	0.00	0.00	0.56	0.00
Sphingobacteriaceae	0.00	0.60	0.70	0.00	0.53	0.00
Peptococcaceae	0.00	0.50	0.54	0.00	0.53	0.00
Acidaminococcaceae	0.00	0.51	0.55	0.00	0.00	0.00

Annex 5. Distribution of the relative abundance (%) of taxa at the genus level above 0.5% abundance cut-off.

Genus	Third-day sampling			Fourth-day sampling		
	Rectal swab	Moore swab	Immersion bottle	Rectal swab	Moore swab	Immersion bottle
Prevotella	10.619	4.114	0.000	10.182	2.271	1.631
Clostridium	5.743	6.445	4.635	7.933	3.434	2.461
Bacteroides	4.670	4.004	1.066	4.579	3.606	2.497
Blautia	3.578	0.909	0.000	3.547	1.115	0.670
Faecalibacterium	3.542	0.522	0.000	2.661	0.000	0.000
Roseburia	2.682	0.613	0.000	2.634	0.739	0.000
Bacteroidales	1.460	0.000	0.000	1.425	0.000	0.000
Treponema	1.284	0.000	0.000	1.580	0.000	0.000
Phocaeicola	1.092	0.000	0.000	1.047	0.000	0.000
Oscillospiraceae	1.062	0.000	0.000	0.915	0.000	0.000
Flavonifractor	0.970	0.511	0.000	0.852	0.000	0.000
Selenomonas	0.957	0.000	0.000	0.760	0.000	0.000
Paenibacillus	0.934	0.000	0.000	0.951	0.000	0.000
Alistipes	0.919	0.000	0.000	0.891	0.000	0.000
Anaerobutyricum	0.882	0.000	0.000	0.826	0.000	0.000
Romboutsia	0.876	0.000	0.000	1.150	0.000	0.000
Bacillus	0.870	0.635	0.649	0.985	0.570	0.000
Clostridioides	0.849	0.000	0.000	1.020	0.000	0.000
Lactobacillus	0.845	0.000	0.000	0.668	0.000	0.000
Proteobacteria	0.817	0.000	0.563	0.726	0.000	0.000
Ruminococcus	0.815	0.000	0.000	0.875	0.000	0.000
Streptococcus	0.789	0.000	0.000	0.876	0.000	0.000
Lachnoclostridium	0.784	0.000	0.000	0.763	0.000	0.000
Turicibacter	0.779	0.000	0.000	0.685	0.000	0.000
Flintibacter	0.778	0.000	0.000	0.633	0.000	0.000
Firmicutes	0.766	0.000	0.000	0.795	0.000	0.000
Pseudomonas	0.755	0.996	5.869	0.763	0.000	0.000
Anaerostipes	0.735	0.000	0.000	0.696	0.000	0.000
Oscillibacter	0.678	1.385	0.000	0.561	1.299	0.832
Campylobacter	0.677	0.000	0.000	0.718	0.000	0.000
Clostridiales	0.660	0.000	0.000	0.666	0.000	0.000
Chryseobacterium	0.633	0.000	0.000	0.666	0.000	0.000
Streptomyces	0.587	0.000	0.000	0.509	0.000	0.000
Alloprevotella	0.580	0.000	0.000	0.624	0.000	0.000
Eubacterium	0.576	0.000	0.000	0.579	0.000	0.000
Terrabacteria	0.573	0.000	0.000	0.582	0.000	0.000
Paraprevotella	0.537	0.000	0.000	0.000	0.000	0.000
Candidatus	0.529	0.000	0.000	0.552	0.000	0.000
Arcobacter	0.000	25.726	36.010	0.000	45.214	54.155
Acinetobacter	0.000	5.344	17.961	0.000	5.217	4.933
Comamonas	0.000	3.591	8.057	0.000	0.884	1.937
Desulfovibrio	0.000	2.186	0.000	0.000	1.015	1.543
Acidovorax	0.000	1.891	5.140	0.000	0.000	0.774
Delftia	0.000	0.513	0.720	0.000	0.000	0.000
Aeromonas	0.000	0.000	0.774	0.000	0.000	0.000
Aliarcobacter	0.000	0.000	0.000	0.000	0.820	0.952
Alicycliphilus	0.000	0.000	0.718	0.000	0.000	0.000
Cloacibacterium	0.000	0.000	1.018	0.000	0.000	0.000
Diaphorobacter	0.000	0.000	0.698	0.000	0.000	0.000
Enterocloster	0.000	0.000	0.000	0.504	0.000	0.000
Glaciecola	0.000	0.000	5.372	0.000	0.000	0.000
Lactococcus	0.000	0.000	1.791	0.000	0.000	0.000
Lysinibacillus	0.000	0.000	2.043	0.000	0.000	0.000
Melaminivora	0.000	0.000	4.691	0.000	0.000	0.000
Pulveribacter	0.000	0.000	1.820	0.000	0.000	0.000
Rheinheimera	0.000	0.000	1.461	0.000	0.000	0.000
Ruthenibacterium	0.000	0.000	0.000	0.000	1.394	0.000
Stenotrophomonas	0.000	0.000	0.658	0.000	0.000	0.000
Variovorax	0.000	0.000	0.555	0.000	0.000	0.000

Annex 6. Distribution of the relative abundance (%) by specie level above 0.2% abundance cut-off.

ESPECIE	Third-day sampling						Fourth-day sampling					
	Rectal swab A	Rectal swab B	Moore swab A	Moore swab B	Immersion bottle A	Immersion bottle B	Rectal swab A	Rectal swab B	Moore swab A	Moore swab B	Immersion bottle A	Immersion bottle B
Faecalibacterium prausnitzii	2.987	2.573	0.429	0.459	0.261	0.399	2.223	2.332	0.299	0.394	0.203	0.659
Prevotella ruminicola	1.885	1.127	0.253	0.292	0.232	0.255	1.698	1.106	0.151	0.187	0.115	0.416
Roseburia intestinalis	1.509	1.103	0.402	0.263	0.294	0.314	1.562	1.063	0.431	0.457	0.242	0.437
Prevotella intermedia	1.424	0.920	0.587	0.444	0.522	0.501	1.427	0.941	0.329	0.445	0.246	0.730
Prevotella melaninogenica	1.419	0.827	0.593	0.394	0.496	0.504	1.299	0.825	0.296	0.444	0.233	0.685
Prevotella dentalis	1.213	0.889	0.236	0.239	0.206	0.228	1.232	0.798	0.129	0.188	0.084	0.312
Roseburia hominis	1.114	0.831	0.183	0.167	0.136	0.181	1.029	0.689	0.282	0.253	0.143	0.228
Flavonifractor plautii	0.963	1.200	0.511	0.540	0.396	0.526	0.852	1.243	0.394	0.466	0.281	0.671
Blautia producta	0.950	0.931	0.419	0.315	0.357	0.434	0.950	1.076	0.472	0.513	0.286	0.601
Anaerobutyricum hallii	0.876	0.767	0.116	0.107	0.099	0.139	0.826	0.809	0.138	0.168	0.084	0.243
Bacteroides fragilis	0.840	0.690	0.595	0.633	0.781	0.889	0.832	0.683	0.576	0.764	0.378	0.763
Clostridioides difficile	0.835	1.061	0.213	0.227	0.212	0.258	1.005	1.132	0.172	0.268	0.126	0.446
Prevotella fusca	0.812	0.539	0.238	0.172	0.197	0.181	0.753	0.478	0.117	0.160	0.091	0.285
Prevotella oris	0.771	0.495	1.116	0.756	1.129	0.986	0.789	0.465	0.634	0.851	0.426	1.288
Prevotella denticola	0.660	0.480	0.134	0.152	0.114	0.124	0.687	0.468	0.076	0.102	0.054	0.197
Prevotella jejuni	0.635	0.374	0.217	0.175	0.213	0.200	0.602	0.407	0.126	0.192	0.088	0.276
Clostridium butyricum	0.575	0.825	0.403	0.262	0.286	0.261	0.608	0.736	0.192	0.249	0.149	0.421
Clostridium chauvoei	0.555	0.661	0.260	0.220	0.189	0.196	0.500	0.726	0.108	0.177	0.088	0.376
Clostridium perfringens	0.495	0.597	0.204	0.201	0.170	0.160	0.523	0.723	0.107	0.152	0.101	0.404
Enterocloster bolteae	0.491	0.500	0.178	0.156	0.191	0.222	0.504	0.517	0.378	0.396	0.168	0.310
Prevotella enoeca	0.453	0.305	0.173	0.136	0.165	0.142	0.425	0.314	0.099	0.157	0.067	0.229
Bacteroides thetaiotaomicron	0.448	0.354	0.393	0.369	0.448	0.482	0.442	0.300	0.307	0.466	0.229	0.468
Phocaeicola dorei	0.403	0.356	0.134	0.151	0.156	0.177	0.383	0.319	0.129	0.185	0.114	0.362
Clostridium botulinum	0.357	0.411	0.372	0.218	0.294	0.266	0.369	0.483	0.177	0.247	0.140	0.348
Arcobacter cryaerophilus	0.039	0.091	23.106	33.555	29.303	24.433	0.062	0.073	32.293	28.854	49.513	12.365
Acinetobacter piscicola	0.000	0.000	1.075	0.660	1.935	0.669	0.003	0.009	1.285	0.368	0.639	0.117
Aliarcobacter faecis	0.007	0.000	0.451	0.796	0.515	0.632	0.007	0.005	0.820	0.764	0.951	0.458
Arcobacter cibarius	0.008	0.000	0.554	0.851	0.694	0.558	0.014	0.013	1.056	0.817	1.313	0.370
Acinetobacter wuhouensis	0.009	0.015	0.777	0.543	1.495	0.502	0.004	0.006	0.943	0.270	0.547	0.099
Arcobacter skirrowii	0.000	0.009	0.425	0.567	0.452	0.426	0.004	0.006	0.617	0.478	0.764	0.232
Bacteroides fragilis	0.840	0.690	0.595	0.633	0.781	0.889	0.832	0.683	0.576	0.764	0.378	0.763
Clostridium butyricum	0.575	0.825	0.403	0.262	0.286	0.261	0.608	0.736	0.192	0.249	0.149	0.421
Bacteroides fragilis	0.840	0.690	0.595	0.633	0.781	0.889	0.832	0.683	0.576	0.764	0.378	0.763
Chryseobacterium gallinarum	0.313	0.232	0.044	0.051	0.029	0.049	0.334	0.252	0.030	0.039	0.025	0.073
Paeniclostridium sordellii	0.306	0.396	0.050	0.062	0.055	0.064	0.398	0.419	0.032	0.057	0.021	0.131
Gammaproteobacteria	0.301	0.305	0.203	0.168	0.140	0.184	0.256	0.299	0.127	0.177	0.114	0.215
Muribaculum intestinale	0.288	0.290	0.054	0.074	0.039	0.055	0.300	0.256	0.031	0.035	0.025	0.105
Collinsella aerofaciens	0.285	0.245	0.021	0.035	0.030	0.052	0.205	0.264	0.034	0.031	0.020	0.055
Clostridium septicum	0.279	0.343	0.126	0.112	0.095	0.100	0.269	0.405	0.056	0.064	0.047	0.211
Bacteroides zoogleoformans	0.263	0.230	0.060	0.074	0.060	0.073	0.260	0.225	0.053	0.088	0.048	0.106
Dysosmobacter welbionis	0.262	0.259	0.057	0.058	0.049	0.065	0.228	0.258	0.052	0.069	0.042	0.083
Lactobacillus johnsonii	0.253	0.351	0.018	0.027	0.019	0.039	0.187	0.662	0.008	0.030	0.012	0.042
Ruminococcus bicirculans	0.250	0.349	0.054	0.064	0.041	0.051	0.278	0.304	0.046	0.052	0.029	0.110
Clostridium isatidis	0.247	0.263	0.117	0.091	0.074	0.074	0.230	0.317	0.039	0.066	0.038	0.165
Massilistercora timonensis	0.246	0.274	0.053	0.059	0.059	0.070	0.229	0.260	0.068	0.081	0.044	0.109
Clostridium bornimense	0.244	0.305	0.120	0.086	0.101	0.085	0.290	0.329	0.065	0.093	0.043	0.172
Succinivibrio dextrinosolvens	0.229	0.420	0.057	0.077	0.038	0.075	0.225	0.253	0.032	0.052	0.023	0.159
Peptacetobacter hiranonis	0.227	0.283	0.046	0.044	0.042	0.042	0.278	0.300	0.032	0.041	0.029	0.109
Parabacteroides distasonis	0.226	0.199	0.128	0.184	0.166	0.174	0.202	0.203	0.099	0.155	0.077	0.209
Ruminococcus albus 7	0.222	0.225	0.038	0.044	0.038	0.044	0.233	0.270	0.038	0.062	0.023	0.088
Bacteroides caccae	0.222	0.161	0.178	0.184	0.237	0.241	0.208	0.184	0.146	0.209	0.116	0.275
Bacteroides ovatus	0.218	0.161	0.841	0.665	0.857	0.991	0.249	0.156	0.895	1.090	0.566	0.775
Clostridium baratii	0.208	0.329	0.136	0.100	0.112	0.099	0.227	0.320	0.055	0.078	0.041	0.184
Bacteroides heparinolyticus	0.206	0.192	0.086	0.092	0.109	0.093	0.220	0.194	0.056	0.075	0.050	0.122
Butyrivibrio proteoclasticus	0.206	0.190	0.070	0.064	0.053	0.062	0.227	0.216	0.053	0.073	0.037	0.109
Anaerostipes caccae	0.206	0.181	0.070	0.052	0.071	0.072	0.209	0.177	0.098	0.082	0.049	0.121
Butyricimonas faecalis	0.206	0.148	0.168	0.157	0.222	0.304	0.178	0.145	0.180	0.276	0.090	0.183
Bacteroides helcogenes	0.204	0.139	0.085	0.105	0.106	0.119	0.182	0.129	0.075	0.119	0.060	0.165
Prevotella denticola	0.200	0.126	0.040	0.025	0.007	0.018	0.193	0.116	0.010	0.015	0.009	0.036