Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE E TECNOLOGIE AGRARIE, AMBIENTALI E ALIMENTARI

Ciclo 34

Settore Concorsuale: 07/H2 – PATOLOGIA VETERINARIA E ISPEZIONE DEGLI ALIMENTI DI ORIGINE ANIMALE

Settore Scientifico Disciplinare: VET/04 – ISPEZIONE DEGLI ALIMENTI DI ORIGINE ANIMALE

MICROBIOME AND RESISTOME OF FOODS OF ANIMAL ORIGIN

Presentata da: Chiara Oliveri

Coordinatore Dottorato

Prof. Massimiliano Petracci

Supervisore

Prof. Gerardo Manfreda

Co-supervisore

Prof.ssa Alessandra De Cesare

Esame finale anno 2022

Table of contents

Abstract

To ensure food safety and to prevent food-borne illnesses, rapid and accurate detection of pathogenic agents is essential. Public authorities are pushing the food industry to develop comprehensive quality management systems to improve food safety. Progress in next-generation sequencing (NGS) technologies have revolutionized the study of bacterial populations, understood those present in foods, supporting the search of methods to better food safety by assessing the presence of specific undesirable microorganisms and genes. It has already been demonstrated that whole-metagenome shotgun sequencing can be used to detect pathogens in food and has the potential to become a powerful tool in the field of modern food safety since it allows the detection, identification, and characterization of a broad range of pathogens and their antibiotic resistance genes. Considering these aspects, in the studies presented in this thesis, the application shotgun metagenomic sequencing has been applied to investigate both the microbiome and resistome of foods of animal origin in order to assess the advantages and disadvantages of shotgun metagenomic sequencing in comparison to the cultural methods used to verify food safety and map ecosystems associated to food systems from farm to fork. The specific food chains addressed are those detailed in the 4 research studies performed to reach the main objective.

In particular, in the first study, to contribute to assess the suitability of shotgun metagenomics to detect a wide range of target microorganisms in foods, we characterized and quantified microorganisms belonging to different domains experimentally spiked in cold-smoked salmon at known concentrations, using shotgun metagenomics. In addition, we compared the sequencing results using four bioinformatic tools, to evaluate the suitability to detect six species of bacteria, including potential food-borne pathogens, as well as *Cryptosporidium parvum*, *Saccharomyces cerevisiae*, and *Bovine alphaherpesvirus 1*. The results of this study showed that shotgun metagenomics can be applied to detect microorganisms belonging to different domains in the same food sample. Nevertheless, a direct correlation between cell concentration of each spiked microorganism and number of corresponding reads cannot be established yet.

In the second study, we investigated whether the efforts of raising chickens without the use of antibiotics make any difference in the microbiome of poultry meat consumers eat. To this aim, we compared the microbiomes characterizing caeca and the corresponding carcasses of two groups of chickens reared one in a conventional farm and one in an antibiotic free intensive farm. Moreover, with the view of planning future studies, we investigated whether checking the correlation between the microbiome and resistome in the caeca and the carcass of the same animal provides more insight than the same analysis performed at flock level. The results showed a clear separation between the taxonomic, functional, and antibiotic resistant genes in the caeca of the birds reared in the

conventional and antibiotic free farms. However, that separation was completely lost on carcasses. The antibiotic free production resulted in statistically significant lower antimicrobial resistance load in the caeca of chickens in comparison to the conventional production. Moreover, the antimicrobial resistance load on carcasses was much higher than in the caeca, without any significant difference between carcasses coming from the two types of farms. All in all, the results of this research highlighted the need to reduce sources of microbial contamination and antimicrobial resistance not only at the farm level but also at the post-harvest one.

In the third study, additional samples of poultry carcasses reared in antibiotic free and conventional flocks were tested, confirming the difference between the three tested groups with are still under investigation for possible correlations with feeding and environmental covariates.

In the fourth study, we started the shotgun metagenomic investigation of an Italian fermented artisanal product, addressing the question if testing one or three aliquots of artisanal food homogenate is representative of the whole homogenate. The results clearly showed that the metagenomes obtained from replicates of the homogenate displayed overlapping taxonomic and functional composition. Therefore, shotgun metagenomics of a single aliquot of an artisanal fermented food is representative of the whole homogenate.

Altogether, the results presented in this thesis confirmed that the application of shotgun metagenomic sequencing represents a powerful tool that can be used in the identification of both spoilage and pathogenic microorganism, their resistome and associated set of functional genes. However, the full implementation of shotgun metagenomics in food safety and inspection of food of animal origin within a regulatory framework required a full standardization of the laboratory and bioinformatic parts. Moreover, a robust relationship between sequence read abundance and concentration of colony-forming unit must be still established.

1. Introduction

1.1 Food Safety investigations: from traditional techniques to metagenomic analysis

Food, an essential part of everyday life, undergoes many processing steps. The quality and safety of a food is largely influenced from relationship with its microbiome (De Filippis *et al*., 2018). The microbiome, play important roles in any food matrix ranging from fermentation, contamination, and spoilage. Deep taxonomic knowledge of the microorganisms and their communities is necessary either to promote desired food processes (i.e., fermentation), to better comprehend microbiological processes involved in food processing and ripening, and to improve microbiological safety by monitoring in situ pathogenic bacteria, hence, to limit damaging events (i.e., contamination and spoilage).

Further evolution was stimulated in the field of microbial ecology by the advent and development of metagenomics. Metagenomics is defined as study of genetic materials from environmental or host-associated microbiota to identify the microbial diversity and its functions (Choi *et al*., 2015) and thus gives a broader description than phylogenetic surveys, which are often based only on the diversity of one gene (i.e., 16S rRNA gene). Within this framework, metagenomics gives genetic information on potentially novel biocatalysts or enzymes, genomic linkages between function and phylogeny for uncultured organisms, and evolutionary profiles of community function and structure. In addition, applies a set of genomic technologies and bioinformatics tools to directly access the genetic content of entire communities of organisms (Thomas *et al*., 2012).

Microbiome outputs are completed with a certain amount of metadata, the more the better. A great potential of taxonomic studies is the option to correlate the abundance of microbial taxa with other variables, taking into account that the results should be considered carefully, as the correlative link is not always due to ecologically meaningful relationships. Nevertheless, the correlations that result statistically relevant can be very useful in assuming a hypothesis on the role of certain microbial species in the food. Correlation analysis between the abundance of microbial taxa and chemical determinations may highlight the possible species responsible for the production of metabolites important for the properties of the final products (Lattanzi *et al*., 2013; De Pasquale *et al*., 2014; De Filippis *et al*., 2017a; De Filippis *et al*., 2018).

Traditionally, conventional techniques including the Gram stain along with plate culture on selective growth media, and subsequently individual biochemical characterization came used for the identification and characterization of bacteria from clinical, food, or environmental origins. In food

systems, estimates of effective microbial diversity are still difficult because the majority of the food-associated microorganisms cannot be cultivated on standard laboratory media. However, these techniques are not feasible for all microorganisms (Amann *et al*., 1995). For example, the selective isolation of microorganisms requires sometimes unknown growth factors and/or growth conditions that cannot be reproduced in the laboratory. It is widely accepted that there are many more microorganisms than those obtainable in plates (Ercolini *et al*., 2001) due to the fact of the inability of detecting novel microorganisms (which might not be cultivable with known media), and of the inability of recovering known microorganisms which are viable but enter a non-cultivable state (Giraffa and Neviani, 2001). Culture-dependent techniques can only detect a little portion (0.1%) of a complex community, indeed, to expand the understanding of the ecological niche of the food, techniques are necessary to identify or characterize microorganisms and predict the functional interactions of different microbiological communities present in the sample. In the last few decades, many culture-independent methods have been developed to solve this limit and extended to the food system. These techniques allow the identification and, in some cases, the quantification of food microbial groups and offer rapid and sensitive methods for determining the composition and diversity of complex microbial communities (Mayo *et al*., 2014). Different methods like denaturing gradient gel electrophoresis (DGGE), temporal temperature gradient electrophoresis (TTGE), and real-time quantitative polymerase chain reaction (qPCR) have been used in food microbial ecology (Cocolin *et al*., 2013).

In this respect advances in multi-omic technologies have permitted the characterization of different microbial species in food matrices, microbial community profiling and monitoring population fluctuations in different microbial ecosystems. The increment of sequencing technologies, the ubiquitous nature, and specificity of nucleic acids over the past four decades have favored the capacity to characterize the microbiomes of matrices food or environment. DNA sequencing progressively evolved, through significant advancements in sequencing chemistries, from low throughput DNA fragment sequencing to high throughput next generation (i.e., NGS) and third-generation sequencing techniques (Loman and Pallen, 2015; Cao *et al*., 2017).

The first shotgun DNA sequencing strategy was the Sanger. The Sanger methodology dominated research and represented a significant breakthrough that permanently changed the way prokaryotes in the environment were analyzed, leading to the advent of the metagenomic analysis era trough sequencing techniques (Hiergeist *et al*., 2015). Sanger *et al*., (1977) introduced the concept of DNA sequencing named the "chain terminator method". This first-generation sequencing technology is focused on incorporation of fluorescently labeled dideoxyribonucleotide (ddNTP) and a primer into a PCR machine that set the stage for automated high-throughput DNA sequencing. With the

information generated from the last terminator base in the four individual base reaction tubes after size separation, the original sequence is determined (Diaz-Sanchez *et al*., 2013). At the same time, Sanger sequencing method presented a high number of disadvantages associated with both the low yield of the DNA sequences obtained and the high cost. Nevertheless, low throughput results and high operational costs, restricted the utilization of the method Sanger (Cao *et al*., 2017).

That said, other sequencing technologies were developed during the following years, including several next-generation sequencing (NGS) technologies (Metzker, 2010; Diaz-Sanchez *et al*., 2013). The study of microbial diversity can now be achieved by using high-throughput sequencing (HTS) approaches after direct nucleic acid extraction from the matrix to be studied. HTS assures higher sensitivity, enabling the detection of nondominant communities that play an important role in the studied ecosystem. The great advantage of these methods is the unprecedented potential for quantitative detection of the structure of microbial communities, in fact, the number of reads detected for a given organism is proportional to its abundance in the sample. Furthermore, HTS greatly reduced the price per base compared to Sanger sequencing (De Filippis *et al*., 2018).

The first commercial HTS platform, the 454 Pyrosequencer, was released in 2000. This technique was based on the combination of single-molecule emulsion PCR with pyrosequencing (shotgun sequencing procedure) (Schadt *et al*., 2010). DNA molecule is first sheared with enzymatic based digestion or sonication, and later ligated with oligonucleotide adapters. Each ligated fragment is attached bead, PCR amplified, and pyro sequenced (Ronaghi *et al*., 1996). The disadvantage of this technique is that prone to high error rates, that is to say, on occasion n nucleotides are read as n-1 nucleotides. This method was used for the first time in an investigation of microbial populations from water, marine, fresh water, fish, corals terrestrial animals (Dinsdale *et al*., 2008).

This way, the development of this technology resulted in advances in the next-generation devices, such as the second-generation (High-throughput next-generation sequencing, HT-NGS) platforms as marketed by Roche, Illumina-Solexa, Life Technologies, and Helicos. The second-generation NGS platforms methods are based on a parallel process in which each single DNA fragment is sequenced individually and separated in clonal amplicons for further analysis among the total sequences generated (Pareek *et al*., 2011). A diversity of sequencing platforms is available from several manufacturers, which vary in sequencing chemistry, read length, and/or throughput (Walsh *et al*., 2018). Presently, Illumina's range of sequencers (MiSeq, NextSeq 500, and the HiSeq series) and the Ion Torrent Personal Genome Machine (PGM) are the most commonly used sequencing platforms (Reuter *et al*., 2015). The Illumina and Ion Torrent sequencers use several sequencing chemistries but follow such principles. Illumina sequencing technology, recent success in its application to metagenomics, is based on reversible dye-terminators. Briefly, adaptor-ligated DNA fragments on

the surface of a flow cell are amplified by bridge PCR, generating clusters of identical DNA fragments. These are then sequenced using a sequencing-by-synthesis approach that involves cyclic rounds of single-base extension with a mixture of fluorescently labeled dNTPs and imaging to identify the incorporated base (Bentley *et al*., 2008). Because all dNTP are present as single, separate molecules, natural competition minimizes incorporation bias. After the incorporation of reversibly terminating nucleotides, a camera capture images of the fluorescence, and the dye along with the terminal 3′ blocker is chemically removed from the DNA allowing the next cycle. Continuous sequence information of nearly 300 bp can be obtained from two overlapping 150 bp paired-reads from a single insert, hence yields of ~60 Gbp can be expected in a single channel. The only limitation of Illumina technology is the limited read length. For this reason, a greater proportion of unassembled reads might be too short for functional annotation (Wommack *et al*., 2008). Although, some current software packages (i.e., MG-RAST) are capable of analyzing unassembled Illumina reads of 75 bp and longer (Thomas *et al*., 2012, Diaz-Sanchez *et al*., 2013; Walsh *et al*., 2017). The Illumina platforms are distinguished in their total output and maximum read length. The Illumina MiSeq is suitable for amplicon sequencing, faster runtime can be achieved, although has a higher error rate (Thomas *et al*., 2012; Salipante *et al*., 2014), can be useful to assess library concentrations, barcode pool balancing, and for sequencing a limited number of samples. Generates low volume of sequence data (15 Gb) and is more suited to whole metagenome shotgun sequencing and metatranscriptomics (high-throughput applications) (Reuter *et al*., 2015; Quince *et al*., 2017). The Illumina NextSeq 500 and the Illumina HiSeq 2500 generate a high volume of sequence data (120 Gb and 1.5 Tb, respectively) and are well suited for metagenomics studies, but NextSeq costs less than half the price of the HiSeq (Quince *et al*., 2017; Walsh *et al*., 2017).

High-throughput next-generation sequencing (HT-NGS) techniques have contributed to change the way to study food microbial ecology, leading to consider microbial populations as consortia (Cocolin and Ercolini, 2015). In recent years, high-throughput sequencing has yielded insights into microbial populations within different environments such as soil (Fierer *et al*., 2012), ocean (Frias-Lopez *et al*., 2008), human (Human Microbiome Project Consortium, 2012), including many foods and food production facilities, although the number of scientific publications on the subject is still rather limited compared to other ecosystems (Ercolini, 2013; Bokulich *et al*., 2016). For the moment, most of the HTS-based studies examined have centred on the monitoring of microbial populations during food fermentations (De Filippis *et al*., 2017b).

The advantage of shotgun sequencing firstly offers the possibility to monitor the abundance of microbial activities directly in the food matrix; secondly, to collect information on the genetic

capacity of the whole microbial community; thirdly, to recover complete or draft microbial genomes from the metagenomes, achieving a strain-level resolution (De Filippis *et al*., 2017b).

Different whole metagenome shotgun sequencing approaches are applied in food. Specifically, these methods can be used to detect and trace outbreaks of food-borne pathogens also through food production chains, study starter cultures and probiotics, and understand the microbial dynamics of food fermentations (Walsh *et al*., 2017). Another key topic for those who interface with the field of food microbiology is the understanding of the evolutionary dynamics of specific spoilage organisms (product spoilage). Most of the studies in the literature focus on fresh meat spoilage, monitoring of spoilage bacteria during storage (De Filippis *et al*., 2018). Beef carcasses included high microbial diversity and the well-known genera associated with meat spoilage (De Filippis *et al*., 2013), with differences due to slaughtering practices and occurring at different areas of the carcass (Korsak *et al*., 2017). De Filippis *et al*., (2013) suggested that bacteria originally present on the carcass colonize the butchery environment, where they become resident at low temperatures. This microbiota constitutes a primary contamination source for fresh meat (De Filippis *et al*., 2013). HTS has been widely used to track contamination sources in different types of food processing plants as in the cooked sausages (Hultman *et al*., 2015), in the salmon fillets (Møretrø *et al*., 2016), and in the fresh meat (Stellato *et al*., 2016). In all cases of application, a resident microbiota in the plant was highlighted and its importance as a primary contamination source was highlighted, as was the need to adopt adequate cleaning and hygiene practices in food handling environments (De Filippis *et al*., 2018). Whole metagenome shotgun (WMS) can also be employed to detect pathogens in food, as revealed by Leonard *et al*., (2015). In this regard, have applied this approach to detect *Escherichia coli* in fresh spinach. Yang *et al*., (2016) stated that WMS is useful for investigating the transmission of pathogens through food production chains (food-borne pathogens). Was used to investigate how food processing affected the microbial composition of beef. Although processing reduced the total number of bacteria in the meat, it was noted that it resulted in an increase in the relative abundance of *Salmonella enterica*, *Escherichia coli*, and *Clostridium botulinum*, potentially because of their ability to survive antimicrobial interventions. Thus, WMS can be used to identify the control points in the food production chain that best reduce contamination by food-borne pathogens (Walsh *et al*., 2017). Nevertheless, beyond simply cataloguing the microorganisms the WMS approaches also elucidate their roles (Hanage, 2014). Several studies have demonstrated that WMS can identify the microorganisms that are most important during fermentation and enhance the qualities of fermented foods. For example, Illeghems *et al*., (2015) detected in a cocoa bean fermentation sample that genes associated with carbohydrate catabolism (specially heterolactic fermentation and pyruvate metabolism) were enriched in *Lactobacillaceae* while genes associated with pectinolysis, and citrate

metabolism were detected in *Enterobacteriaceae*. This implies that these bacteria might contribute to degradation of cocoa pulp and flavour formation. Similar, WMS can provide insights into the role of specific microorganisms in flavour production during fermentation and cheese ripening. Monitoring microbial gene expression during ripening of a surface-ripened cheese (Dugat-Bony *et al*., 2015) revealed the presence of microbial genes involved in flavour production from amino acids and highlighted that these activities were enhanced in the first phase of ripening. In another study, Wolfe *et al*., (2014), conducted an analysis of bloomy-rind, natural-rind, and washed-rind cheeses microbial communities. These included cysteine and methionine metabolism pathways (associated with the production of volatile sulfur compounds) and valine, leucine, and isoleucine degradation pathways (associated with putrid and sweaty aromas). Furthermore, in the same study genes encoding lipases, proteases, and methionine-γ-lyase (important enzyme in the production of sulfur compounds) were identified in *Pseudoalteromonas*, suggesting that this genus is involved in flavour production in cheese. These studies provided an in-depth analysis of the cheese maturation process and allowed us to better understand the metabolic activities of the different community members and their possible interactions (De Filippis *et al*., 2017b). WMS approach it can also be employed to identify, microbes associated with defects. For example, was used to determine the causal agent of a pinking defect in cheeses. Quigley *et al*., (2016) have discovered that *Thermus thermophilus* (which had not previously been associated with the cheese microbiota) was enriched in defect cheeses and hence associated genes involved in carotenoid production were enriched in these samples. This implies, that this approach could be employed to identify the causes of other defects in cheeses and eventually to inform control measures to prevent such defects.

The application of HT-NGS sequencing is emerging and moving toward the development and the improvement of the poultry industry, improving animal production, increasing the food safety measures and preventing food-borne pathogens. Multiple studies have focused on the sequencing of food production animals including chickens. HT-NGS techniques have permitted the research diversity and functions of microbiota from the gastrointestinal tract (GIT) of various livestock animals creating great volumes sequence data comprising genetic information (Metzker, 2010). For this reason, allows hypothesis-driven research on chicken GIT microbiota, thereby highlighting the roles of previously unknown and rare microbial GIT species (Medinger *et al*., 2010; Diaz-Sanchez *et al*., 2013; Choi *et al*., 2015).

1.2 Historical overview and definitions of the term's microbiome and microbiota

Etymologically the term microbiome comes from ancient Greek, "micro" (small) and "biome" is composed of the word bíos (life) and modified by the ending "ome", while the term microbiota is a

combo of “micro”, (small), with the term “biota”, which means the living organisms of an ecosystem or a particular area. Microbial communities were termed as the collection of microorganisms living together (multi-species assemblages) that interact with each other in a bordering environment (Konopka, 2009). The variety of perspectives on the term microbiome was a discussion for many years. The currently most cited definition in various papers stated that the term microbiome and microbiota were coined by Nobel laureate-microbiologist Joshua Lederberg in 2001 (Podolsky, 2017) in which it is defined microbiomes (within an ecological context) as a community of commensal, symbiotic, and pathogenic microorganisms within the environment. Actually, holding the search to pre-2001, the term microbiota is a term that is use from at least 50 years in basic microbiology (Lane-Petter, 1962) while, the term microbiome was used for defining a very small ecological niche incorporating plant and animal life. Particularly, in 1988 Whipps, Lewis and Cooke, working on the ecology of microorganisms provided the first definition of the term microbiome: “A convenient ecological framework in which to examine biocontrol systems is that of the microbiome. This may be defined as a characteristic microbial community in a reasonably well-defined habitat which has distinct physio-chemical properties. The term thus not only refers to the microorganisms involved but also encompasses their theatre of activity”. This notion represents an advancement of the definition of the microbial community, considering that it defines a microbial community with distinct properties and functions and its interactions with its environment, which implies the formation of specific ecological niches. Nevertheless, there are many other microbiome definitions that have been published over the years. Marchesi and Ravel, 2015 focused on their definition of microbiota as the set of microorganisms present in a defined environment while the term microbiome refers to the entire habitat, including the microorganisms' bacteria, archaea, fungi, protozoa, and viruses, their genomes (i.e., genes/microbial gene expression), and the nearby environmental conditions and its prevailing biotic and abiotic conditions. Sender *et al*., (2016) and Roto *et al*., (2015), have described the microbiota as a microbial community composed of commensal, symbiotic and pathogenic microorganisms, which usually colonize an area of human and animal (live in an environment such as the intestine) and are around 2 times more plentiful than somatic and germinal cells of the host. Hence, collective genomes and gene, and gene products (host and microbiota metabolites and proteins) of these symbionts is known as the microbiome. Nevertheless, when discussing the microbiota in the chicken intestine one is refers to the bacterial population (Marchesi and Ravel, 2015; Prescott, 2017; Clavijo and Flòrez, 2018; Kogut, 2019; Berg *et al*., 2020).

The microbiota includes all living members constitute the microbiome. Microbiome, defined by Whipps *et al*., (1988) includes the concept “theatre of activity”, involving the of molecules produced by the microorganisms, including their nucleic acids, proteins, lipids, polysaccharides (structural

elements), organic and inorganic molecules, signaling molecules and toxins (metabolites) and molecules generated by coexisting hosts and structured by the neighboring environmental conditions. Several definitions limit the definition of microbiome as encompassing the genes and genomes of microorganisms only. It is claimed that this is the definition of metagenome, which combined with the environment constitutes the microbiome. Metagenome and microbiome are often used interchangeably, but the metagenome is defined as a collection of genomes of microbial populations, the genes within the microbial populations and includes the plasmids within the different bacterial population from the members of a microbiota (Roto *et al*., 2015). In contrast to the microbiota, which can be studied separately, the microbiome is always composed by all members living, which interact with each other, live in the same habitat, and form their ecological niche together. The microbiome is characterized by the application of multi-omics technologies (metagenomics, metatranscriptomics, metaproteomics, or metabolomics approaches) combined with environmental metadata provide detailed information on microbial activities in the environment (Marchesi and Ravel, 2015; Berg *et al*., 2020).

Historically, of microbiome research has emerged from environmental microbiome study (microbial ecology) and has evolved rapidly over the past few decades providing an interdisciplinary platform for agriculture, food science, biotechnology, and human medicine. This rapid evolution of microbiome research covering different fields, faces a variety of challenges due to lack of a clear or agreed of the vocabulary used to describe the term microbiome, and of lacking consensus on best practices in microbiome research is missing. Research progress has been driven by the development of new techniques and equipment. From the development of the first microscopes (that allowed the discovery of a new and unknown ecosystem and the identification of microorganisms), shifting the focus of the research community on the role of microorganisms as disease-forming agents that needed to be eliminated. Nevertheless, extensive research has shown that only a small proportion of microorganisms are associated with diseases, differently, the majority of microorganisms were known for beneficial interactions with other well microorganisms hence essential for ecosystem functioning. Subsequently, find of DNA, PCR, cloning techniques and the development of sequencing technologies, allowed the investigation of microbial communities using cultivation-independent, DNA-based and RNA-based approaches (Brul *et al*., 2008). New sequencing technologies and sequence data showed both the critical roles of microorganisms in human and animals and the ubiquity of microbial communities in association within higher organisms (Lozupone *et al*., 2012).These new potential (at the beginning of this century and continues through today), have reshaped microbial ecology, because the analysis of genomes and metagenomes (in high-throughput) manner provides powerful methods for addressing the functional potential of individual

microorganisms and of whole communities in their natural habitats (Venter *et al*., 2004; Liu *et al*., 2012; Berg *et al*., 2020).

1.3 Whole shotgun metagenomic sequencing of food products

Whole shotgun metagenomic sequencing is an alternative approach that seeks to avert all 16S rRNA gene metataxonomic sequencing limitations. Shotgun metagenomics consists in the sequencing of bacterial DNA isolated from the whole microbial community (after fragmentation and library preparation), without any prior PCR step, avoiding the possibility of amplification biases (De Filippis *et al*., 2017b; Laudadio *et al*.,2018). DNA molecules are randomly broken by enzymatic or mechanical methods into tiny fragments that are then independently sequenced (Sharpton, 2014; Durazzi *et al*., 2021). The final output represents the metagenome of the microbial populations present in the sample. The sequencing reads obtained are aligned to various genomic locations for the myriad genomes present in the sample, including non-microbes or microbes with unknown taxonomically informative genetic markers. The DNA sequences are sampled from taxonomically informative genomic loci (e.g.,16S rRNA gene) while others from coding sequences that provide insight into the biological functions encoded in the genome (Sharpton, 2014). As a result, metagenomic data deliver knowledge on mapping of the taxonomic composition of the ecosystem under study, allowing a more accurate resolution at the species level and potentially strain-level. Moreover, it is possible to tracking and comparing the abundance of microbial activities directly in the food matrix, and hence to collect information on the genic contribution of the whole microbial community in terms of functional genes (De Filippis *et al*., 2017b; Truong *et al*., 2017; Laudadio *et al*., 2018; Durazzi *et al*., 2021). But even so, metagenomic sequence data include several challenges. Metagenomic data, other than being relatively complex, and large, entail computational problems in informatics analysis. Furthermore, it can be difficult to determine the genome from which a read belongs. Indeed, communities are so different that genomes are not fully represented by reads and two reads from the same gene may not overlap and are thus unimaginable directly compare through sequence alignment (Schloss and Handelsman, 2008; Sharpton *et al*., 2011). Once the reads overlap, it does not mean that they belong to the same genome: in fact, when there is a correspondence between reads, they could also derive from different genomes that have modified their assembly during sequencing (Mavromatis *et al*., 2007; Mende *et al*., 2012). Furthermore, metagenomic analysis provide a large volume of data to obtain significant results, entail computational problems in information processing. To help researchers, the development and progress of informatics software is very rapid and allows to improve the efficiency of metagenomic analysis (Sharpton *et al*., 2014). Bioinformatics is in fact a new discipline that deals with the development and integration of information science applications at the

service of research in the biotechnology field. Future explorations of this technique will allow us to study how the microbiota of food modulates in relation to the different conditions of storage temperature, heat treatment, pH, and aw, all thanks to the data obtained from sequencing. Another challenge is characterized by the fact that DNA extracted many times contains unwanted DNA inside. Sometimes, this DNA can exceed the one of interest and this makes the method to be used more complex (enriching the microbial DNA before to sequencing). Bioinformatics software and molecular methods is able to filter this DNA from the metagenomic one of interest (Woyke *et al*., 2006; Chew and Holmes, 2009; Delmotte *et al*., 2009; Schmieder and Edwards, 2011a; Garcia-Garcerà *et al*., 2013). Finally, another challenge of metagenomic analysis that should not be underestimated is represented by the eventual contamination of the metagenomic sequence. Once the genetic material is sequenced, is difficult to determine which reads belong contaminant's genome. This contamination could misdirect analyses of community genetic diversity if the contaminant's genome is enriched of genes that are uncommon in the community (above all when the contaminant is abundant) or has a large genome. Nevertheless, other than limiting the contamination by applying good sampling and DNA extraction practices, exist bioinformatic software allows identification and filter of contaminant sequences in the metagenomic data (Schmieder and Edwards, 2011b). Whole shotgun metagenomic sequencing, thanks to the large amount of data generated, has become the largely used method in laboratories, even because the limits being overcome by the development in bioinformatics. In recent years, metagenomic sequencing has been utilized to determine new viruses, reveal novel and ecologically important proteins, determine taxa and metabolic pathways that differentiate gut microbiota, and characterize the genomic diversity and function of uncultured bacteria (Godzik, 2011; Yozwiak *et al*., 2012; Wrighton *et al*., 2012; Sharpton, 2014). Generally, a classic shotgun metagenomic study contains five phases:

1. the sampling, processing and sequencing of the samples
2. preprocessing of the sequencing reads
3. sequence analysis to profile taxonomic, functional and genomic features of the microbiome
4. statistical and biological post-processing analysis
5. validation.

To carry out each phase, researchers can interface with the choice of different experimental and computational approaches (Qiunce *et al*., 2017).

1.3.1 Sampling and processing

Collection and storage methods are the first and decisive step in any metagenomics project. The DNA extracted should be representative of all cells present in the sample and in the same way sufficient amounts of high-quality nucleic acids must be obtained for successive library production and sequencing. Key objectives of this phase are the collection of sufficient microbial biomass for sequencing and to minimize contamination of samples. Sample collection and preservation protocols should specifically protocol for each sample type, due to the fact can affect both quality and accuracy of metagenomics data. Collection and storage methods that have been validated for one sample type cannot be assumed to be optimal for other sample types. As such, careful preliminary work to optimize processing conditions for sample types is often necessary. Cuthbertson *et al*., (2014) have shown that length of time between sample collection and freezing and the number of freezes–thaw cycles a sample undergoes can affect the microbial community profiles that are detected. DNA extraction methodology must be able to cause cell lysis of the diverse microbial taxa, otherwise, sequencing findings may be dominated by DNA derived from easy-to-lyse microbes. DNA extraction methods that comprise mechanical lysis (or bead beating) are considered to be more effective than those that rely on chemical lysis (Yuan *et al*., 2012). Extraction techniques, based on bead beating result in shortened DNA fragments. This way can contribute to DNA loss during library preparation methods that use fragment size–selection techniques. Contaminants that could interfere with the analysis may derive from kit reagents or from the laboratory environment (Tanner *et al*., 1998), from the sequencing reads of previous analyzes or from human DNA of other hosts (Thomas *et al*., 2012; Quince *et al*., 2017).

1.3.2 Library preparation and sequencing

A wide variety of library preparation protocols exist, but they all have in common the fact that fragments of DNA, with a length of 50-500 bp, at the ends are fused with platform-specific adapters. The choice of library preparation and sequencing method depends on the availability of materials, cost, ease of automation and DNA sample quantification. The Illumina platform has become dominant as a choice for shotgun metagenomics due to its wide availability, high outputs (over 1.5 Tb per run), and high accuracy (with a typical error rate of 0.1-1%). There are several methods for generating Illumina sequencing libraries, depending on the method of fragmentation used. Transposase-based "tagmentation", for example in Illumina Nextera and Nextera XT instruments, is a popular method due to its low cost, and dilution methods could reduce these costs ulteriorly (Baym *et al*., 2015). Tagmentation approaches require small DNA inputs (1 ng of DNA), due to the subsequent PCR amplification step. PCR is performed to choose for molecules containing adapters at both ends and to produce sufficient quantities for sequencing. The adapters, at their extremities,

usually contain elements for the immobilization of library molecules on a solid surface (surface of a glass slide) and amplification. One or both adapters may contain sequencing priming sites. In Illumina libraries one adapter includes a 'read 1' (sequencing primer site), in which only one end is sequenced (single-end sequencing) or in which both ends are sequenced (paired-end sequencing) containing a 'read 2' (primer site) and a site for an index primer, which is applied to read the unique index sequence, allowing libraries in a multiplexed run to be distinguished (Quince *et al*., 2017; Van Dijk *et al*., 2014).

1.3.3 Analysis of taxonomic diversity

One of the first classifications that are carried out in a microbial community is the quantification of its taxonomic diversity, which consists in determining which microorganisms are present in a community and their abundance. Taxonomic diversity serves to identify the similarity between two communities. Furthermore, when the biological functions of taxa are known, taxonomic analysis links these biological functions to the microbial community. Taxonomic diversity is quantified through different strategies:

1) **Analyzing taxonomically informative marker genes**. The analysis of marker genes is one of the simplest and most computationally efficient ways to quantify the taxonomic diversity of a metagenome. Each read is compared to a reference database in which there are sequences that provide taxonomic or phylogenetic information (marker genes), through an algorithm able to recognize if a read is equal to the marker gene and to classify the reads based on their similarity with the sequences of the marker genes. Since this approach determines the comparison of metagenomic reads with a relatively small database for the purpose of a similarity search, marker genes analysis can be a quick method for estimating the diversity of a metagenome.

2) **Binning**: grouping sequences into defined taxonomic groups.

3) **Assembling sequences into distinct genomes**.

These strategies are not mutually exclusive and may be synergistic. In some situations, it may be suitable to bin sequences into taxonomic groups and then subject each group's sequences to assembly, while other instances may warrant conducting an initial assembly and then subjecting the assembled sequences to binning (Sharpton, 2014).

1.3.4 Assembly

When conducts shotgun metagenomics, the complete sequences of protein coding genes (previously characterized or novel) as well as full operons in the sequenced genomes can offer inestimable functional knowledge about the community. Hence the first approach is an assembly of shorter reads

into genomic contigs and orientation of these into scaffolds is often executed to offer a more compact and concise view of the sequenced community under investigation. Two strategies can be employed for assembly shorter reads into contigs: reference-based assembly (co-assembly) and de novo assembly. Reference-based assembly relates for use of one or more reference genomes as a "map" in order to create contigs, which can constitute genomes or parts of genomes belonging to a specific species or genus. Tool's software such as Newbler (Roche), MIRA 4 (Chevreux *et al*., 2004) or AMOS, and MetaAMOS (Treangen *et al*., 2013) are commonly used in metagenomics for performing referenced-based assemblies. These tools are not computationally intensive and perform well when metagenomic samples contains sequences where closely related reference genomes are available. In such cases, sequences from closely related organisms would have already been deposited in online databases, enabling them to be used as references for the assembly procedure (Oulas *et al*., 2015).

De novo assembly refers to the generation of assembled contigs using no prior reference to known genome(s) (Paszkiewicz and Studholme, 2010). Assembly merges collinear metagenomic reads from the same genome into a single contiguous sequence (i.e., contig) and is helpful for produce longer sequences, which can simplify bioinformatic analysis relative to unassembled short metagenomic reads. If utilized to quantify taxonomic abundance, one must be careful to track contig coverage (i.e., the number of assembled reads that align to the average base in the contig), as contigs are later treated as a single sequence in most downstream analyses, likely may thus not accurately quantify the abundance of the taxon as it is represented in the raw data (Sharpton, 2014). This task is computationally expensive, requires larger computational resources and relies heavily on de Bruijn graphs tools. Despite these tools were built with the assumption of assembling a single genome, often underperform when used for metagenome assemblies. Metagenome assembly is more difficult because the coverage of each constituent genome depends on the abundance of each genome in the community. Low abundance genomes may end up fragmented if overall sequencing depth is insufficient to form connections in the graph (Quince *et al*., 2107). Which is why, were developed of the next generation of assembly tools, the assemblers MetaVelvet (Namiki *et al*., 2012) and Meta-IDBA (Peng *et al*., 2011). These tools employ a combined binning and assembly approach to create more accurate assemblies from datasets containing a mixture of multiple genomes. Both assemblers, they make use a multiple k-mer approach to detect kinks in the de Bruijn graph and then use these k-mer thresholds to decompose the graph into subgraphs. In addition, avoid the task of choosing a k-mer length that works well for both low- and high-abundance species. These tools further assemble contigs and scaffolds based on the decomposed subgraphs, and thus perform a more efficient grouping/ assembly of contigs, effectively separating those belonging to different species (Oulas *et al*., 2015).

There are various considerations associated assembling metagenomic sequences. First, assembly tends to be restricted to the most abundant taxa in the community. Second, assembly may produce in silico chimeras, so it should have been applied cautiously and with consideration. Repetitive regions within a genome are also difficult to assemble. Third, assembly can be computationally intensive, especially in its requirements for RAM. For these reasons, binning sequences prior to assembly can be a good way to cut reduce on the computational complexity (Sharpton, 2014).

1.3.5 Binning

Binning refers is the process of grouping reads or contigs into individual genomes and assigning the groups to specific species, subspecies, or genus. Binning plays an essential position in the analysis of metagenomes. Firstly, depending on the method used, binning can provide indications of the presence of new genomes that are difficult to identify. Second, it can provide information on the distinct number and type of taxa present in the community. Furthermore, binning provides a way to reduce the complexity of the data, so that post-binning analyzes (e.g., mounting) can be performed independently for each read rather than the totality of the data. Binning can be performed on assembled or unassembled data, although most algorithms manage to be more precise as the length of the reads increases. Two methods have been developed, on the information used to group the sequences: 1) compositional based binning, makes use of the fact that genomes have conserved nucleotide composition (e.g., a certain GC or the abundance distribution of k-mers). Using this conserved species-specific nucleotide composition, these methods are able of grouping sequences into their respective genomes. Composition-based binning when performed on short reads (i.e., 150 bps), is not reliable as they do not contain enough information, in fact, is performed composition-based binning on assembled datasets. Thus, longer contigs can provide the required k-mer distribution information, which will allow effective binning and taxonomic assignment of the binned fragments. 2) similarity or homology-based binning, in which the unknown DNA fragment might encode for a gene and the similarity of this gene with known genes in a reference publicly available database (e.g., NCBI's nonredundant database - nr or PFAM) can be used to classify and hence bin the sequence. This way may provide higher annotation accuracy and resolution compared to compositional based binning. This method requires more computational research since each read is aligned and compared with a large number of sequences. This method is obviously not ideal for the identification of new genomes, however similarity-based binning of the reads allows for greater accuracy and resolution (Thomas *et al*., 2012; Sharpton, 2014; Oulas *et al*., 2015). Composition based binning algorithms feature the tools TETRA (Teeling *et al*., 2004), S-GSOM (Chan *et al*., 2008) PhylopythiaS (Patil *et al*., 2012), and ClaMS (Pati *et al*., 2011). Exist there are also tools that employ similarity-based

binning algorithms in their metagenomics analysis pipelines such as MG-RAST and MEGAN. MEGAN compares the reads to the database in which the sequences are inserted according to the NCBI taxonomic classification (Huson and Weber, 2013). MG-RAST uses phylogenomic reconstruction of database sequences to which a read is similar (Meyer *et al*., 2008; Glass *et al*., 2010).

1.3.6 Annotation of biological functions

Metagenomics seeks to clarify what are the collective functions that are encoded in the genomes of organisms that are part of a community. The functional diversity of a community can be quantified by correlating metagenomic sequences to particular functions. Reads that contain coding sequences for specific proteins are identified through particular algorithms (*gene prediction*). Each coding sequence is then compared with other genes, proteins, protein families or metabolic pathways present in the database, of which some functional annotation is known (*functional annotation*). This metagenomic analysis produces a profile describing the number of distinct functions and their relative abundance and can be used to compare various metagenomes and identify communities that are metabolically similar (Human Microbiome Project Consortium, 2012), knowing that specific treatments affect the functional composition of the community (Looft *et al*., 2012). The metagenomes resulting from shotgun sequencing may also reveal the presence of novel genes (Nacke *et al*., 2012) or provide information into the ecological conditions associated with those genes whose function is not yet known (Buttigieg *et al*., 2013; Sharpton, 2014).

Gene prediction determines which metagenomic reads contain coding sequences. Once identified, these coding sequences can be related to a function (Sharpton, 2014). Feature prediction is the process of labelling sequences as genes or genomic elements. For completed genome sequences a number of algorithms were developed (Delcher *et al*., 1999) that identify CDS with more than 95% accuracy and a low false negative ratio. Gene prediction can be carried out on assembled or unassembled metagenomic sequences. For the assembled sequences, gene prediction is very similar to the sequence analysis that is performed on the whole genome, through prediction algorithms that require species-specific parameters. A number of tools were specifically designed to handle coding DNA sequences, including MetaGeneMark (McHardy *et al*., 2007), FragGeneScan (Rho *et al*., 2010), Orphelia (Hoff *et al*., 2009) and Metagene (Noguchi *et al*., 2008) and all of which utilize ab initio gene prediction algorithms. Often, annotation pipelines use an intersection of these tools to obtain a more informative prediction of the protein coding genes. Gene prediction tools utilize codon information (i.e., start codon-AUG) to identify potential open reading frames and hence to classify sequence stretches as either coding or non-coding. Most tools can be trained by using the desired training sets (Thomas *et*

al., 2012; Oulas *et al*., 2015). For example, FragGeneScan is trained for prokaryotic genomes only and is used by IMG/MER and MG RAST as well as EBI Metagenomics. It is believed to be one of the most accurate gene-prediction tools currently available. The unassembled sequences, the problem is when there are partially coding sequences in which the coding part, or the gene, starts upstream or stops downstream of the read. One of the simplest methods of identifying coding sequence is binning which map the reads to a database of gene sequences. This method also allows to provide the functional annotation of the analysed gene (Deshpande *et al*., 2013).

This method has been used to quantify the genetic diversity in marine microbial communities (Rusch *et al*., 2007), and the gut microbiota (Qin *et al*., 2010) and is helpful for cataloguing the specific genes present in a metagenome. This method is high-throughput gene prediction since it relies on algorithms that are able to associate a genomic fragment with a very similar sequence present in the database. If the objective is to identify novel or highly divergent genes within a microbial community, this is not the best method. Another method for gene prediction involves transduction of each read into six possible proteins coding and the confrontation of each of the resulting proteins in a database of protein sequences by aligning the sequences. Hence, the sequence alignment can be useful to identify those metagenomic sequences that encode peptides homologous to proteins present in the database (Sharpton, 2014).

The coding sequences identified, can be associated with a precise biological function, by classifying the predicted metagenomic proteins into protein families. A protein family is a group of evolutionarily related protein sequences or subsequences (Finn *et al*., 2014). On the other hand, the proteins within a family share a common predecessor and hence encode similar biological functions and to that, the sequence encodes the family's function. Altogether, metagenomic annotation involves functional assignment to the protein coding genes, achieved by homology-based searches of query sequences against databases containing known functional and/or taxonomic information (Oulas *et al*., 2015). Classification of a metagenomic protein sequence within a protein family demands its comparison within a database in which there are known sequences that code for certain proteins (sequence database) or a comparison of the sequence with a probabilistic model describing the diversity of proteins in a family (HMMs database). Comparing metagenomic reads to a database of sequences tends to be fast and produces results more specific hits for reads that are closely related to sequences in the database while comparing metagenomic reads to a database of HMMs tends to identify more distantly related and diverged members of a family. Many functional annotation databases are available to obtain annotation for metagenomic datasets, such as SEED subsystems (Overbeek *et al*., 2005), eggnog (Muller *et al*., 2010), KEGG (Kanehisa *et al*., 2004), and COG/KOG (Tatusov *et al*., 2003), that are sequence database. For example, SEED annotation system (employed by MG-RAST)

links specific family level functions into higher-order functional subsystems (Overbeek *et al*., 2014) while EggNOG is a database of non-supervised orthologs groups of proteins that tends to be frequently updated so as to include a relatively large amount of sequence diversity (Powell *et al*., 2014). Finally, HMM databases in metagenomic analyses tend to be limited to PFAM, which uses HMMs to model protein domains (Finn *et al*., 2014). On the other hand, no reference database covers all biological functions, the ability to visualize and merge the interpretations of all database searches within a single framework is important. For these reasons, there are several web servers such as MG-RAST, IMG/M and CAMERA (Glass *et al*., 2010; Markowitz *et al*., 2008; Sun *et al*., 2010) that interface with distributed computing clusters to conduct gene prediction, the database search, family classification and annotation.

Protein family classification of reads requires complex calculations as all metagenomic peptides are compared to all protein sequences in the database. Each comparison is independent so that the calculation clusters and servers can distribute the computational load in parallel and improve the throughput achieved. Protein family classification of metagenomic reads has some drawbacks. Since a microbial community is very heterogeneous, the functional diversity encoded in the metagenome can only approximate the functional activity of a community. Furthermore, most databases contain families whose biological function is not yet known, so the reads homologous to these families of proteins still cannot find a functional association. Finally, each database uses different approaches so each can annotate different gene portions, which can then produce different functional profiles to describe a microbial community (Sharpton, 2014).

A complementary approach to metabolic function profiling of metagenomes is an in-depth characterization of specific functions of interest. For example, the identification of genes involved in antibiotic resistance, in a microbial community can inform on the spread of antibiotic resistance (Pehrsson *et al*., 2016). Ad hoc methods and manually curated databases of antibiotic-resistance genes have been crucial to this approach. ARDB (Liu and Pop, 2009) was the first widely adopted resistance database and is now complemented by additional resources, such as Resfams (Gibson *et al*., 2015). Comparably large efforts are also devoted to reporting the virulence repertoire of a metagenome; targeted analyses of metagenomes for specific gene families of interest can also be used to validate findings from single, cultivation-based isolate experiments (Quince *et al*., 2017).

1.3.7 Online software for metagenomic sequencing data

There is several online software that can be useful for processing metagenomic data, some of them are free servers. In order to meet the computational needs for metagenomic data analysis, **MG-RAST** (http://metagenomics.anl.gov) (Meyer *et al*., 2008) was launched in 2007 at Argonne National

Laboratory as a free server and has grown into a dominant resource for the metagenomic research around worldwide. MG-RAST is open-source bioinformatics software for processing, analysing, sharing and disseminating metagenomic datasets. Receives user DNA sequence data and processes the data utilizing the databases (inside) for annotation mapping, as well as the NCBI taxonomy. When the data obtained from the sample is uploaded, these are comparable with those contained in the database. The primary data product displayed to the user by MG-RAST is in the form of abundance profiles for specific taxa, and taxonomic information is projected against this data. The sequences are normalized and processed, and the results are generated automatically. Supported are the comparison of NCBI taxonomies derived from whole genome shotgun data and the comparison of relative abundance for KEGG, eggNOG, COG/KOG and SEED subsystems on multiple levels of resolution. Both IMG/MER and MG-RAST are widely used data management repositories and comparative genomics environments. They are fully automated pipelines that provide quality control, gene prediction, and functional annotation. The system makes the results visible on the display and allows the user to change the parameters. Access to data is password protected and all data generated by the automatic pipeline is available through the download of data products generated, as well as optional sharing and publishing within the respective portals. MG-RAST predicts all genes in the metagenome and identifies the homologs of those genes present in the isolate genomes using a tool named BLAT (BLAST-like alignment tool) (Kent, 2002). BLAT misses' similarities below 70% identity, so many strong hits to other genes are missed. After the best hits to genes from an isolated genome are identified, all subsequent analysis is done using the genes of the isolate genomes and not the genes of the metagenome at hand. This creates many limitations since the analysis is not performed on the original genes of the metagenome, but on the "proxy" genes to the isolated genomes instead. The advantage of this method is its computational speed. The only computationally intensive step is to find the best hits of the metagenomes against the isolates. Once this is done, all other comparisons are pre-existing. The MG-RAST web interface enables comparison using a number of statistical techniques and enables the incorporation of metadata into the statistics. MG-RAST system supports shotgun and amplicon metagenomes from any platform in FASTQ or FASTA format and, in addition, requires a minimum read length of 75 bp and a minimum dataset size of 1 megabase (Thomas *et al*., 2012; Oulas *et al*., 2015; Wilke *et al*., 2016). Hence, within this framework the tasks can be divided into three conceptual steps: 1. data cleansing/quality evaluation, 2. data transformation/reduction, and 3. data analysis/interpretation. In the first step, the objective is to filter the noisy sequences and includes "Preprocessing" removes some ambiguous sequences or low-quality sequences; "Dereplication" removes duplicated sequences; "Screening" removes human genomic sequences that may have been mixed accidentally, using a sequence alignment tool, Bowtie (Langmead *et al*., 2009).

In the second step, provides for “gene prediction” finds genes within the DNA sequence, using an HMM-based gene prediction tool for short and error-prone sequences, FragGeneScan (Rho *et al*., 2010); “Clustering” further compresses the data by grouping similar sequences and presenting only one consensus sequence, using a search-based clustering tool, Uclust (Edgar, 2010). In the third step, provides for “identify proteins or RNAs” by comparing with data in public databases and then to do “annotation,” using the sequence alignment tool BLAT (Tang *et al*., 2014). **MGmapper** is software used to process raw sequence data and ensure access for routine analysis of complex datasets. Performs reference-based sequence assignment, followed by a post-processing analysis to make reliable taxonomy annotation at species and strain levels. MGmapper pipeline enabling usage of any custom database based on a set of fasta sequences, the usage of databases antimicrobial resistance genes and 16S rRNA. The current version of MGmapper includes 18 databases. MGmapper processes fastq reads in four steps: 1) pre-processing (an optional trimming and filtering) of raw reads to remove potential positive control reads, by using the Cutadapt program (Martin, 2011); 2) mapping of reads to specified reference databases and alignment-based filtering. However, in doing so, properly paired read may align to more than one reference sequence, located in different reference sequence databases; 3) identification the best hits, assigning a read-pairs to only one specific reference sequence; 4) post-processing of taxonomy annotations and preparation of excel and text files which includes normalized abundance statistics, read and nucleotide count (Petersen *et al*., 2017). The **CosmosID** platform a simple and easy-to-follow algorithm, has the potential to provide fast, reliable bacterial detection and identification down to the species and strain level from metagenomic shotgun sequencing data. Identifies metagenomic sequences using statistical and computational methods, uses raw and not assembled reads as inputs and compares them with the sequences present in the reference databases of bacteria, viruses, fungi, protists and antibiotic resistance genes. Nevertheless, strategies for metagenomic detection of antibiotic resistance genes need further development (Yan *et al*., 2019). The **One Codex** data platform developed for the taxonomic and functional analysis of metagenomes (WGS), 16S and other sequencing data, is used freely to analyse public data and was drawn up to identify microbial sequences employing a “k-mer based” taxonomic classification algorithm through a web-based data platform (Wood and Salzberg, 2014). One Codex uses two reference databases, the full One Codex database which currently includes approximately 40.000 genomes of bacteria, viruses, fungi, archaea and protists and a smaller database, NCBI RefSeq database which includes over 8.000 microbial genomes. Samples are uploaded (through a graphical upload tool) in FASTA or FASTQ format to the One Codex platform with both drag-and-drop and folder navigation options. Once uploaded, reads are taxonomically classified and the interactive report is populated and linked to the user’s account (Minot *et al*., 2015).

1.3.8 Biostatistical analysis of metagenomic sequencing data

Independently of the methods used for shotgun metagenomic sequence analyses, the outputs will understand data matrices of samples versus microbial features (i.e., species, taxa, genes and pathways). Analysis statistical uses tools to interpret matrices and decipher how the findings relate to the test sample metadata. Numerous tools and software packages exist for performing statistical analysis of metagenomic data which include multivariate statistics and machine learning. Unsupervised methods include clustering and correlation of samples, and visualization techniques such as heat maps, ordination (e.g., principal component analysis and principal coordinates analysis) or networks, which allow patterns in the data to be revealed graphically whereas supervised methods include statistical methods. For example, multivariate analysis of variance (ANOVA) for direct hypothesis testing of differences between groups, or machine learning classifiers that train models to label groups of samples (Pasolli *et al*., 2016). Both of these methods deem the community as a whole. In recent years, the R statistical programming language (Team, 2008) has gained popularity and is currently used for multivariate statistics. Packages such as vegan (Oksanen *et al*., 2007) Bioconductor (Gentleman *et al*.,2004) and phyloseq (McMurdie and Holmes, 2013) provide multiple in-built functions and libraries for performing a wide range of statistical analyses required for metagenomic datasets (Oulas *et al*., 2015). Numerous tools and software packages exist for performing statistical analysis of shotgun metagenomic data. For example, alpha diversity (which measures variability, richness, dominance, and evenness within a single population) is supported through software such as Shannon entropy (Gorelick, 2006) and Phylogenetic Diversity (PD) (Chao,1984). Rarefaction analysis is used to assess the coverage of the microbial community in the sample. Rarefaction curves plot the sample size with respect to the esteemed number of genera (Jaenicke *et al*., 2011). Beta diversity (which is the diversity across many populations or samples) is calculated using several matrices, such as unweighted and weighted UniFrac (Lozupone *et al*., 2006) and PCoA (Principal Coordinate Analysis) (Oulas *et al*., 2015; Ghosh *et al*., 2019). There are two main approaches for quantifying β-diversity: those that take into account the evolutionary differences between communities (phylogenetic β-diversity), and those that do not (taxon-based or non-phylogenetic methods) (Jovel *et al*., 2016). Finally, one of the aims of metagenomics is to link functional and phylogenetic information to the chemical, and physical that characterize an environment. While measuring all these parameters can be time-consuming and cost-intensive, it allows retrospective correlation analysis of metagenomic data that was perhaps not part of the initial aim of the project or might be of interest for other research questions. The value of such metadata cannot be overstated

and, in fact, has become mandatory or optional for the deposition of metagenomic data into some databases (Markowitz *et al*., 2008; Thomas *et al*., 2012).

1.4 Safety and hygiene of poultry meat production

Microbial diversity is shaping the ecology of diverse ecosystems. Studying microbial diversity and its interactions is a challenge due to the variability, which can occur between sources. In relation to poultry meat production, the microbial diversity and its dynamics can influence the product shelf life and safety, if spoilage bacteria are favoured and pathogenic bacteria are present and able to grow (Rouger *et al*., 2017a). In the poultry meat agri-food industry, contaminating bacteria came from faeces, skin, or feathers (animal microbiota), air, equipment, surfaces (production plant environment) and from human manipulators (Chaillou *et al*., 2015). Moreover, the microbiological quality of meat depends on various factors, such as the level of contamination during slaughter and processing stages, the temperature and conditions of storage, and the physiological status of the animal at the time of slaughter (Nychas *et al*., 2008). During food processing, the carcasses come into contacts with various surfaces that can result in cross-contamination events which implies that during processing of poultry carcasses, microbial contamination often occurs as a result of the processing procedures employed. Another source of contamination can be represented by other birds which can enter in contact with carcasses (Rouger *et al*., 2017b). Most food-borne outbreaks are caused by cross-contamination phenomena due to food contact with contaminated surfaces (Griffith *et al*., 2015). In particular, at the slaughterhouse, the microbial contamination that occurs during the path of the carcasses along the line occurs mainly at the surface level and is not limited to zoonotic agents only, although the remaining part of microorganisms (such as the spoilage ones) is also very important (Bolder, 2007; Luber, 2009). During the scalding and de-feather removal phases, the epidermis, already colonized by different microbial communities, is usually damaged or removed, and consequently, the exposed portion is more vulnerable to contamination caused by gram-negative bacteria from the intestinal contents or from other animals (Thomas and McMeekin, 1980). At the time the carcasses are conveyed along the slaughter line, they can be subject to cross-contamination phenomena by aerosols or condensation formed on the equipment or on the ceiling, thus compromising the shelf life of the final product. Another way of cross-contamination can be represented by the evisceration phase, following the expulsion of faecal material or through the contact plant. Furthermore, in the evisceration phase, it is important that during the removal of the gut it remains intact, in order to prevent the spread of faecal material and bacteria on the carcasses, suggesting that evisceration is a phase that, if performed incorrectly, can cause a significant increase in the levels of microbial load on the carcasses (Mead, 2004; Bolder, 2007). The gastrointestinal tract of poultry hosts many bacteria,

such as *Campylobacter* spp. or *Salmonella* spp., that can be potentially dangerous. Poultry gut microbiota has been studied to correlate gut microbiota, animal feeding and their health (Shaufi *et al*., 2015; Ranjitkar *et al*., 2016). According to Hue *et al*., (2011) and Pacholewicz *et al*., (2016), there is a correlation between the number of *Campylobacter* in the ceca and the contamination level seen on carcasses by measuring an average contamination level of 8.05 log CFU/g of ceca and 2.39 log CFU/g of carcasses. Some of the microorganisms originated from the animal's gastrointestinal tract as well as from the environment with within the animal had contact at some time before or during slaughter (Koutsoumanis *et al*., 2004). This can be assessed with studies on the origin of the contaminants showing an association between work surfaces and the presence of Enterobacteriaceae on meats. Psychrophilic bacteria (capable of developing at low temperatures) are also regained from hides and work surfaces (within the slaughterhouse) just like from carcasses and slaughtered meat at all stages of processing (Gill, 2005). As regards the chilling phase of the carcasses, air-chilling is becoming popular all over the world, although according to various studies the latter does not show any reduction in the total or pathogenic microbial load (Allen *et al*., 2000; Fluckey *et al*., 2003). The storage temperatures of the carcasses and their maintenance during the chill chain are equally important, as this allows to keep under control the growth of spoilage bacteria, especially psychrophilic ones. The incorrect maintenance of the chill chain could lead to the formation of condensation on the surfaces of the carcasses, which can constitute a favourable environment for the development of bacteria; however, even in the absence of condensation, the activity water (aw) of these surfaces must still be adequate in order to avoid unwanted bacterial proliferation (Hinton, 2000).

1.5 Microbiological criteria and sampling plans: Regulation (EC) No. 2073/2005

Developing microbiological criteria for a food is a complex process that requires considerable effort since their application demands considerable demands (Van Schothorst *et al*., 2009). According to the International Commission on Microbiological Specifications for Foods defined the term "Microbiological criterion" as the presence of microorganisms or their toxins or the number of organisms per unit of mass/volume/ area, determined by use of defined procedures and applied in acceptance sampling of food (ICMSF, 1986). The microbiological criterion, it is expressed (whatever type of criterion is used) as a definite value, which may be a number of microorganisms or the absence of a given organism in a defined quantity of sample (i.e., no Salmonella in 25 g) (Simonsen, 1995).
At the level of the European Union, the EU legislation, containing the "microbiological criteria for foodstuffs", is represented from Commission Regulation (EC) No 2073/2005 and its amendments, lays down the microbiological criteria for certain microorganisms and the implementing rules to be complied with by food business operators that they must respect during the performance of the

activities. The Regulation provides in Annex I of Chapter 1 the food safety criteria of products placed on the market during their shelf life. Definition of food safety criterion as established in the Commission Regulation (EC) No 2073/2005 means a criterion defining the acceptability of a product or a batch of foodstuff applicable to products placed on the market. In Annex I of Chapter 2 reports the process hygiene criteria to assess the adequacy and the level of hygiene of the production process. Definition of process hygiene criterion as established in the Commission Regulation (EC) No 2073/2005 means a criterion indicating the acceptable functioning of the production process. Such criterion is not applicable to products placed on the market. It sets an indicative contamination value above which corrective actions are required in order to maintain the hygiene of the process in compliance with food law (European Commission, 2005).

All the analytical reference methods listed in the EU Regulation 2073 to verify the compliance of food lots to both food safety criteria and process hygiene criteria are based on ISO cultural methods to both detect and/or quantify bacteria pathogens or groups of indicator microorganisms. Alternative methods, including molecular methods such as PCR, can be applied. However, whenever a sample positive for a pathogen is detected, the result must be confirmed using the ISO reference method. In this context, the possibility to detect all pathogens in the same sample using sequencing methods such as metataxonomics or metagenomics should be always couple with the possibility to double check the presence of detected pathogens using the reference ISO method.

1.6 Foods of animal origin and food-borne outbreaks

The EU regulations are structured to prevent the occurrence of biological hazards in feed and food to prevent the risk those hazards can cause for human health. Unfortunately, the implementation of such regulations cannot prevent the occurrence of food-borne outbreaks that in the EU Member States in 2019 were 5.175 involving 49.463 cases of illness, 3.859 hospitalizations and 60 deaths. In addition, 117 outbreaks, 3,760 cases of illness and 158 hospitalizations were communicated by six non-MS. *Salmonella* was the agent most identified in food-borne outbreaks and caused the highest number of hospitalizations (49.6% of all outbreak-associated hospitalizations). At the EU level, the consumption of food of animal origin (i.e., fish and fishery products, eggs and egg products, meat and meat products, milk and milk products) was associated with most of the food-borne strong-evidence outbreaks (EFSA and ECDC, 2021a).

At the EU level and reported to ECDC, *Campylobacter* was the third most frequently reported causative agent for food-borne outbreaks. Campylobacteriosis has been the most commonly reported zoonosis in humans in the EU since 2005. The specific notification for campylobacteriosis is mandatory in Iceland, Norway, Switzerland and in twenty-one EU Member States. The reporting of

food-borne campylobacteriosis disease outbreaks in humans is binding according to the Zoonoses Directive 2003/99/EC. In 2019, 220.682 confirmed cases of human campylobacteriosis were reported by 28 EU MS with notification rates of 59.7 cases per 100.000 population, confirming a decrease of 6.9% compared with 2018 (Table 1). The lowest country-specific notification rates (≤8.6 per 100.000) in 2019 were ascertained in Bulgaria, Cyprus, Greece, Latvia, Poland, Portugal and Romania while the highest rates were ascertained in Czechia, Slovakia, Denmark and the United Kingdom (EFSA and ECDC, 2021a).

Table 1: Reported human cases by country and year of campylobacteriosis and notification rates per 100.000 population in the EU/EFTA, from 2015 to 2019 (EFSA and ECDC, 2021a).

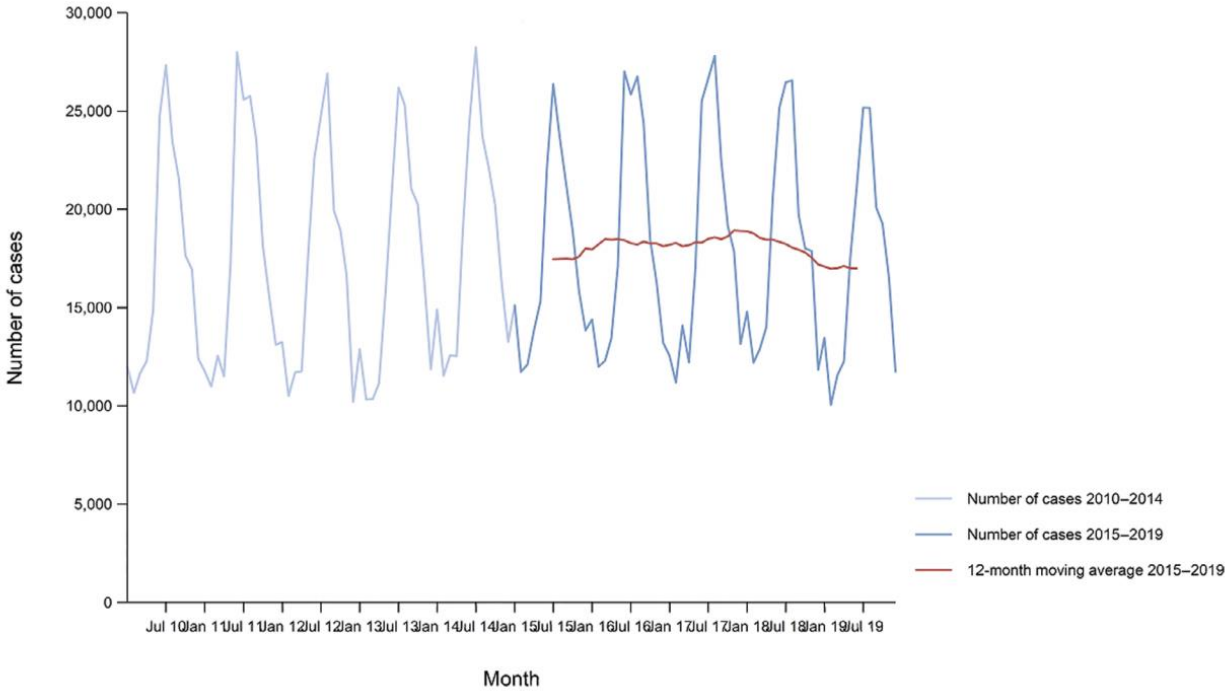
Country	2019					2018		2017		2016		2015	
	National coverage[a]	Data format[a]	Total cases	Confirmed cases & rates		Confirmed cases & rates		Confirmed cases & rates		Confirmed cases & rates		Confirmed cases & rates	
				Cases	Rate	Cases	Rate	Cases	Rate	Cases	Rate	Cases	Rate
Austria	Y	C	6,573	6,573	74.2	7,999	90.7	7,204	82.1	7,083	81.4	6,258	72.9
Belgium	Y	A	7,337	7,337	64.0	8,086	70.9	8,649	76.2	10,055	88.9	9,066	80.7
Bulgaria	Y	A	231	229	3.3	191	2.7	195	2.7	202	2.8	227	3.2
Croatia	Y	C	1,732	1,722	42.2	1,965	47.9	1,686	40.6	1,524	36.4	1,393	33.0
Cyprus	Y	C	21	21	2.4	26	3.0	20	2.3	21	2.5	29	3.4
Czechia	Y	C	23,169	22,894	215.0	22,895	215.8	24,326	230.0	24,084	228.2	20,960	198.9
Denmark	Y	C	5,402	5,402	93.0	4,559	78.9	4,255	74.0	4,712	82.6	4,327	76.5
Estonia	Y	C	348	347	26.2	411	31.2	285	21.7	298	22.6	318	24.2
Finland	Y	C	4,382	4,382	79.4	5,099	92.5	4,289	77.9	4,637	84.5	4,588	83.8
France[b]	N	C	7,712	7,712	57.5	7,491	56.0	6,579	49.2	6,698	50.3	6,074	45.7
Germany	Y	C	61,526	61,254	73.8	67,585	81.6	69,251	83.9	73,736	89.7	69,921	86.1
Greece	Y	A	366	366	3.4	357	3.3	.	.	.	.	.	.
Hungary	Y	C	6,441	6,400	65.5	7,117	72.8	7,807	79.7	8,556	87.0	8,342	84.6
Ireland	Y	C	2,776	2,776	56.6	3,044	63.0	2,779	58.1	2,511	53.1	2,453	52.4
Italy[d]	N	C	1,633	1,633	–	1,356	–	1,060	–	1,057	–	1,014	–
Latvia	Y	C	133	133	6.9	87	4.5	59	3.0	90	4.6	74	3.7
Lithuania	Y	C	1,225	1,221	43.7	919	32.7	990	34.8	1,225	42.4	1,186	40.6
Luxembourg	Y	C	271	271	44.1	625	103.8	613	103.8	518	89.9	254	45.1
Malta	Y	C	298	278	56.3	333	70.0	231	50.2	212	47.1	248	56.4
Netherlands[c]	N	C	3,415	3,415	34.1	3,091	34.6	2,890	32.5	3,383	38.3	3,778	43.0
Poland	Y	C	715	715	1.9	719	1.9	874	2.3	773	2.0	653	1.7
Portugal	Y	C	942	887	8.6	610	5.9	596	5.8	359	3.5	271	2.6
Romania	Y	C	805	805	4.1	573	2.9	467	2.4	517	2.6	311	1.6
Slovakia	Y	C	7,829	7,690	141.1	8,339	153.2	6,946	127.8	7,623	140.5	6,949	128.2
Slovenia	Y	C	1,085	1,085	52.1	1,305	63.1	1,408	68.2	1,642	79.5	1,328	64.4
Spain[d],[f]	N	C	9,723	9,723	–	18,411	–	18,860	–	15,542	–	13,227	–
Sweden	Y	C	6,693	6,693	65.4	8,132	80.4	10,608	106.1	11,021	111.9	9,180	94.2
United Kingdom	Y	C	58,718	58,718	88.1	65,246	98.4	63,267	96.1	58,901	90.1	59,797	92.2
EU Total	–	–	**221,501**	**220,682**	**59.7**	**246,571**	**64.1**	**246,194**	**64.9**	**246,980**	**66.4**	**232,226**	**63.0**
Iceland	Y	C	136	136	38.1	145	41.6	119	35.2	128	38.5	119	36.2
Norway	Y	C	4,154	4,154	78.0	3,668	69.3	3,883	73.8	2,317	44.5	2,318	44.9
Switzerland[e]	Y	C	7,223	7,223	84.0	7,675	90.1	7,219	85.4	7,980	94.4	7,070	84.5

(a): Y: yes; N: no; A: aggregated data; C: case-based data.
(b): Sentinel surveillance: notification rates calculated with estimated coverage of 20%.
(c): Sentinel surveillance: notification rates calculated with estimated coverage 52%.
(d): Sentinel surveillance; no information on estimated coverage. So, notification rate cannot be estimated.
(e): Switzerland provided data directly to EFSA. The human data for Switzerland include data from Liechtenstein.
(f): Data not complete in 2019, rate not calculated.

Between 2015 and 2019, the number of confirmed campylobacteriosis cases reported in the EU/EEA they have had a characteristic seasonality, with peaks cases in the summer months. However, between 2012 and 2019 in the EU a smaller but distinct winter peak has become apparent in Austria, Germany,

Finland, Belgium, Luxembourg, the Netherlands, Switzerland and Sweden. The EU /EEA trend was stable from 2015 to 2019, the only Member State that indicated decreasing trend was Hungary, while Italy, Latvia, Romania and Portugal indicated increasing trends (Figure 1). One Member State advised that the reported number of campylobacteriosis cases is lower than expected (due to the COVID-19 situation in 2020) (EFSA and ECDC, 2021a).

Figure 1: Trend in reported between 2015 to 2019 by month human cases of campylobacteriosis in the EU /EEA (EFSA and ECDC, 2021a).



Source(s): Austria, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Romania, Slovakia, Slovenia, Sweden and the United Kingdom. Belgium, Bulgaria, Croatia, Greece, Portugal and Spain did not report data to the level of detail required for the analysis.

Salmonellosis is the second most common zoonosis in humans in the EU after campylobacteriosis and an important cause of food-borne outbreaks in the EU/EEA. In the EU, notification and surveillance of food-borne salmonellosis in humans is mandatory in accordance with European Commission Decision 2000/96/EC and Decision No 2119/98/EC in Iceland, Norway and Switzerland and twenty-two EU Member States, whereas in six Member States (Belgium, France, Luxembourg, Netherlands and the United Kingdom) is based on other systems. Since 2007, EU and EEA countries report their public health surveillance data to The European Surveillance System (TESSy) at European Centre for Disease Prevention and Control (ECDC). The reporting of food-borne salmonellosis disease outbreaks in humans is binding according to Zoonoses Directive 2003/99/EC. In 2019, 87.923 confirmed cases of human salmonellosis were reported by 28 EU Member States

with a notification rate of 20.0 cases per 100.000 population, confirming at the same level as in 2018 (20.1 cases per 100.000 population) (Table 2) (EFSA and ECDC, 2021a).

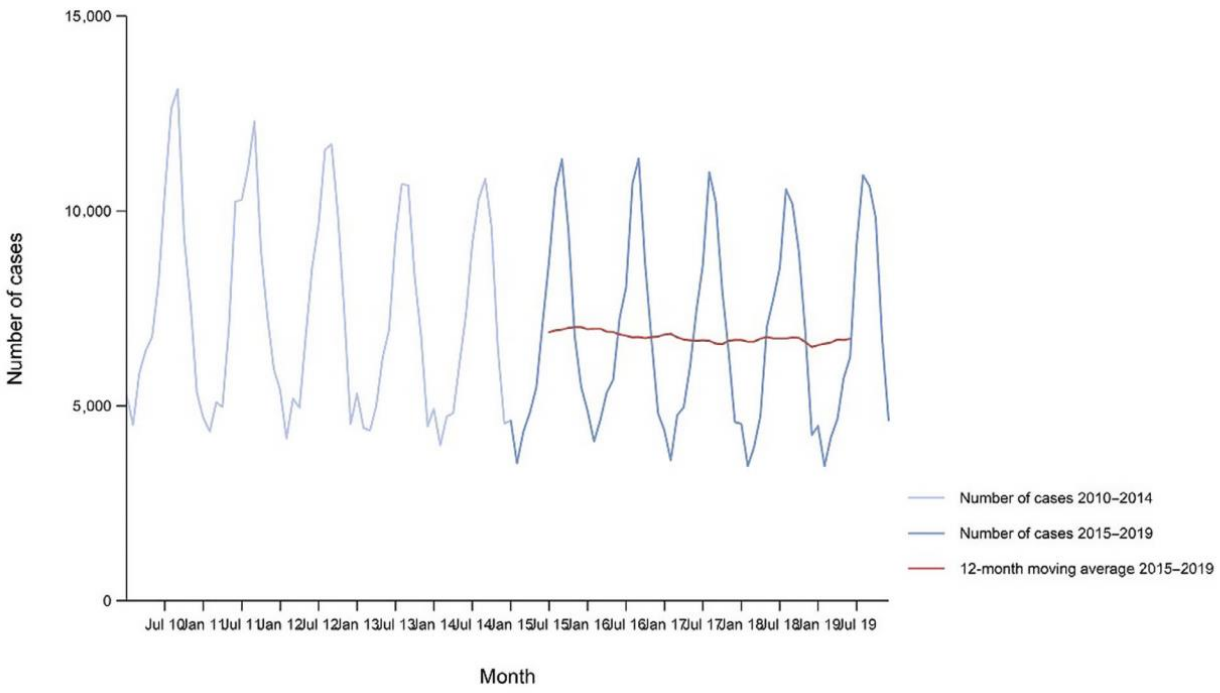
Table 2: Reported human cases by country and year of salmonellosis and notification rates per 100.000 population in the EU/EFTA, from 2015 to 2019 (EFSA and ECDC, 2021a).

Country	2019					2018		2017		2016		2015	
	National coverage[a]	Data format[a]	Total cases	Confirmed cases & rates		Confirmed cases & rates		Confirmed cases & rates		Confirmed cases & rates		Confirmed cases & rates	
				Cases	Rate	Cases	Rate	Cases	Rate	Cases	Rate	Cases	Rate
Austria	Y	C	1,868	1,866	21.1	1,538	17.4	1,667	19.0	1,415	16.3	1,544	18.0
Belgium	Y	C	2,527	2,527	22.1	2,958	26.0	2,298	20.2	2,699	23.9	3,050	27.1
Bulgaria	Y	A	596	594	8.5	586	8.3	796	11.2	718	10.0	1,076	14.9
Croatia	Y	C	1,318	1,308	32.1	1,323	32.2	1,242	29.9	1,240	29.6	1,593	37.7
Cyprus	Y	C	62	62	7.1	44	5.1	59	6.9	77	9.1	65	7.7
Czechia	Y	C	13,306	13,009	122.2	10,901	102.7	11,473	108.5	11,610	110.0	12,408	117.7
Denmark	Y	C	1,119	1,119	19.3	1,168	20.2	1,067	18.6	1,081	18.9	925	16.3
Estonia	Y	C	154	150	11.3	314	23.8	265	20.1	351	26.7	112	8.5
Finland	Y	C	1,175	1,175	21.3	1,431	26.0	1,535	27.9	1,512	27.6	1,650	30.2
France[b]	N	C	8,935	8,935	27.8	8,936	27.8	7,993	24.9	8,876	27.7	10,305	32.3
Germany	Y	C	13,692	13,495	16.3	13,293	16.1	14,051	17.0	12,858	15.6	13,667	16.8
Greece	Y	C	642	642	6.0	640	6.0	672	6.2	735	6.8	466	4.3
Hungary	Y	C	5,172	4,452	45.6	4,161	42.6	3,922	40.0	4,722	48.0	4,894	49.7
Ireland	Y	C	356	347	7.1	352	7.3	379	7.9	299	6.3	270	5.8
Italy	Y	C	3,268	3,256	5.4	3,635	6.0	3,347	5.5	4,134	6.8	3,825	6.3
Latvia	Y	C	472	438	22.8	409	21.1	225	11.5	454	23.1	380	19.1
Lithuania	Y	C	745	736	26.3	779	27.7	1,005	35.3	1,076	37.3	1,082	37.0
Luxembourg	Y	C	131	131	21.3	135	22.4	118	20.0	108	18.7	106	18.8
Malta	Y	C	131	131	26.5	116	24.4	107	23.2	162	36.4	126	29.3
Netherlands[c]	N	C	1,197	1,197	10.8	1,061	9.6	954	8.7	1,150	10.6	974	9.0
Poland	Y	C	8,919	8,373	22.0	9,064	23.9	8,921	23.5	9,718	25.6	8,245	21.7
Portugal	Y	C	500	432	4.2	302	2.9	462	4.5	376	3.6	325	3.1
Romania	Y	C	1,413	1,383	7.1	1,410	7.2	1,154	5.9	1,479	7.5	1,330	6.7
Slovakia	Y	C	5,234	4,992	91.6	6,791	124.8	5,789	106.5	5,299	97.7	4,841	89.3
Slovenia	Y	C	362	362	17.4	274	13.3	275	13.3	311	15.1	401	19.4
Spain[d],[f]	N	C	5,103	5,103	–	8,730	–	9,426	–	9,818	–	9,015	–
Sweden	Y	C	1,990	1,990	19.5	2,041	20.2	2,280	22.8	2,247	22.8	2,312	23.7
United Kingdom	Y	C	9,718	9,718	14.6	9,466	14.3	10,105	15.3	9,900	15.1	9,490	14.6
EU Total	–	–	**90,105**	**87,923**	**20.0**	**91,858**	**20.1**	**91,587**	**19.7**	**94,425**	**20.5**	**94,477**	**21.0**
Iceland	Y	C	50	50	14.0	63	18.1	64	18.9	39	11.7	44	13.4
Norway	Y	C	1,093	1,092	20.5	961	18.2	992	18.9	865	16.6	928	18.0
Switzerland[e]	Y	C	1,547	1,547	18.0	1,467	17.2	1,848	21.9	1,517	17.9	1,375	16.4

(a): Y: yes; N: no; A: aggregated data; C: case-based data.
(b): Sentinel system; notification rates calculated with an estimated population coverage of 48%.
(c): Sentinel system; notification rates calculated with an estimated population coverage of 64%.
(d): Sentinel surveillance; no information on estimated coverage 2015–2018. So, notification rate cannot be estimated. (e): Switzerland provided data directly to EFSA. The human data for Switzerland include data from Liechtenstein.
(f): Data not complete in 2019, rate not calculated.

The highest notification rates in 2019 were ascertained in Czechia and Slovakia (122.2 and 91.6 cases per 100.000 population) while the lowest rates were ascertained by Cyprus, Greece, Ireland, Italy, Portugal and Romania (≤7.1 cases per 100.000 population). Between 2010 and 2019, the number of confirmed salmonellosis cases in the EU/EEA they reported with more cases during the summer months (Figure 2). The EU/EEA trend for salmonellosis was flat from 2015 to 2019, with no Member States with an increasing trend. One Member State (Finland) was the only one reporting a significantly decreasing trend (EFSA and ECDC, 2021a).

Figure 2: Trend in reported between 2015 to 2019 by month human cases of salmonellosis in the EU/EEA (EFSA and ECDC, 2021a).



Source: Austria, Belgium, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Lithuania, Luxembourg, Latvia, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Sweden and the United Kingdom. Bulgaria, Croatia and Spain did not report data to the level of detail required for the analysis.

As in previous years and 2019, the four most commonly reported *Salmonella* serovars in the humans, food and animals were *S. Enteritidis* (50.3%), *S. Typhimurium* (11.9%), monophasic *S. Typhimurium* (1,4,[5],12:i:-) (8.2%) and *S. Infantis* (2.4%) (Table 3). The fifth most common serovar *S. Newport* decreased by 20.0% compared with 2018. Serovar *S. Mikawasima* increased by 92.1% and 137.1% compared with 2018 and 2017, respectively (EFSA and ECDC, 2021a).

Table 3: Distribution from 2017 to2019 of reported confirmed cases of human salmonellosis, by the 20 most frequent serovars in the EU/EEA in 2019 (EFSA and ECDC, 2021a).

Serovar	2019			2018			2017		
	Cases	MSs	%	Cases	MSs	%	Cases	MSs	%
Enteritidis(*)	39,865	27	50.3	39,781	27	49.9	38,780	27	49.2
Typhimurium(*)	9,404	27	11.9	10,395	27	13.0	10,589	27	13.4
Monophasic Typhimurium 1.4.[5].12:i:-(*)	6,491	18	8.2	6,427	17	8.1	6,322	16	8.0
Infantis(*)	1,924	26	2.4	1,859	26	2.3	1,803	26	2.3
Newport	870	24	1.1	1,086	21	1.4	920	24	1.2
Derby	721	23	0.9	710	23	0.9	612	23	0.8
Stanley	560	19	0.7	521	22	0.7	554	21	0.7
Kentucky	545	24	0.7	663	22	0.8	617	19	0.8
Napoli	508	18	0.6	457	15	0.6	406	17	0.5
Agona	503	20	0.6	602	18	0.8	645	20	0.8
Virchow(*)	477	21	0.6	541	24	0.7	510	21	0.6
Coeln	455	18	0.6	443	20	0.6	265	21	0.3
Bovismorbificans	454	19	0.6	465	18	0.6	344	20	0.4
Java	440	14	0.6	415	16	0.5	387	16	0.5
Mikawasima	415	15	0.5	216	13	0.3	175	13	0.2
Chester	350	17	0.4	369	19	0.5	329	18	0.4
Bareilly	321	17	0.4	299	16	0.4	427	18	0.5
Saintpaul	302	20	0.4	324	20	0.4	330	21	0.4
Branderup	300	18	0.4	259	17	0.3	260	18	0.3
Hadar(*)	298	17	0.4	312	20	0.4	334	19	0.4
Other	14,097	–	17.8	13,556	–	17.0	14,288	–	18.1
Total	**79,300**	**27**	**100.0**	**79,700**	**27**	**100.0**	**78,897**	**27**	**100.0**

MS: Member State.
(*): Target Salmonella serovars in poultry populations.
Source(s): 27 MS: Austria, Belgium, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, the United Kingdom; and two non-MS: Iceland and Norway.

1.7 Antimicrobial resistance and animal, environmental and human resistomes

The beginning of the antibiotic era began with the discovery of penicillin in 1928 and later with the introduction to the medical industry in the 1940 (Fleming, 2001). Antibiotics have revolutionized modern medicine and are specific antimicrobial substances that either inhibit or kill the growth of bacteria without damaging host cells and tissues (Aminov *et al*., 2010). Antibiotics have been used in human and veterinary medicine for more than 70 years and have greatly contributed to tackling pathogenic bacteria and protecting human and animal health; for this reason, antibiotics are widely used in the prevention and treatment of bacterial infections (Oniciuc *et al*., 2018; Zhao *et al*., 2021). Initially, most antibiotics (naturally occurring substances) were used for treating infections in humans, but subsequently, the use was extended to disease treatment/prevention in food animals (Van Boeckel *et al*., 2015). One of the main threats to the worldwide public health and food safety is

considered the increasing incidence of antimicrobial resistance (AR) (Roca *et al*., 2015; World Health Organization Antibiotic resistance, 2018). Although AR is not a new phenomenon (D'Costa *et al*., 2011), over the past several decades, the emergence of new antibiotic resistance genes (ARGs) that can spread worldwide have been discovered. The collective pool of ARGs (intrinsic and acquired resistance genes) in a given environment, their precursors in pathogenic and non-pathogenic bacteria and potential resistance mechanisms within microbial communities is termed resistome (Wright, 2007; Kim and Cha, 2021). Each year, hundreds of thousands of people die due to infections by antibiotic resistant bacteria (ARB) and it is estimated that 10 million people will die each year due to a continued rise in AR (Yang *et al*., 2021). Therefore, the controlling the spread of antibiotic resistant bacteria (ARB) is one of the most urgent tasks for human health management in the 21st century (Laxminarayan *et al*., 2013). Over of years, several microorganisms have modified the ability to make chemicals such as antibiotics or antimicrobials, used to prevent microorganisms from growing and multiplying and even to kill them and they are commonly used in human and veterinary medicine to treat a variety of infectious diseases. Antibiotic resistance (also known as antimicrobial resistance, AMR) a gloomy and very common phenomenon around the world, is defined as the reduced ability or inability of an antimicrobial agent to inhibit the growth of a bacterium, such as pathogenic bacteria leading to therapy bankruptcy (EFSA and ECDC, 2021b). A bacterial strain can acquire resistance in different ways: by the uptake of exogenous genes, by the activation/triggering of a genetic cascade, (inducing the expression of resistance mechanisms) and by horizontal gene transfer (HGT) from other bacterial strains or by mutation (EMA and EFSA, 2017). ARGs and ARB have been widely detected in various environments, such as in livestock farming environments, that can be transmitted to humans through the food chain, air, and water; in wastewater treatment plants (Zhang *et al*., 2011); in surface water, drinking water (Chao *et al*., 2013); in soil and also animal and human faeces (Smith *et al*., 2002). Undoubtedly, animal and human guts may constitute reservoirs for ARG and ARB (Wang *et al*., 2021). The same Wang *et al*., (2020) have conducted studies of gut resistance groups in humans, pigs, and chickens, deepening the evolution, distribution, and transmission of ARGs in the gut microbiota. The emergence and spread of AMR can be sparked by several factors such as poor hygiene conditions/practices in the food chain, that can may facilitate the transmission of resistant microorganisms or the continuous use of antibiotics in human medicine and animal husbandry, considering livestock an important source of AMR (Oniciuc *et al*., 2018). Chantziaras *et al*., (2014) have shown that the use of antimicrobials in animals will lead to an increase of AMR and that the reduction of usage will lead to reduced resistance. Nevertheless, overuse of antibiotics in farm animals and humans is speeding up the enrichment and spread of ARB and ARGs, contributing to the resistance crisis (Raffatellu, 2018; Munk *et al*., 2018). The inappropriate use of antimicrobials has

been linked to the appearance/spread of microorganisms which are resistant to them, over time, rendering treatment less effective or ineffective and representing a serious risk to public health and food safety. In Europe, the antibiotics used in food-producing animals and in human medicine belong to the same classes or are frequently the same while the administered quantities and the route of administration may differ between humans and food-producing animals. There are important variations between different countries globally and also within Europe and between and within food-producing animal populations (EFSA and ECDC, 2021b).

1.8 Whole shotgun metagenomic sequencing to investigate the resistome in the food of animal origin ecosystems

Over the years, the research of AR has grown from focusing on single pathogenic organisms in culture to research AR in pathogenic, environmental and commensal bacteria at the level of microbial communities (Crofts *et al*., 2017). As not all microorganisms can be grown in the standard laboratory conditions, as in the case of the microbiome, culture-independent strategies are urgently needed to explore the resistomes in living and non-living environments. High-throughput sequencing-based metagenomic analysis has become important tool for characterizing resistomes, and has been widely used to explore the diversity, abundance and distribution of ARGs between human and food animal environments. For example, metagenomic sequencing has been employed to study ARGs in several environmental microbial community (Schmieder and Edwards, 2012), poultry and pig faeces (Tong *et al*., 2017; Munk et al., 2018), cattle rumen (Singh *et al*., 2012), air (Yang *et al*., 2018), soil and wastewater (Tang *et al*., 2016), drinking water (Ma *et al*., 2017) and human gut (Hu *et al*., 2013). Using shotgun sequencing, Wang *et al*., (2019) have revealed that poultry faeces include multiple ARGs, and their abundance in poultry is greater than in humans. However, metagenomics is also a powerful tool that allows potential applications in AMR selection and surveillance and therefore could assist the tracking of AMR genes and mobile genetic elements, giving the information to identify hotspots and routes of transmission of AMR across the food chain and to carry out quantitative risk assessments (Oniciuc *et al*., 2018; Wang *et al*., 2020) for example, in swine herds (Munk *et al*., 2017). The EFSA BIOHAZ Panel (2019) a SWOT analysis of different uses of metagenomics for risk assessment of food-borne microorganisms, understood metagenomics-based AMR monitoring, was performed concluding that metagenomics can be used for risk assessment of food-borne pathogens, especially in relation ARGs. Although the screening AMR genes in food samples within surveillance schemes are encouraging, there are still difficulties and obstacles to overcome. One of the greater obstacles in WMS reads is the databases used for ARGs which produce a high rate of false negatives. In addition, despite the recognized role of the food production chain in

the spread of ARB, WMS has hardly been used to study AMR in food ecosystems, because the number of available scientific publications is limited. Nevertheless, represents 13% of all WMS related reviewed publications treating of the AMR. For this reason, some studies have followed functional metagenomics approaches to detect novel and differing ARGs in the food chain reservoir (Oniciuc *et al*., 2018). Here are some publications of WMS treating of the AMR. Specifically, Naik *et al*., (2018) carried out an inducted into the microbiota of five marine fish species. *Photobacterium*, *Vibrio*, *Acinetobacter*, *Psychrobacter* and *Flavobacterium* were the most and human pathogens detected. Many different AMR genes were detected, including some on mobile plasmids and Class I integrons, which could be transferred to other bacteria in the food chain. Furthermore, metagenomic approaches (i.e., WMS) can be useful in studies on the use of antibiotics in animal husbandry. AMR genes can spread from this reservoir to the environment and from there to humans, either directly or indirectly via the food chain (Capita and Alonso-Calleja, 2013). Noyes *et al*., (2016a) investigated how the profile of the resistome is widespread in cattle and how these genes are passed along the meat processing chain to retail. Over and above, demonstrated the importance of this type of study to understand the mechanisms that leading to the spread of ARB in foodborne. Pitta *et al*., (2006), using a WMS approach, identified 18 AMR gene classes in dairy agroecosystems, with the most abundant AMR genes such as multidrug transporters, tetracycline, vancomycin, β-lactam resistance determinants, bacitracin. Munk *et al*., (2018) investigated the abundance, diversity and structure of the acquired pig and broiler resistomes in Europe, including Italian farms, through metagenomic shotgun sequencing strategy. This study represents the largest metagenomic AMR monitoring effort of livestock and sequencing effort >5.000 Gb. Were able to observe that the number of unique AMR genes predicted significantly correlated between pig and poultry farms across countries and that the metagenomic resistome varied significantly between the pig and poultry reservoirs, but also within each species, in a country-dependent manner.

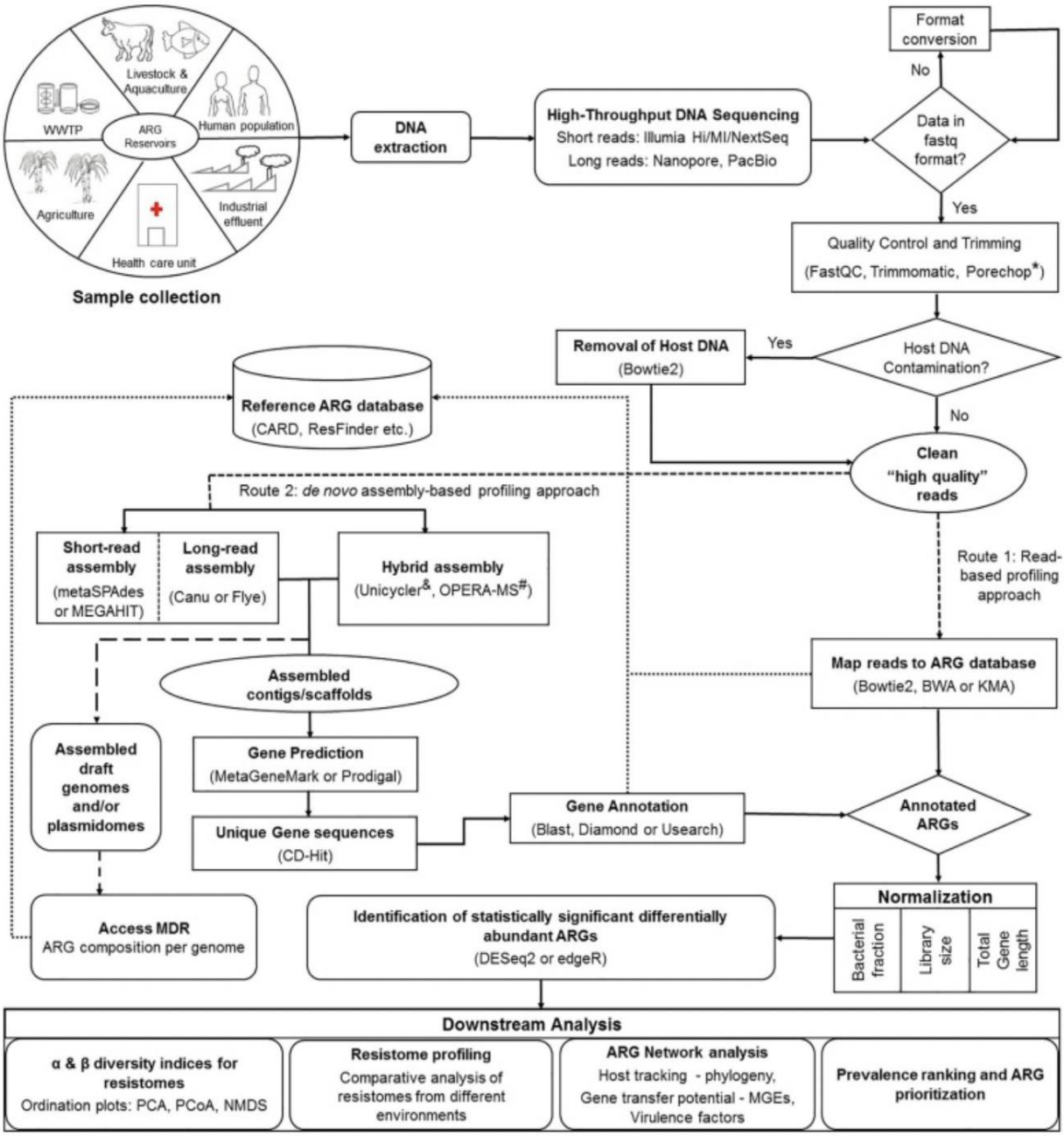
1.8.1 Bioinformatics processing for establish antibiotic resistance in sequencing data

Through the study and characterization of ARGs metagenomes, we are now able to obtain important insights into antibiotic/antimicrobial resistance (AR) supplying novel ecological and epidemiological perspectives. A suite of bioinformatics pipelines and ARG databases are currently available for metagenomic data analyses. However, it is crucial to choose the tools that are most suitable for the specific analysis being conducted, since platforms may significantly vary (Gupta *et al*., 2020).

1.8.2 Quality Control

Using HTS data, an advisable bioinformatics pipeline for determining environmental resistomes is presented in Figure 3. Initially, FastQC and Trimmomatic are used for removing low-quality sequences and contamination (Breitwieser *et al*., 2019). Moreover, if the metagenomics data comes from humans, animals, plants etc. it is important to filter out host sequences before the analysis. After quality control, reads can either be assembled into longer contigs and subsequently mapped (for characterization/quantification of ARGs). In recent times, hybrid assembly approaches that combine more accurate (short) and less accurate (long) read sequences have been implemented to achieve better resolution of AR profiles (Wick *et al*., 2017); in contrast, reads can either be directly mapped to ARG databases.

Figure 3: A workflow for determining the distribution of resistome in complex environments using metagenomic data (Gupta *et al*., 2020).



Tools for specific analysis are shown in the parenthesis. MDR: Multi-drug resistance; * tool for adapter trimming from long reads; & tool for hybrid assembly of short and long reads from bacterial genomes; # tool for hybrid assembly of short and long reads from complex metagenomes.

1.8.3 Read-based vs assembly approaches

Sequencing data can either be subjected to direct profiling (unassembled reads) or de novo assembly-based profiling, for the characterization of ARGs in metagenomic datasets (Chen *et al*., 2019). Read-based approaches have earning attention in clinical ARG surveillance due to their speed and ease of computation. Significantly, read-based approaches allow users to detect ARGs (that might be

undetectable by assembly-based methods) from low-abundance organisms present in complex communities (Boolchandani *et al*., 2019). However, direct mapping of unassembled reads to large data sets can result in high false positive predictions due to spuriously alignment of reads to other genes because of local sequence homology (Jovel *et al*., 2016). High-quality reads are directly aligned to reference ARG database applying pairwise alignment tools such as BWA (Li and Durbin, 2010), Bowtie2 (Langmead and Salzberg, 2012) and KMA (Boolchandani *et al*., 2019). De novo assembly allows better exploration of upstream and/or downstream factors and more accurate detection of protein-coding genes, but can may result in loss of data (Hendriksen *et al*., 2019). The reads are firstly assembled using de novo metagenome assemblers such as MetaSPAdes (Nurk *et al*., 2017) Velvet (Zerbino and Birney, 2008) and MEGAHIT (Li *et al*., 2015), generate contiguous fragments (contigs) and identified protein-coding regions on assembled contigs. After which the predicted protein-coding regions are annotated for resistance determinants and then compared them against antimicrobial resistance reference databases that using similarity-based search tools, such as DIAMOND (Buchfink *et al*., 2015), BLAST (Altschul *et al*., 1990) and USEARCH (Edgar, 2010).

1.8.4 Determining e-values, identity, query coverage levels, cut-off values and data normalization

The stringency of e-values, identity and query coverage levels applied to compare reads and/or contigs to reference databases affect ARG classification. Therefore, choosing a proper sequence identity cut-off becomes important for characterizing a matching read as a resistance gene (Sabino *et al*., 2019). Many of the available ARG databases are biased to genes in clinically associated pathogens and commensals. This stems from the fact that the selection of ARGs (using stringent cut-offs) can increase the probability of targeting genes that are actually functional but may omit environmentally relevant ARGs that can be more diverse. Differently, the selection of ARGs (using less stringent cut-offs) can increase the probability of false positives, by increasing the scope of identified hits to encompass environmentally relevant ARGs (Bengtsson-Palme *et al*., 2017). Even if a read is 100% identical to a documented resistance gene and is short reads, this fragment may be identical to similar genes that do not confer resistance or may be part of a truncated non-functional gene. Annotation of metagenomic samples references a list of putative ARGs and subsequently, the metagenomic data can be analysed to indicate the total abundance and diversity of ARGs. As the number of generated metagenomic reads can vary between samples, the normalization of data is fundamental to make libraries from different samples comparable (McMurdie, 2018).

1.8.5 Biostatistical analysis of metagenomic dataset of ARGs

In recent years there have been significant advancements in statistical methods for analysis of metagenomic dataset of ARGs (Waldron, 2018). Significance in a given sample is based on a cut-off of probability value (i.e., p-value of 0.05 or 0.01). However, the probability of committing false observations increases when multiple tests are multiple performed (Noble, 2009; Chen *et al*., 2017). To characterize significant differences in the abundance of ARGs in metagenomic data Jonsson *et al*., (2016) evaluated the capacity of 14 currently available statistical methods, finding large differences in the performance of the methods; ShotgunFunctionalizeR (software package for R) allows for regression type approaches applying generalized linear models (Kristiansson *et al*., 2009). Such models are also implemented in DESeq2 (Love *et al*., 2014a) edgeR (Robinson *et al*., 2010) packages; IMG/M is based on Gaussian approximation (Markowitz *et al*., 2008) while STAMP focuses on comparisons of pairs of metagenomes using Fisher's exact test, but also Welch's t-test and MetaStats (White *et al*., 2009; Gupta *et al*., 2020).

1.8.6 Abundance and diversity in the resistome analysis

Not only the abundance of resistance genes may be of importance for determining risks, but also the diversity of such genes found (Bengtsson-Palme *et al*., 2017). It is still not very much clear which diversity indices would be preferable measure for total resistome diversity from metagenomic sequencing, because extrapolated resistome data are represented by thousands of different ARGs, and therefore the significant results are dependent on the diversity of associated ARGs (i.e., scope of the resistome) but also on the abundance of individual ARGs. Similar to approaches used for determining differences in the microbiome, the variation in ARG distribution within and between samples is evaluated by calculating beta and alpha diversity indices (Munk *et al*., 2018). Beta diversity indices used to compare the feature dissimilarity between different samples can provide essential data how specific treatments/environments influence of the resistome. To generate a distance matrix between pairs of samples different metrics can be employed such as Bray-Curtis, Canberra and weighted UniFrac (quantitative metrics) for feature abundance data, or binary-Jaccard and unweighted UniFrac (qualitative metrics) for the presence or absence of features (Barwell *et al*., 2015). Furthermore, to assess the significant clustering between groups, are employed non-parametric permutation tests such as PERMANOVA and/or ANOSIM (Pehrsson *et al*., 2016) will come displayed in easy and interpretable 2D or 3D illustrations through principal component analysis (PCA), principal coordinate analysis (PCoA) and ordination techniques such non-metric dimensional scaling (NMDS) (Calle, 2019). For the alpha diversity can be employed indices ecological that combine richness and evenness

(i.e., Shannon index) (Knight *et al*., 2018a) or ACE (Chao and Lee, 1992) and Chao1 (Chao, 1984) sensitive to the number of sequences per sample. Bengtsson-Palme, 2018 has identified that Chao1 index performed very well advising use for estimating resistance gene diversity. In contrast, Shannon, Simpson and ACE indices were not very satisfactory. In literature, it has also been reported that VEGAN (a software package in R) is used to perform beta and alpha diversity analyses (Dixon, 2007).

1.8.7 Web tools to detection of ARG

Based on the user's experience, the detection of the ARG can be reached on commercial platforms with open access accessible online or downloadable for installation and usage. The most applied and currently suitable open access pipelines for identifying ARGs in metagenomic data are: ARGs-OAP, Resistance Gene Identifier (RGI), Graphing Resistance Out of meTagenomes (GOOT), Mutation Mapping in Metagenomes (Mumame) and fARGene. Below, the online bioinformatics pipelines that were used in this thesis project are described. **ARGs-OAP** uses its own non-redundant ARG database, Structured ARG reference database (SARG) constructed from CARD (Jia *et al*., 2017) and ARDB (Liu and Pop, 2009) with a hybrid BLASTX and UBLAST algorithm developed for rapid annotation and classification of ARG-like sequences from metagenomic data (Yang *et al*., 2016b). **RGI,** applies one resistance detection approach based on functional resistance homologues (Protein Homologue Model) and one resistance detection approach based on detecting mutations conferring resistance in sensitive targets (Protein Variant Model). RGI uses CARD's being organized AR detection model to predict ARGs, to predict intrinsic antimicrobial resistance genes and can be used within the conda open-source package through the Bioconda project (https://bioconda.github.io) (Jia *et al*., 2017; Boolchandani *et al*., 2019; Gupta *et al*., 2020).

1.8.8 ARG databases, online public platforms

ARG databases along-with web tools for sequence analysis and annotation of ARGs offer antimicrobial resistance related reference data. The selection of reference databases has consequences for the quality of the information obtained. For this reason, annotation based on bioinformatics analysis of sequence similarity will never be more accurate than that of reference sequences, so it is fundamental to select a reference database with high-quality annotations (Bengtsson-Palme *et al*., 2016). These databases contain the function of genes and describe phenotypic information accumulated from various studies that include antimicrobial susceptibility testing of bacteria harbouring specific antimicrobial resistance genes. A number of public databases containing ARGs information exist may vary in the scope of the resistance mechanisms and the type of annotations that they contain (Bengtsson-Palme *et al*., 2017; Arango-Argoty *et al*., 2020; Gupta *et*

al., 2020). For these reasons, can be divided into **specialized AR databases** and **generalized AR databases**. Specialized AR databases were created to provide complete information for specific gene families such as β-lactamase encoding genes (Boolchandani *et al*., 2019) a family of antimicrobial resistance enzymes that facilitate hydrolyzation of the key β-lactam rings in β-lactam antimicrobials, thus protecting the bacteria from the antimicrobial activity. Specialized AR databases incorporate the Comprehensive β-lactamase Molecular Annotation Resource (CBMAR) (Srivastava *et al*., 2014), Lactamase Engineering Database (LacED) (Thai *et al*., 2009) and β-lactamase database (BLDB) (Naas *et al*., 2017). The CBMAR β-lactamases database includes molecular and biochemical information that could divulge understanding of known and novel β-lactamases. LacED offers integrated tools for sequence analysis and features structural data specific for SHV and TEM β-lactamases (Boolchandani *et al*., 2019; Gupta et al., 2020). Generalized AR databases incorporate a wide range of ARGs and mechanistic information and include MEGARes (Lakin *et al*., 2017), Resfinder (Zankari *et al*., 2012), Resfams (Gibson *et al*., 2015), Comprehensive Antibiotic Resistance Database (CARD) (Alcock *et al*., 2020), Antibiotic Resistance Genes Database (ARDB) (Liu and Pop, 2009), Antibiotic Resistance Gene Annotation (ARG-ANNOT) (Gupta *et al*., 2014), Functional Antibiotic Resistance Metagenomic Element (FARME) database (Wallace *et al*., 2017), SARG(v2) (Yin *et al*., 2018), Mustard (Ruppé *et al*., 2019) and Antibiotic Resistance Gene miner (ARG-miner) (Arango-Argoty *et al*., 2020). Among the ARG databases, **ARDB** and **ARG-ANNOT** are currently archived. ARDB database is the first established manually curated resource of ARG sequences established in 2008. Each gene was annotated to include resistance mechanism, resistance type, COG characterization and ontology but its last update was in July 2009, meaning that any resistance gene discovered after that date is not included in the database. However, all the ARDB sequences were integrated into CARD. In addition, ARDB does not make any distinction between resistance genes with a confirmed resistance function and those predicted to confer resistance based on homology, for this it is possible to define it as a database containing, therefore, sequences that in fact are not functional resistance genes (Bengtsson-Palme *et al*., 2017; Gupta *et al*., 2020). Similar problems also haunt the **ARG-ANNOT** database, however, employs the relaxed search criteria to identify resistance genes, meaning that the database includes many sequences with poor annotation information and that many entries are unlikely to be functional resistance genes by limiting the identification of true resistance genes (Bengtsson-Palme *et al*., 2017). **ARG-miner** is a resource for the inspection and curation of ARGs based on crowdsourcing. Integrates multiple ARG databases including SARG, CARD, etc. into a common nomenclature by removing redundant information (Arango-Argoty *et al*., 2020; Gupta *et al*., 2020). **MEGARes,** can easily detect AR determinants in large metagenomic datasets. It has an accessible, user-friendly service (called AmrPlusPlus) for analyzing metagenomic

data. At the nucleotide and protein levels, each entry in MEGARes has been manually validated and comprises several sources, including the curated CARD database (Bengtsson-Palme *et al*., 2017; Doster *et al*., 2019).

Resfinder (Zankari *et al*., 2012) is one of the oldest databases, a web-based and standalone tool that keeps its sequences up to date and only contains sequences of acquired antibiotic resistance genes that also accept short reads as input for comparison against known acquired resistance genes. In addition, it extracts information from other databases (i.e., ARDB and Lahey database) and also from published literature, including reviews (De Abreu *et al*., 2021). The **Mustard** (Ruppé *et al*., 2019) antimicrobial resistance determinants database, characterizes ARGs based on three-dimensional protein structures to help predict resistance genes, suggesting higher sensitivity (Boolchandani *et al*., 2019; Gupta *et al*., 2020). **FARME** is the first ARG database that comprises a curated set of microbial sequences functionally screened to confer resistance in various functional metagenomics studies in environmental samples. One of the benefits of the FARME database is that it contains over seven times the number of non-redundant protein sequences compared with other ARG databases (i.e., CARD and ARDB) and also includes regulatory elements, predicted proteins flanking antimicrobial resistance genes and mobile genetic elements (Boolchandani *et al*., 2019; Gupta *et al*., 2020). **CARD** database focuses on providing high-quality reference data and molecular sequences within a controlled vocabulary, the Antibiotic Resistance Ontology (ARO); is the most comprehensive resource for ARGs information available and uses its own tool RGI, which employs curated AR detection models. This resource manually curated ontology-based provides extensive information on ARGs and their resistance mechanisms. CARD curation occurs monthly with an interplay of text mining and contain a user-friendly graphical interface that provides function-based classification of ARGs (Jia *et al*., 2017). In CARD, the use of a single reference sequence for every resistance gene increases the likelihood that each sequence has been confirmed to confer resistance in at least some species (Bengtsson-Palme *et al*., 2017; Boolchandani *et al*., 2019; Alcock *et al*., 2020; Gupta *et al*., 2020). For AGR identification across less commonly studied bacteria are used Hidden Markov model-based databases (HMM). **Resfams,** a curated resource of protein families that are linked to their profile HMM associated with antimicrobial resistance functional, derived from multiple AR protein sequence alignments collectively acquired from the LacED, CARD, and Lahey databases. Relative to the other platforms concentrated on pathogen associated ARGs, Resfams provides an ecological/environmental overview of the resistome. For this reason, it can identify a greater number of novel ARGs and remote homologues of known ARGs than other databases such as CARD and ARDB that rely on BLAST-based methods for gene identification. This increased sensitivity relative to other ARG databases demonstrates the versatility of the HMM (Boolchandani *et al*., 2019; Gupta

et al., 2020). It is important to emphasize that the choice of databases can influence the interpretation of risk associated with AMR in public health and thus, it is important to keep these databases up to date and standardize them for further research (Bengtsson-Palme *et al*., 2017).

1.9 Food chain: the role of the environment as source of antibiotic resistance

As recognized by all international organizations and WHO, antibiotic resistance is currently a major global threat, with estimates of 33.000 annual fatalities human in the EU (Cassini *et al*., 2019). Since 2011, the European Commission (EC) published covering the 2011-2016 period, its first Action Plan against the rising threats from AR, including the intention to contain the risks of spreading AR via the environment (Smith *et al*., 2016). The EC action plans were building on addressing the threat from a holistic approach (One Health approach), also considering animal, environmental and human sectors. Nevertheless, the international organizations such as FAO, OIE and WHO have also recognized the need to further investigate (http://www.who.int/foodsafety/areas_work/antimicrobial-resistance/tripartite/en). In the last years, increasing importance has been given to the role of the environment as a source of ARGs for both animals and humans, but still their uncertainties about the role played by the spread and persistence of AR. Food-producing environments are spaces where the food of animal origin is produced/processed, first at the primary production level (i.e., preharvest) and later processing level (i.e., slaughterhouses, processing plants) (postharvest). In general, these environments can be contaminated by ARG/ARB deriving from several environmental sources, such as from terrestrial/aquatic food-producing animals and residues from post-harvest food plants (i.e., slaughterhouses and food processing plants). In this way, if ARG/ARB contaminate food-producing spaces, they can spread throughout the food chain through several routes and constitute a threat to public health (EFSA BIOHAZ Panel, 2021).

It is approved that the usage of antimicrobials, (AMU) is an important factor for the selection/diffusion of antimicrobial resistance in food-producing systems. In general, the major risk factor of AMR is related to antimicrobial use on the farms (EMA and EFSA, 2017; Jayarao *et al*., 2019). For this reason, in the food animal sector in the EU, the sales of antimicrobials declined by 34.6% in the period between 2011 to 2018 (EMA, 2020; More, 2020). Regardless from prudent of AMU the most important measures to limit AMR, both at pre- and post-harvest, involve the correct implementation of effective measures such as hygiene practices and biosecurity. We can say that ARB/ARGs present in food are related to the AMU, the livestock management and density (i.e., housed or free-range animals), type of commercial feeds, the influence of waterborne pollution and management strategies in the production environment. Nevertheless, the antimicrobial resistance in food can associate with processing methodologies and the probability of faecal and environmental

contamination of the final food item. Li *et al*., (2020) examined the microbiome of the chicken breasts through shotgun metagenomic sequencing to allow ARG profiling and to inform multiple food safety and quality characteristics. The samples represented different production and processing practices, including "No Antibiotics Ever" (ABF) versus conventional (CONV) products. This study suggests that resistome comparison between CONV and ABF samples identified no significant difference in ARG abundance and composition between the two production practices. Furthermore, ARG abundance of the chicken breast samples was considerably lower than that of any other livestock samples and environmental samples analysed in the current study. This is demonstrated clearly these results indicate a low risk of ARG accumulation on chicken breasts regardless of antibiotic usage in live production. Finding is consistent with AR surveys in beef production examined by Noyes *et al*. (2016b). This study suggests that no ARG was detected on beef products and that interventions during slaughtering and beef processing might reduce the risk of ARG transmission to consumers.

1.10 Sources and transmission routes AMR in the poultry production sector

The total poultry population within 2030 will reach 8.5 billion, defining the poultry production as one of the most rapidly expanding global industries (DeSa, 2015) with 13.3 million tonnes of poultry meat (carcass weight) https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agricultural_production_-_livestock_and_meat&oldid=470510#Poultry (Eurostat). In 2019, poultry meat production was largest in France, Germany, Italy, Spain and Poland (Eurostat, 2021). Antibiotic resistance bacteria and then antibiotic resistance genes may be found in meat, in particular from those regions (poultry meat exporters) where the control of antimicrobials is less present (Rozman *et al*., 2019; Rabello *et al*., 2020; EFSA BIOHAZ Panel, 2019; Xu *et al*., 2020; EFSA and ECDC, 2021).

Previously to Regulation (EC) No 1091/2005, antibiotics were used in the EU to eliminate *Salmonella* infection at the farm level and in meat broiler prior to slaughter. Nevertheless, despite dramatically reduced usage in most EU countries, routine use selected other bacteria such as *Campylobacter* and fluoroquinolone resistant *E. coli* which are still circulating in the poultry industry (Mouttotou *et al*., 2017; Perrin-Guyomard *et al*., 2020). Nhung *et al*., (2016) showed that AR commensal bacteria and zoonotic pathogens are widespread in poultry farming and at the production level for two fundamental reasons, both due to persistence in the farming environment (due to suboptimal cleaning conditions and disinfection) and in response to the current selection given by drugs, and which both reasons may select for AMR in surviving bacteria. Aspects of biosecurity/farm hygiene have also always been identified as factors responsible for the introduction/environmental persistence of AMR on farm, even date there are some gaps in understanding of the risk factors and the most effective interventions

(Davies and Wales, 2019). However, the intensity of antimicrobial use is the major driver for complex patterns of resistance and associated metagenomes on poultry farms (Pesciaroli *et al*., 2018; Xiong *et al*., 2018). The main factors that contribute to the introduction of AMR and allow the transfer of AMR between different food production systems and in food products from terrestrial animals, include pre-harvest and post-harvest factors. Often, pre-harvest data based on the presence of AMR in various sources, but often not on transmission of AMR from such a source to a food production system. These sources could introduce AMR into food production environments. Differently, in the post-harvest, the information can be only found for the transmission of pathogens pretty than transmission of ARB and ARG. In the following, the factors contributing to the introduction of AMR from environmental sources (the production pyramid environment) into poultry sector are presented.

1.10.1 Pre-harvest factors

Mixed farms or co-grazing: in poultry farms in which co-grazing occurs (i.e., with sheep, horses and cattle) animals may be potential sources of ARB in the food-producing environment, as they can be a reservoir for different pathogens. One study showed that cattle were potential reservoirs of *Campylobacter* for commercial broilers (Frosth *et al*., 2020). **Wildlife, rodents and arthropods:** arthropods and wildlife have been shown to source and transmission route of AMR (Darwich *et al*., 2019; Dolejska, 2020). Insects and flies attracted to manure stores, or the droppings pits can be vectors of ARB and ARGs (Zurek and Ghosh, 2014; Poudel *et al*., 2019). Migratory birds can transport ARB and ARGs internationally (Cao *et al*., 2020). Starting from the fact that wildlife species can acquire ARB from flies (Royden *et al*., 2016) but are relevant sources of ARB as they are more likely to be able to access housing and feed or bedding stores (Jahan *et al*., 2021). All in all, from the limited studies of the wildlife microbiome the likelihood of a particular species being acquired within the microbiome by wildlife species and the ability to transmit it locally is unknown, limiting efforts to assess AMR dissemination. **Waste, litter, soil and surface water in poultry farm:** with poor cleaning and disinfection of the house surrounding and litter or waste from other animals, external environmental contamination with *Campylobacter* that carry varying AMR determinants around poultry farms can happen (Graham *et al*., 2009; Battersby *et al*., 2017). In species in which bedding is frequently used (if contaminated with faecal material), can be also a potential source of AMR, some studies, especially chickens, have emphasized the role of litter as a vector where ARB from faeces persists in large numbers and becomes source of contamination for animals which come into contact with this ARB-contaminated bedding (Yang *et al*., 2006). The application of animal manure as a fertilizer and sewage sludge application can introduce AMR into soils, so leading to dissemination of ARB and ARGs of animal origin to the soil. Subsequently, from the soil, the ARB and ARGs can

enter different production systems, not limiting the transmission within-sector but between sectors. Furthermore, the ARB and ARGs from manure can spread to surrounding watercourses and if farm animals are exposed to waterways contaminated exposure to ARB and ARGs that originated from another animal sector can occur (Moore *et al*., 2010), but there is no evidence of the impact of this on the EU level. **Drinking water and feed:** drinking water can be a source of ARB entry into farms, especially if the drinking water is drawn from a well or from local surface waters (Tanner *et al*., 2019; O ' Dwyer *et al*., 2018). Feed contamination has been primarily identified for pathogens rather than for ARB or ARGs, for this reason, feed processing practices can result in the presence of pathogens and ARB in feed during cooling, transportation, storage and finally, on farm (Sapkota *et al*., 2007). Heat treatment reduces the risk of the presence of pathogenic bacteria and consequently reduces the risk of AMR (Torres *et al*., 2011). **Dust/air:** with respect to AMR in dust and air samples, research has focused on the presence of ARB and ARGs within animal housing and in the nearby environment (Luiken *et al*., 2020). In the literature, it has been reported that ARB can be found in the air within and around poultry houses, including during the empty inter-crop period (albeit at lower levels than during the life of the flock) and can disseminate organisms over a wide area (where powerful extractor fans are used) (Brooks *et al*., 2010). Poultry house manure, dust, or litter includes a matrix of faecal bacteria, nutrients for microbial growth such as feed additives and antimicrobial agents, and heavy metals (Deng *et al*., 2020). In addition, dust contaminated by ARB is also created after the spreading of poultry litter and manure on land, such as on poultry farms observed in the UK (APHA/Defra, 2020). **Human:** human as a breeder, veterinarian or worker is a primary source of ARB for farm animals, involving organisms such as LA-MRSA (Monecke *et al*., 2013).

1.10.2 Post-harvest factors

In many European countries, most of the poultry meat used for processing is imported from third countries (https://ec.europa.eu/info/food-farming-fisheries/animals-and-animal-products/animal-products/poultry_en), where problems with the widespread emergence of AmpC resistance are increasing in the Salmonella Heidelberg, representing a threat of introduction (vertically transmitted) into the EU poultry production environment (Souza *et al*., 2020). Data on the introduction and spread of AMR relate also to pathogen bacteria, such as *Salmonella* or *Campylobacter* (if no information is available on AMR). A wide variety of ARB and ARGs have been marked from the poultry intestinal tract and, to a lesser extent, the poultry farm environment and production chain, suggesting a relationship between the finding of the same organism in the poultry environment and its occurrence in birds. These AMR threats have rarely been described in poultry environments in European countries. Although there is a small amount of information on the occurrence of AMR, so all current

evidence is based solely on observations of potential correlations. Most of post-harvest studies are focused on the study of carcass contamination and therefore are necessary more research is needed to evaluate the impact of the post-harvest environment on meat contamination of ARB. **Transport:** a fundamental route of transmission could be transport, in a particular vehicle, crates and modules. Trucks usually visit different farms (mixing animals from different origins) and can spread pathogen contamination between farms where partial the thinning is practiced, between slaughterhouse and poultry farms (Rasschaert *et al*., 2020). **Slaughterhouse:** the lairage area, where several batches of animals are housed within the same structures, may be a source of AMR for the animals entering the abattoir, furthermore, the animals can come into contact with ARBs or pathogens that are present in the environment. Therefore, there is a risk of introducing ARBs and pathogens into the slaughterhouse environment (i.e., from contaminated skin and intestines) as their load can be high within this area (EFSA BIOHAZ Panel, 2021). **Slaughter line and equipment:** all activities carried out within the slaughterhouse can introduce/spread pathogens through the carcasses into the food chain and thus can act as a source of antimicrobial resistance. Broilers transport ARBs on the integument and internally even when antimicrobials have not been given that's why cross-contamination of poultry at slaughter contributes helps to greater microbial diversity in retail chicken than in live birds (Althaus *et al*., 2017; Montoro-Dasi *et al*., 2020). Even the semi-automated or automated processes are possible sources of disseminating ARB. For example, in the scalding, defeathering and evisceration stages they are the main sources of release of microorganisms in the slaughtering environment and of carcass contamination via spillage of intestinal contents and there is no stage of the process that can reduce contamination of carcasses and indirectly AMR transmission (Pacholewicz *et al*., 2015; Rasschaert *et al*., 2020). A study by Cornejo *et al*., (2018) argued that despite the drug withdrawal period being observed, feathers can serve as a means of introducing antimicrobial contamination into the poultry slaughter process. **Waste, wastewater and air:** contamination of wastewater, slaughtering process water, the abattoir waste, the airflow and aerosols from slaughter facilities may also introduces and spread of ARB and ARGs. (Savin *et al*., 2020). Slaughterhouse waste (low-grade category 3) are used for the manufacture of feed for farmed fish and pets, indicating further dissemination of AMR beyond the food chain (Groat *et al*., 2016). Occupational transmission of **workers** through their hands or equipment (i.e., cross-contamination) may act as sources of AMR and so how sources of ARB and ARG, influencing the further processing stages for chicken carcasses at slaughterhouses (Mulders *et al*., 2010; Van Gompel *et al*., 2020), often with disinfection and cleaning procedures being inadequate to eliminate bacteria contamination between working days, as shown for pathogens (Samapundo *et al*., 2019; Obe *et al*., 2020). In addition, mitigation measures must be implemented to reduce contamination during transport, lairage and post-harvest processing.

For example, during transport and lairage, management measures (batch separation), good hygiene practices and disinfection may reduce the spread of ARB (Obe *et al*., 2020). Differently, in the post-harvest environment are based on ordinary general measures such as good manufacturing practices and hygiene (are not specific for ARB/ARGs). Even though useful, such measures require validation regarding their impact on AMR (West *et al*., 2018).

1.11 Targets for monitoring of AMR

AMR monitoring in zoonotic and indicator bacteria in food-producing animals and their food products involves continuous data collection, analysis and reporting. In addition, allows you to comprehend the development/diffusion of resistance, to follow temporal trends in the occurrence/distribution of AMR and well as provide risk assessment data, and evaluates targeted interventions. The European Union Summary Report on AR in zoonotic and indicator bacteria from animals, food and humans in 2018/2019 gives the findings of the AMR monitoring in the food-producing animal populations (carcasses and meat) and in humans. Data on AMR in zoonotic and indicator bacteria are collected annually from EU MS, and analysed by EFSA and ECDC. The 2018 monitoring and reporting has been concentrating on poultry, while the monitoring and reporting 2019 has been concentrating on pigs and calves (under 1 year of age) and included data regarding resistance in zoonotic *Campylobacter* and *Salmonella* from animals and food and humans, and resistance in indicator *Escherichia coli* and meticillin-resistant *Staphylococcus aureus* (MRSA) from food and animals. In line with Commission Implementing Decision 2013/652/EU, monitoring of AMR from a public health perspective is obligatory in *Campylobacter jejuni*, *Salmonella*, and indicator commensal *E. coli* in the domestically produced animal populations and their derived meat (EFSA and ECDC, 2021b).

2. Aims of the work

Food safety and the protection of consumer's interests is of increasing concern to the European community, including policy makers, food business operators (FBOs), citizens, non-governmental groups, professional associations, international trade organizations. The SarsCov2 pandemic increased the concern on emerging food safety risks and threats, highlighting that early identification of existing as well as emerging risks is at the heart of protecting public health and the environment, while avoiding both short and medium term negative economic and societal impacts. In this framework untargeted throughout analytical methods able to detect and mitigate all microorganisms within a food system can help to quickly detect the increase of specific biological hazards in foods.

The **main objective** of this research project has been to apply shotgun metagenomic sequencing to investigate both microbiome and resistome of foods of animal origin in order to assess advantages and disadvantages of shotgun metagenomic sequencing in comparison to the current analytical methods used to guarantee food safety and map ecosystems associated to food systems.

In order to achieve this main objective four studies have been performed with 5 specific objectives (**SO**). To contribute to assess the suitability of shotgun metagenomics to detect a wide range of target microorganisms in foods, in the study 1 a proficiency test (PT) was organised as part of the COMPARE project which used to be a multidisciplinary research network that has the common vision to become the enabling analytical framework and globally linked data and information sharing platform for the rapid identification, containment and mitigation of emerging infectious diseases and food-borne outbreaks. Within the activities scheduled in COMPARE study 1 of my research project aimed at characterising and quantifying microorganisms belonging to different domains experimentally spiked in cold- smoked salmon at known concentrations, using shotgun metagenomics (**SO1**). To analyse the metagenomic results, many European and global institutions use internal pipelines which are not publicly available or pipelines which are in the public domain but combined in an unknown way. Among the few bioinformatic tools publicly available there are MG RAST (Keegan *et al*., 2016), which is public and free (https://www.mg-rast.org); OneCodex (Minot *et al*., 2015) (www.onecodex.com) and CosmosID (Yan *et al*., 2019) (https://app.cosmosid.com/) which are public but not free; MGmapper (Petersen *et al*., 2017), hosted at the CGE, now call CCMetagen 1.0 (https://cge.cbs.dtu.dk/services/MGmapper/) which is public, free but not always updated in the web version. The outputs obtained testing the metagenomes of study 1 using these different bioinformatic tools were comparatively assessed to evaluate the suitability of the different tools to detect all microorganisms spiked in the salmon (**SO2**).

In 2019 the European Union (EU) produced an estimated 13.3 million tonnes of poultry meat, representing an increase of around 27% in comparison to 2010 (Eurostat, 2121). Poultry meat is characterized by high-quality proteins, vitamins, and minerals important for the human diet (Marangoni *et al*., 2015). Since the poultry rearing cycle lasts 35 to 42 days, poultry meat can be produced without the use of antimicrobials more easily than pork and beef meat having rearing cycles of months. Moreover, the mean values, expressed in number of defined daily doses (DDDvet)/biomass for poultry, of antimicrobial agents obtained from the technical estimates of the sales of veterinary antimicrobials in the European Union in 2016 were 0.5 for poultry versus 1.3 for pigs (ECDC, EFSA, EMA, 2021). There are different strategies to achieve antibiotic free poultry flocks as the implementation of effective biosafety measures and management options as well as promoting beneficial microbes in the chicken gastrointestinal (GI) tract of chicken to enhance animal health and inhibit pathogen colonization. To this aim feed can be supplemented with probiotics and prebiotics, also blended in the same supplement (i.e., a symbiotic), to ensure diversity and stability of the GI microbial community, as well as positive interactions with host's gastroenteric epithelium and immune system (Brugaletta *et al*., 2020). As an alternative, probiotics can be also supplemented in the litter and up take by the animals (De Cesare *et al*., 2019). Whenever a poultry disease occurs in an antibiotic free flock the animals are treated with antibiotics and the flock turns into conventional. Therefore, poultry farms can occasionally rear antibiotic free flocks along with conventional ones. In the studies 2 and 3 it was investigated whether the efforts of raising chickens without the use antibiotics make any difference in the microbiome of poultry meat consumers eat (**SO3**). To this aim in study 2 I compared the microbiomes characterizing caeca and corresponding carcasses of two groups of chickens reared in one conventional and one antibiotic free intensive farm. Moreover, I investigated if the correlation between the microbiome and resistome in the caeca and the carcass of the same animal provide more insights than the same analysis performed at flock level before planning future studies (**SO4**). In study 3, additional samples of poultry carcasses reared in antibiotic free and conventional flock were tested within the EU project CIRCLES which has the overall aim to investigate how the microbiomes associated to poultry food system interact and impact one with the other.

To date, the microbial composition associated to fermented foods has been investigated using amplicon-based sequencing (Knight *et al*., 2018b; Diaz *et al*., 2019) and few studies have applied shotgun metagenomic sequencing to obtain a higher insight into the taxonomy and functional potential of fermented foods (Wu *et al*., 2017; Walsh *et al*., 2020). Within the PRIMA project ArtiSaneFood, in study 4 I started the metagenomic investigation of an Italian artisanal product and

the question I tried to answer was if testing one aliquot of artisanal food homogenate is representative of the whole homogenate (**SO5**).

3. Materials and methods

3.1 Materials and methods of study 1

Samples tested

The samples tested in the Study 1 were 3 aliquots of 0.2 g of Salmo salar, cold-smoked and vacuum packed. Each aliquot was spiked with 50 µL of a mock microbial community consisting of six bacteria (*Staphylococcus aureus*, *Propionibacterium freudenreichii*, *Bacteroides fragilis*, *Escherichia coli*, *Fusobacterium nucleatum* and *Salmonella enterica*), one parasite (*Cryptosporidium parvum*), one fungus (*Saccharomyces cerevisiae*) and 10 µL of heat-inactivated virus (*Bovine alphaherpesvirus* 1) (Table 4). The spiked samples were shipped frozen by the Technical University of Denmark (DTU) to the University of Bologna for a proficiency test organized within the H2020 COMPARE project. At the arrival, the 3 samples were submitted to DNA extraction.

Table 4: Combination of the mock community used to spike the samples of cold-smoked salmon and concentration of each microorganism.

Taxon	Feature	Quantity per subsample (cells/virus gene copies)
Bacteria		
Staphylococcus aureus subsp. *aureus* NCTC 8325	Gram +	500.000.000
Propionibacterium freudenreichii subsp. *freudenreichii* DSM 20271	Gram +	500.000.000
Bacteroides fragilis NCTC 9343 / DSM 2151	Gram -	50.000.000
Escherichia coli ATCC 25922	Gram -	50.000.000
Fusobacterium nucleatum subsp. *nucleatum* ATCC 25586 / DSM 15643	Gram -	50.000.000
Salmonella enterica subsp. *enterica* serovar *Typhimurium* str. ATCC 14028S / DSM 19587	Gram -	50.000.000
Parasite		
Cryptosporidium parvum IOWA II isolate		1.000.000
Fungus		
Saccharomyces cerevisiae S288C		5.000.000
Virus		
Bovine alphaherpesvirus 1	Ds DNA visus	1.20E+10

DNA extraction from smoked salmon

Total DNA from each cold-smoked salmon sample was extracted using PowerFood® Microbial DNA Isolation Kit (MoBio-Qiagen) according to manufacturer's directions with some modifications.

Briefly, each sample was suspended in 1 ml of sterile physiological solution (NaCl 0.90%) and homogenized in vortex for 1 minute. At the end of the homogenization step, the liquid part of the sample was transferred inside a 2 ml tube and centrifuged at room temperature for 5 minutes at 11000 rpm. Once the supernatant was removed, the pellet was suspended again in 1 ml of sterile physiological solution (NaCl 0.90%), homogenized in vortex for 1 minute and centrifuged at room temperature for 5 minutes at 11000 rpm. At the end of the centrifugation step, the formed pellet was suspended in 450 ml of PF1 solution. Subsequent steps of the extraction protocol were performed in accordance with the handbook that came with the extraction kit.

Quantification of total DNA

The total DNA isolated from each cold-smoked salmon sample were quantified on a BioSpectrometer® (Eppendorf, Milan, Italy) and Fragment Analyzer Automated CE System (Advanced Analytical). In particular, in the first quantification obtained through BioSpectrometer® (Eppendorf, Milan, Italy), was assessed DNA yield, in terms of quantity and quality. The parameters of interest were the concentration of DNA in µg / ml, an absorbance value at 260 nm and the value of DNA quality that was assessed in terms of absence of contaminants according to the value of the A260 / A280 nm ratio, used for indication of sample purity. This protocol is generally accepted for DNA with absorbance ratio values of 1.8-2.0, which indicates a pure DNA sample (Manual instruction PowerFood Microbial DNA Isolation Kit). All samples complying with DNA quality scores were subjected to library preparation.

Library preparation and metagenomic sequencing

The DNA extracted and assessed for quality and quantity was submitted to the library preparation procedure with the Nextera® XT DNA Library Preparation Kit (Illumina, San Diego, CA). Illumina's Nextera technology provides an input DNA enzymatic fragmentation (termed tagmentation) and adds partial adapter sequences (in a single step). During tagmentation, the transposase enzyme fragments the DNA and simultaneously adds specific adapters to both ends of the fragments, that allow for subsequent PCR amplification to introduce index sequences (barcode) (McElhoe *et al*., 2014). Nextera® XT DNA Library Preparation Kit was chosen because this method improves traditional protocols, is the fastest for sample preparation for any Illumina sequencing platform, reduces the sample handling and finally allows to analyse of up to 96 samples (Marine *et al*., 2011; McElhoe *et al*., 2014).

In particular, after to the second quantification obtained through Fragment Analyzer Automated CE System (Advanced Analytical), the DNA was diluted, in molecular-grade water, to the concentration

required (0.2 ng/µL) through others intermediate dilutions (i.e., 10 and 2 ng/µL). After the DNA tagmentation step with Amplicon Tagment Mix (ATM), Tagment DNA Buffer (TD) and Neutralize Tagment Buffer (NT), the samples were amplified using a limited-cycle PCR program in which the sequencing primers (Index 1 and 2) and adapters were inserted. Before moving to the next step of library quantification, a "clean up" step was carried out using the AMPure XP beads, to purify the library DNA and remove short library fragments. Once the libraries have been prepared, it was necessary to validate the concentration, nano molarity and average length of the DNA with Fragment Analyzer Automated CE System (Advanced Analytical) and Quant-iT™ PicoGreen^{â} dsDNA Assay Kit and through Chip DNA Hi Sensitivity analysis on Bioanalyzer 2100 (Agilent Technologies). Typical libraries indicate a broad size distribution of ~250-1000 bp. Various libraries can be sequenced with average fragment sizes as small (250 bp) or large (1500 bp) (Nextera XT DNA Library Prep Kit). Each library pool was adjusted to a nano molarity of 4 to be sequenced. A total of 3 µL of each library (4 nM) were pooled together. Each pool was loaded into a flow cell of the glass slide. Each fragment of DNA library was anchored on complementary oligo-adapters placed on the flow cell and clonally amplified through a solid-phase amplification called bridge amplification and then sequenced by synthesis. Libraries were sequenced using the NextSeq 500 sequencer (Illumina) with NextSeq 500/550 High Output Reagent Kit v2 (300 cycles) (Illumina) at 2x150 bp in paired-end mode. The metagenomes were featured by an average output of 7 Gbp.

Bioinformatics and statistical analysis

Filtering, trimming, and taxonomic classification of raw reads were analysed using 4 different bioinformatics software represented by MG-RAST (Keegan *et al*., 2016) (https://www.mg-rast.org), MGmapper (Petersen *et al*., 2017) (https://cge.cbs.dtu.dk/services/MGmapper/), CosmosID (Yan *et al*., 2019) (https://app.cosmosid.com/) and OneCodex (Minot *et al*., 2015) (www.onecodex.com). In MG-RAST, the abundances of taxonomic composition were performed using the RefSeq reference database (NCBI reference sequences) (Pruitt *et al*., 2005), RDP (Ribosomal database project), Silva LSU (large subunit-23S/28S), Silva SSU (small subunit-16S/18S) and Greengenes and narrowing down the analysis only to those annotations marked as Bacteria. In all databases, the following quality parameters were set: maximum e-value cut-off <1e-5, minimum identity cut-off 60%, and minimum alignment length cut-off 15 bp. In MGmapper the database selected was Silva and in OneCodex the One Codex database. Finally, for CosmosID the GenBook database was used with the following parameters set: taxonomy switcher and results-total. The results of abundance of each taxonomic level for each sample were analysed using the Statistical Analysis of Metagenomic profile Software v 2.0.9 (STAMP) (Parks *et al*., 2014). The statistical differences between the outputs of different

bioinformatics tools were not assessed because only three samples were available for each combination of databases.

3.2 Materials and methods of study 2

Animals and Treatments

In this study a total of 30 broilers were investigated: 15 were reared in a conventional farm and 15 belonged to a poultry flock reared in a conventional farm but never treated with antibiotics. The farming types did not differ with respect to animal age at slaughter. All tested animals were female, ROSS, fed with no medicated feed and slaughtered at 47 (antibiotic free) and 48 (conventional) days. According to the European legislation, in the conventional production a treatment with antibiotics was performed for therapeutic purposes upon veterinary prescription. On the contrary, any antibiotic was not administered to the animals reared in the antibiotic free farm. In the group reared in the conventional farm amoxicillin was administered at 8 and 29 days at the recommended doses of 20–30 mg/kg of live weight in drinking water for three to five days and sulfadimethoxine/trimethoprim was administered at 24 days as an antimicrobial for the coccidiosis treatment and/or prevention, according to the manufacturer's recommendation (100 mg/20 mg in 1–2 L of drinking water once a day for five days). All 30 slaughtered broilers were randomly sampled at the slaughterhouse within the first group of animals processed at the start of the working day, when the slaughterhouse was still clean and disinfected to avoid bias due to cross-contamination.

Sample collection

In each sampled animal both the gastrointestinal tract and the carcass were collected. For each selected animal, the GIT was collected at the evisceration step (during slaughtering,) and immediately stored at 4°C in a refrigerated box; at the same time a sterile plastic flag was attached to the hook transporting the carcass from which the GIT was collected in order to pick up that specific carcass at the end of the refrigeration tunnel. After sampling, the carcass was kept refrigerated in a box different from that containing the GIT samples. All samples were transported to the laboratory and processed. Each gastrointestinal tract was dissected out and a small portion (i.e., 0.5 to 2 g) of caeca content was collected, transferred into a 2 ml sterile plastic tubes, flash-frozen in liquid nitrogen before storage at -80°C until DNA extraction (De Cesare *et al*., 2017). In a sterile environment, approximately 10 g of neck and breast skin were collected from each carcass and placed inside a sterile bag. Subsequently, each sample was diluted in 90 ml of sterile physiological solution (NaCl 0.90%), homogenized in the stomacher (MAYO homogenius HG400V) at normal speed for 1 minute. At the end of the

homogenization step, all the liquid part of the sample was transferred inside a 50 ml sterile falcon tube and centrifuged at 4°C for 20 minutes at 9980x g. The supernatant that formed was removed from the sterile falcon tube, and the remaining pellet was resuspended in 5 ml of sterile physiological solution cold before storage of the pellet at -80°C until DNA extraction. Two different protocols and kits were applied depending on the type of sample to be performed.

DNA extraction from broiler cecal contents

The DNA was extracted from each sample of cecal content using a modified QIAamp® DNA Stool Mini Kit protocol (Qiagen, Milan, Italy), as previously described (De Cesare *et al*., 2017). One major modification is the addition of a bead-beating step at the beginning of DNA extraction. Briefly, 0.25 g of caecal content were suspended in 1 ml lysis buffer (500 mM NaCl, 50 mM Tris-Cl, pH 8.0, 50 mM EDTA, 4 % SDS) with steel beads, 2 mm size and 5 mm size (Retsch, Germany) and homogenized on the Mixer Mill MM 400 (Retsch, Germany) for 2 minutes at frequency 1/s 300. The samples were then heated at 70°C for 15 minutes, followed by centrifugation at 4°C for 5 minutes at 13000 rpm, to separate the DNA from the bacterial cellular debris. This process was repeated with a second 300 µL aliquot of lysis buffer. The samples were then subjected to 10 M v/v ammonium precipitation and a 70% ethanol (Carlo Erba, Milan, Italy) wash and re-suspended in 100 ml TE buffer (10 mmol l -1 Tris-HCl; 1 mmol l -1 EDTA, pH 8.0). The samples were treated with DNase-free RNase (Roche) and incubated overnight at 4°C, before being processed through the QIAamp® DNA Stool Mini Kit (Qiagen, Milan, Italy) according to manufacturer's directions with some modifications.

DNA extraction from broilers carcasses

Total DNA from each carcass sample was extracted using PowerFood ® Microbial DNA Isolation Kit (MoBio-Qiagen) according to manufacturer's directions with some modifications. Briefly, 2 ml of solution was taken from each sample and centrifuged at 4°C for 20 minutes at 6800 rpm. Once the supernatant was removed, the formed pellet was suspended in 450 ml of lysing solution, which includes a detergent to break the cell walls (PF1) (Manual instruction PowerFood Microbial DNA Isolation Kit). Subsequent steps of the extraction protocol were performed in accordance with the handbook that came with the extraction kit.

Quantification of total DNA

The total DNA isolated from each cecal content samples and each carcasses sample were quantified on a BioSpectrometer® (Eppendorf, Milan, Italy) and Fragment Analyzer Automated CE System

(Advanced Analytical). All samples complying with DNA quality scores were subjected to library preparation.

Library preparation and metagenomic sequencing

Total DNA from caeca and carcass samples was fragmented and tagged with sequencing indexes and adapters using the Nextera® XT DNA Library Preparation Kit (Illumina, San Diego, CA). Shotgun metagenomic sequencing was performed using the NextSeq500 (Illumina) with NextSeq500/550 High Output Reagent Kit v2 (300 Cycles) (Illumina) at 2x150 bp in paired-end mode. One caeca sample was removed later in the process for technical reasons (linked to sequencing yield) resulting in a total of 59 samples (broilers caeca: N=29, broiler carcasses: N=30).

Bioinformatic and biostatistics analysis

Filtering and trimming of raw reads were performed using MG-RAST https://www. mg-rast.org (accessed on 14 December 2021) (Keegan *et al*., 2016) bioinformatics pipelines. In the taxonomic analysis, only taxa from the bacterial domain were considered. Moreover, taxa present in less than 4 samples or represented by less than four reads were discarded. Analogously, in the functional analysis, functional genes present in less than four samples or represented by less than 4 reads were discarded.

The statistical analysis of both the taxonomic and functional gene composition was performed using R 3.6.3 and the libraries phyloseq 1.30.0 (McMurdie and Holmes, 2013) and DESeq2 1.26.0 (Love *et al*., 2014b). Relative abundances displayed in the bar plots were computed normalizing to sum to 1 the read counts obtained from MG-RAST. Then, in the bar plots only the first (at most) 20 taxa/functional elements with relative abundance greater than 1% were shown.

Before proceeding with the statistical analysis, read counts were normalized with DESeq2 to take into account the different sizes of the samples. In this step, size factors were estimated using the function estimateSizeFactors of DESeq2 with the "poscount" option, as suggested when dealing with sparse data. Then, DESeq2 was used to assess whether the taxa/function abundances differed between groups. Specifically, the Wald test was used to determine the statistical significance and the Log Fold Changes were shrunk using the apeglm method (Zhu *et al*., 2019). Finally, p-values were adjusted for multiple testing using the Benjamini-Hocheberg procedure (Benjamini and Hochberg, 1995). A threshold of 0.05 was used in all analyses to assess their statistical significance.

Alpha diversity was estimated using the InvSimpson/Shannon/Chao1 index, and differences in alpha diversity between groups were evaluated fitting a linear regression model and using the Student's t-test to assess whether the linear relationship between alpha diversity and the grouping was negligible.

Beta diversity was estimated starting from the read counts normalized with DESeq2 and computing the Bray-Curtis dissimilarity (Bray and Curtis, 1957) among samples. Principal Coordinate Analysis (PCoA) was used to visualize the results. After applying the rlog transformation (Love *et al*., 2014b) to DESeq2 normalized counts, Principal Component Analysis (PCA) was performed using the function prcomp of the library stats 3.6.3 in R and the correlation between samples was computed using Kendall's coefficient.

The resistome of each sample was predicted using the Resistance Gene Identifier (RGI) (Alcock *et al*., 2020). Fastq reads were aligned using the bowtie2 algorithm (Langmead and Salzberg, 2012) to the 'canonical' curated CARD reference sequences (Alcock *et al*., 2020), as well as to the in silico predicted allelic variants available in CARD's Resistomes & Variants data set (Alcock *et al*., 2020), as suggested in the resistance gene identifier (RGI). The alignments were obtained at the allele level and were filtered so that only entries with >95% identity to the CARD reference sequences and with more than 50 base pairs of reference allele covered by reads were kept. RGI mapping counts were adjusted for differences in both gene lengths and bacterial sequence abundances by computing fragments per kilobase reference per million bacterial fragments (FPKM). Results at the AMR gene family and drug class level were obtained by aggregating the counts at the allele level. The beta diversity of the samples based on the resistome was obtained by computing the PCoA. To this aim, the counts were normalized with DESeq2 as previously described, and the Bray-Curtis distances between all samples were calculated using the R packages vegan 2.5.7 (Legendre and Gallagher, 2001) and phyloseq 1.28.0 (McMurdie and Holmes, 2013). The PCoA was computed separately for caeca and carcass samples, and the effect of the origin of the sample on the sample dissimilarities were determined using permutational multivariate analysis of variance using distance matrices (the 'adonis2′ function in the vegan v2.5.7 package). Finally, conventional and antibiotic free AMR gene families were compared using the same DESeq2 pipeline previously described for the taxonomic and functional analysis.

3.3 Materials and methods of study 3

Broiler farms and sample collection

In this study, three groups of carcasses (i.e., Group 1, Group 2 and Group 3) reared in three broiler farms have been tested. All farms were located in Northern Italy. The animals of group 1 were slaughtered at 35 days and treated with antibiotics at 28 days; the animals of group 2 were slaughtered at 34 days and never treated with antibiotics; the animals of group 3 were slaughtered at 33 days and never treated with antibiotics. A total of 15 carcasses from each tested group were randomly sampled

at the slaughterhouse at the end of the refrigeration tunnel and immediately stored in a sterile plastic bag kept at 4°C in a refrigerated box. After each sampling, the samples were transported to the laboratory within two hours and immediately processed. Approximately 10 g of neck and breast skin were collected from each carcass and diluted in 90 ml of sterile physiological solution (NaCl 0.90%) inside a sterile bag and homogenized in the stomacher (MAYO homogenius HG400V) at normal speed for 1 minute. Subsequently, all the liquid part of the sample was transferred inside a 50 ml sterile falcon tube and centrifuged at 4°C for 20 minutes at 6800 rpm. The supernatant that formed was removed from the sterile falcon tube, and the remaining pellet was resuspended in 5 ml of sterile physiological solution cold before storage of the pellet at -80°C until DNA extraction.

Microbiological analysis of the carcasses

Sample enrichment and enumeration were both used to detection *Campylobacter* from carcasses according to parts 1 and 2 respectively of the EN ISO 10272-1:2006 standard. In brief, in the first stage of enrichment, a total of 10 g of skin from the breast and the neck were put into a sterile bag and diluted 1:10 in selective liquid medium, Bolton broth (Oxoid) (Bolton *et al*., 1984), incubated in microaerobic atmosphere (85% N2, 10% CO2, 5% O2) obtained by Oxoid CampyGen™ 3.5L Sachet (ThermoFisher) inside an Oxoid™ Anaerobic 3.5L Jar at 37°C for 4 h to 6 h followed by 42°C for 44 h ± 4 h. At the end of the incubation period (isolation and selection for conformation phase) carried out to give the pathogen cells the possibility to multiply if they were present, the enrichment broth was streaked with a sterile loop on superface of selective isolation medium, modified charcoal cefoperazone deoxycholate agar (mCCD agar) (Oxoid) in triplicate and incubated at 42°C in a microaerobic atmosphere (as previously described) for 44 ± 4 h. ISO 10272 requires that the enrichment culture is streaked onto two selective media after 48 h, however, due to constraints on resources, only modified charcoal cefoperazone deoxycholate agar (mCCDA) was used in this study. The reference method for the detection of *Salmonella* in carcasses was performed according to the international reference standard method EN ISO 6579:2002/Amd.1:2007. In brief, a further aliquot of 10 g was put into a sterile bag and dilute 1:10 in non-selective broth, Buffered Peptone Water (BPW). The mix was then homogenized for 1 minute in a stomacher and incubated at 37 °C ± 1° for 18 h ± 2 h (non-selective pre-enrichment phase), followed by a phase of selective enrichment in Modified Semi-solid Rappaport Vassiliadis Agar (MRVS) (Oxoid) incubated at 41.5 ± 1°C for 24 h ± 3 h and Muller-Kaufmann tetrathionate-Novobiocin Broth (MKTTn) incubated at 37°C ± 1°C for 24 h ± 3 h. The RVS plates were inoculate with 3 drops of incubated BPW culture. The 3 drops were total 0,1 ml and were placed separately and equally spaced on the surface of the medium. Both enrichment broths were streaked on selective solid media, xylose lysine deoxycholate (XLD) and

Brilliant Green Agar (BGA); overall, were identified furthest point of spread of opaque growth from the inoculation points and was dip a 1 µl loop just inside the border of the opaque growth and streaked in the XLD and BGA plate, so that well-isolated colonies will be obtained. The XLD and BGA plates were Incubated at 37°C ± 1°C for 24 h ± 3 h (selective plating and identification phase).

DNA extraction and quantification from broilers carcasses

DNA was extracted from each carcass sample using the PowerFood® Microbial DNA Isolation Kit (MO BIO-Qiagen), as reported in Study 2. The purity and concentration of DNA samples were determined according to the 260/280 nm absorbance ratio using a biospectrophotometers, BioSpectrometer® (Eppendorf, Milan, Italy) and Fragment Analyzer Automated CE System (Advanced Analytical). All samples complying with DNA quality scores were subjected to library preparation.

Library preparation and metagenomic sequencing

Total DNA from carcasses sample was fragmented and tagged with sequencing indexes and adapters using the Nextera® XT DNA Library Preparation Kit (Illumina, San Diego, CA). Shotgun metagenomic sequencing was performed using the NextSeq 500 sequencer (Illumina) with NextSeq500/550 High Output Reagent Kit v2 (300 Cycles) (Illumina) at 2x150 bp in paired-end mode. Three carcasses breast and neck skin samples were removed later in the process for technical reasons (linked to sequencing) resulting in a total of 42 samples (Group 1 carcasses: N = 12; Group 2 carcasses: N = 15; Group 3 carcasses: N = 15).

Bioinformatic analysis

The metagenomic sequences (filtered and trimmed) were analysed using the MG-RAST metagenomics analysis server, bioinformatics pipeline. MG-RAST is a software built to provide users of a complete analysis of environmental DNA (metagenomic sequences), in terms of alignment of sequences and taxonomic and functional assignments (Meyer *et al*., 2008). The system provides answers to several applications, and in particular, for our purpose, it helps to identify the composition of a microbial community and function by deriving community composition from shotgun metagenomic data using sequence similarities. For this purpose, the MG-RAST software performs quality control, protein prediction, clustering and similarity-based annotation on nucleic acid sequence datasets using several bioinformatics tools (Wilke *et al*., 2017).

After applying the quality control procedure, following the instructions of the MG-RAST manual (Keegan *et al*., 2016) (https://www.mg-rast.org), were identified the mean values of the relative

frequency of abundances of taxonomic composition of the sequencing data, performing a search the RefSeq database (NCBI reference sequences) (Pruitt *et al*., 2005) and narrowed down the analysis only to those genera annotations marked as Bacteria, in order to study the relevant component of chicken carcasses. Moreover, mean values of the relative frequency of abundances of the functional composition of the sequencing data were performed using the SEED Subsystems database (Overbeek *et al*., 2014) using database hits at the function level with a focus on ARGs (level 3 and function level). In both cases, the following quality parameters were set: maximum e-value cut-off <1e-5, minimum identity cut-off 60%, and minimum alignment length cut-off 15 bp.

Statistical analysis

The relative frequency of abundance of each taxonomic level, of the functional genes and ARGs were obtained by the normalized read counts and compared in the software Statistical Analysis of Metagenomic Profile v2.1.3 (STAMP) (Parks *et al*., 2014). The first step was to convert the tsv format to the txt format (MS-DOS). Subsequently, the metadata file necessary to associate the reads obtained with the belonging samples (Group1, Group 2, Group 3) was created. Statistical processing for comparing three groups of profiles was carried out applying the ANOVA test, without multiple test correction method. For the statistically significant features were further examined with Tukey-Kramer (post hoc tests) setting a significance value of the p-value ≤0.05, determining which groups of profiles differ from each other. The software was asked to remove unclassified reads. Additionally, we chose ≥1% relative abundance as an arbitrary cut-off to compare taxa or pathways, whereas, in reality, genera, species, function level of functional genes, level 3 and function level of ARGs may be present below this threshold.

Diversity indices were analysed at genus and species level in free software for statistical computing and graphics, R 3.6.3 (McMurdie and Holmes, 2013). The vegan package (version 2.5.6) (Oksanen *et al*., 2020) was used for alpha diversity analysis by choosing Shannon, Simpson, and Inverse Simpson indices, while for beta diversity analysis Bray-Curtis distance matrix method. Alpha indices were analysed with one-way ANOVA and Tukey's post-hoc test by considering "group" as the experimental factor and the individual sampled animal as the experimental unit. Beta diversity was graphically examined through principal coordinates analysis (PCoA) and analysed with permutational multivariate analysis of variance (PERMANOVA) by using "adonis" function with 10.000 permutations, followed by pairwise permutation MANOVA with RVAideMemoire package (Hervé, 2021).

3.4 Materials and methods of study 4

In this very preliminary study, 1 artisanal salame manufactured in a local meat factory in the Emilia-Romagna area was investigated. Meat batter was then stuffed into casings, resulting in the sausage of about 20 cm long and about 500 g in weight. Fermentation and ripening for 126 days (18 weeks) were carried out in a climatic chamber before collecting the sample for analysis. The sample was placed in a sterile plastic bag and transferred to the laboratory immediately in a refrigerated cold box of about 4°C and processed. A portion of 25 g of salami (casing and meat) was sampled at three different points using a sterile scalpel and placed in a stomacher bag. Then, 225 ml of BPW (Oxoid) was added to the bag. The content was homogenized in a stomacher (MAYO homogenius HG400V) at normal speed for 1 minute. Next, the homogenized solution was transferred in 3 falcons of 50 ml. Three aliquots of 100 µl each were obtained from each tube and spread plated onto Baird Parker (Oxoid) with added egg yolk tellurite emulsion (Oxoid) incubated at 37°C for 48 h according to the ISO 6888-1/A1:2003. The rest of the tube content was centrifuged at 5000 rpm at 4°C for 20 minutes. The supernatant was discarded, and the pellet was resuspended in 6 ml of cold sterile physiological solution, pelleted again before storage at -80°C until DNA extraction.

DNA extraction and quantification from pellet obtained from the sample homogenate

The DNA was extracted from each of the three pellet aliquots using the PowerFood® Microbial DNA Isolation Kit (MO BIO-Qiagen, Milano, Italia) as previously described (De Cesare *et al*., 2019). The purity and concentration of DNA samples were determined according to the 260/280 nm absorbance ratio using a biospectrophotometers, BioSpectrometer® (Eppendorf, Milan, Italy) and Fragment Analyzer Automated CE System (Advanced Analytical).

Library preparation and shotgun metagenomic sequencing

Total DNA from each aliquot of pellet was fragmented and tagged with sequencing indexes and adapters using the Nextera® XT DNA Library Preparation Kit (Illumina, San Diego, CA). Shotgun metagenomic sequencing was performed using the NextSeq 500 sequencer (Illumina) with NextSeq500/550 High Output Kit v2.5 (300 Cycles) (Illumina) at 2x150 bp in paired-end mode.

Bioinformatic analysis

The metagenomic sequences were analysed using the MG-RAST metagenomics analysis server (Keegan *et al*., 2016) (https://www.mg-rast.org), to identify the relative abundances of taxonomic composition using a search in the RefSeq database (NCBI reference sequences) (Pruitt *et al*., 2005). Moreover, the relative abundances of the functional composition of the sequencing data, was

performed using the SEED Subsystems database (Overbeek *et al*., 2014) using database hits at the level1and function level. Statistical differences in MG-RAST outputs were not evaluated because only three samples were analysed.

4. Results

4.1 Study 1

In the study 1, three metagenomes were sequenced and analysed with different bioinformatic tools to assess their performance in the detection and quantification of all microorganisms spiked in the salmon samples. The output of each metagenome is shown in Table 5. Each metagenome was analysed using the bioinformatic tool named MG-RAST selecting the databases RefSeq, SILVA LSU, SILVA SSU, and Greengenes. Moreover, the same metagenomes were analysed using OneCodex, MG-mapper and CosmosID. Among these tools, MG-RAST is freely available online at the link https://www.mg-rast.org/; Mg-mapper is freely available at the link https://bitbucket.org/genomicepidemiology/mgmapper/src/master/ but is not always updated in the on-line version; OneCodex, at the link https://www.onecodex.com/, and CosmosID, at the link https://www.cosmosid.com/, are available online but they use requires the payment of a fee.

Table 5: Sample ID with the corresponding metagenome and outputs.

Sample ID	Length of the library (bp)	N. of reads	Gbp
1216	386	104251249	9.81
1217	391	108914800	8.44
1218	541	128316548	7.97

Reads belonging to microorganisms different from those spiked in the salmon were identified in the analysed metagenomes because the salmon was not sterilized. However, the description of the results focuses on the spiked microorganisms.

Metagenomes analysis using MG-RAST

Table 6 summarises the relative abundance (%) calculated in the tested metagenomes using MG-RAST and the database RefSeq. The database RefSeq identified all the taxonomic groups belonging to the bacteria species spiked in the salmon. Among the spiked bacteria, *P. freudenreichii* and *S. aureus* showed the highest relative abundance values, reflecting the highest concentration of both these species spiked in the salmon at Log 8 cfu/g. For the other bacteria species, spiked in the salmon at Log 7 cfu/g, a clear correspondence with the relative abundance values was not observed. Indeed, *F. nucleatum* showed a relative abundance of 1.62% while *E. coli* a relative abundance of 4.79%. Moreover, *S. enterica* and *B. fragilis* displayed comparable values but around 9%.

Among the other microorganisms (i.e., non-bacteria microorganisms) spiked in the salmon (Table 7), RefSeq was able to identify the yeast with a mean value of relative abundance of 0.01% and the parasite with a mean value of relative abundance of 0.15%. On the contrary, the spiked virus was not detected.

Table 6: Relative abundance of the bacteria spiked in the tested samples calculated using MG-RAST and the database RefSeq.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Phylum	Proteobacteria	29.08	35.99	28.37	31.15	3.43
	Firmicutes	36.31	35.73	34.13	35.39	0.92
	Actinobacteria	31.05	37.18	29.52	29.25	1.59
	Bacteroidetes	16.54	18.57	18.43	17.84	0.92
	Fusobacteria	2.56	1.31	1.47	1.78	0.55
Class	Gammaproteobacteria	28.79	35.70	28.08	30.86	3.43
	Bacilli	26.13	25.56	23.95	25.21	0.92
	Actinobacteria	31.26	27.33	29.75	29.44	1.61
	Bacteroidia	16.60	18.62	18.52	17.91	0.93
	Fusobacteria	2.58	1.31	1.48	1.79	0.56
Order	Enterobacterales	23.45	23.69	22.29	23.14	0.60
	Bacillales	18.00	19.44	23.83	20.42	2.48
	Propionibacteriales	25.71	22.49	24.49	24.23	1.41
	Bacteroidales	16.56	18.58	18.49	17.88	0.93
	Fusobacteriales	2.57	1.31	1.48	1.79	0.56
Family	Enterobacteriaceae	23.35	23.62	22.17	23.05	0.62
	Staphylococcaceae	15.88	15.33	13.73	14.98	0.91
	Propionibacteriaceae	25.71	22.49	24.49	24.23	1.31
	Bacteroidaceae	16.34	18.37	18.23	17.65	0.92
	Fusobacteriaceae	2.56	1.31	1.47	1.78	0.55
Genus	*Escherichia*	5.87	5.33	3.72	4.97	0.91
	Staphylococcus	11.56	11.57	9.95	11.03	0.76
	Propionibacterium	25.69	22.48	24.47	24.21	1.32
	Bacteroides	16.33	18.37	18.22	17.64	0.92
	Salmonella	8.56	8.56	9.10	8.74	0.25
	Fusobacterium	2.55	1.30	1.46	1.77	0.55
Species	*E. coli*	5.65	5.14	3.58	4.79	0.88
	S. aureus	10.62	10.63	9.14	10.13	0.70
	P. freudenreichii	25.15	21.10	23.01	23.08	1.65
	B. fragilis	8.70	9.82	9.72	9.41	0.50
	S. enterica	8.55	8.54	9.08	8.72	0.25
	F. nucleatum	2.34	1.19	1.13	1.62	0.51

Table 7: Relative abundance of the fungus, parasite and virus spiked in the tested samples calculated using MG-RAST and the database RefSeq.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Species	*Saccharomyces cerevisiae*	0.02	0.02	0.01	0.01	>0.00
	Cryptosporidium parvum	0.16	0.18	0.11	0.15	0.02
	Bovine alphaherpesvirus 1	-	-	-	-	-

Table 8 summarises the relative abundance values (%) calculated using MG-RAST and the database SILVA LSU for the bacteria groups spiked in the salmon samples. The database Silva SSU identified all the bacteria species spiked in the salmon. Moreover, using this database *P. freudenreichii* and *S. aureus* displayed a comparable level of relative abundance reflecting their equal level of concentration in the salmon. A good level of correspondence was also observed for the species spiked at Log 7 cfu/g as *E. coli, B. fragilis, S. enterica* and *F. nucleatum*. Among the other microorganisms spiked in the salmon (Table 9) Silva LSU identified the yeast and the parasite but the virus was not detected.

Table 8: Relative abundance of the bacteria spiked in the tested samples calculated using MG-RAST and the database SILVA LSU.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Phylum	Proteobacteria	10.72	23.12	6.17	13.33	7.16
	Firmicutes	3.34	2.96	1.30	2.53	0.89
	Actinobacteria	4.141	5.022	4.80	4.65	0.37
	Bacteroidetes	2.23	2.99	2.05	2.42	0.41
	Fusobacteria	4.31	2.26	2.16	2.91	0.99
Class	Gammaproteobacteria	10.71	23.09	6.18	13.32	7.15
	Bacilli	3.24	2.89	2.62	2.91	0.25
	Actinobacteria	4.14	5.02	2.81	3.99	0.91
	Bacteroidia	2.23	2.99	2.06	2.42	0.40
	Fusobacteria	4.32	2.26	2.16	2.91	0.99
Order	Enterobacterales	4.97	6.68	3.83	5.16	1.17
	Bacillales	3.61	3.07	1.40	2.69	0.94
	Propionibacteriales	4.09	4.94	2.76	3.93	0.90
	Bacteroidales	2.58	3.33	2.41	2.77	0.40
	Fusobacteriales	4.98	2.52	1.36	2.95	1.50
Family	Enterobacteriaceae	4.30	6.00	3.26	4.52	1.13
	Staphylococcaceae	3.12	2.75	1.18	2.35	0.84
	Propionibacteriaceae	4.04	4.88	2.72	3.88	0.89
	Bacteroidaceae	2.23	2.99	2.05	2.42	0.41
	Fusobacteriaceae	4.31	2.26	2.16	2.91	0.99
Genus	*Escherichia*	2.13	2.68	1.42	2.07	0.52
	Staphylococcus	3.12	2.75	1.18	2.35	0.84
	Propionibacterium	4.04	4.88	2.72	3.88	0.89
	Bacteroides	2.23	2.99	2.05	2.42	0.41
	Salmonella	1.19	1.54	1.06	1.26	0.20
	Fusobacterium	4.30	2.26	2.09	2.88	1.00
Species	*E. coli*	2.12	2.67	1.42	2.07	0.51
	S. aureus	3.08	2.72	1.18	2.32	0.82
	P. freudenreichii	3.97	4.82	2.67	3.82	0.88
	B. fragilis	1.42	1.90	1.30	1.54	0.26
	S. enterica	1.18	1.53	1.05	1.25	0.20
	F. nucleatum	1.78	1.56	1.43	1.59	0.14

Table 9: Relative abundance of the fungus, parasite and virus spiked in the tested samples calculated using MG-RAST and the database SILVA LSU.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Species	*Saccharomyces cerevisiae*	0.005	0.006	0.006	0.005	<0.001
	Cryptosporidium parvum	0.019	0.018	0.018	0.018	<0.001
	Bovine alphaherpesvirus 1	-	-	-	-	-

Table 10 summarises the relative abundance (%) calculated using Mg-RAST and the database SILVA SSU for the bacteria groups spiked in the salmon samples. The database SILVA SSU identified all the bacteria species spiked in the salmon. Among the bacteria, *P. freudenreichii* and *S. aureus* spiked at Log 8 cfu/g showed comparable mean relative values of abundance, while for the species spiked at Log 7 cfu/g comparable values of abundance were not detected. Indeed, S. *enterica* and *F. nucleatum* showed lower relative abundance than *B. fragilis* and *E. coli*, although were spiked at higher concentrations. Among the other microorganisms spiked in the salmon (Table 11) the database SILVA SSU was able to detect the yeast with a relative abundance as low as 0.005%, while *Cryptosporidium parvum* was quantified with a relative abundance of 0.137%. The spiked virus was not identified by the database Silva SSU

Table 10: Relative abundance of the bacteria spiked in the tested samples calculated using MG-RAST and the database SILVA SSU.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Phylum	Proteobacteria	34.23	53.67	39	42.30	8.27
	Firmicutes	16.91	10.43	10.61	12.65	3.01
	Actinobacteria	22.67	19.98	25.65	22.76	2.32
	Bacteroidetes	10.85	10.42	17.68	12.98	3.33
	Fusobacteria	14.84	5.39	6.12	8.78	4.29
Class	Gammaproteobacteria	34.12	53.44	37.17	41.57	8.48
	Bacilli	16.92	10.48	11.05	12.81	2.91
	Actinobacteria	22.69	20.09	26.72	23.16	2.73
	Bacteroidia	10.83	10.44	18.42	13.23	3.67
	Fusobacteria	14.85	5.42	6.38	8.88	4.24
Order	Enterobacterales	16.09	15.19	20.85	17.37	2.48
	Bacillales	16.39	10.24	10.49	12.37	2.84
	Propionibacteriales	22.66	20.32	26.44	23.14	2.52
	Bacteroidales	10.87	10.61	18.29	13.25	3.56
	Fusobacteriales	14.90	5.51	6.34	8.91	4.24
Family	Enterobacteriaceae	16.09	15.19	21.02	17.43	2.56
	Staphylococcaceae	12.32	9.47	9.88	10.55	1.26
	Propionibacteriaceae	21.18	19.47	25.20	21.95	2.40
	Bacteroidaceae	10.55	10.35	11.17	10.69	0.35
	Fusobacteriaceae	6.90	5.51	6.39	6.26	0.57
Genus	*Escherichia*	7.51	5.88	8.45	7.28	1.06
	Staphylococcus	11.23	9.45	9.59	10.09	0.81
	Propionibacterium	21.24	19.59	24.81	21.88	2.18
	Bacteroides	10.63	10.44	18.09	10.72	0.27
	Salmonella	1.79	1.65	2.46	1.96	0.35
	Fusobacterium	5.77	5.43	6.24	5.81	0.332

Species	*E. coli*	4.49	3.24	5.23	4.32	0.82
	S. aureus	6.86	5.87	5.96	6.23	0.45
	P.freudenreichii	14.52	12.33	15.90	14.91	1.47
	B. fragilis	4.82	4.40	9.25	4.49	0.24
	S. enterica	1.39	1.20	1.98	1.52	0.33
	F. nucleatum	1.49	1.46	1.87	1.60	0.19

Table 11: Relative abundance of the fungus, parasite and virus spiked in the tested samples calculated using MG-RAST and the database SILVA SSU.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Species	*Saccharomyces cerevisiae*	0.006	0.007	0.003	0.005	0.002
	Cryptosporidium parvum	0.16	0.11	0.14	0.137	0.021
	Bovine alphaherpesvirus 1	-	-	-	-	-

Table 12 summarise the relative abundance (%) calculated using MG-RAST and the database RDP for the bacteria groups spiked in the salmon samples. This database identified all spiked bacteria species. Moreover, *P. freudenreichii* was detected as the most abundant species while *S. aureus* spiked at Log 8 cfu/g as *P. freudenreichii* was detected with a relative abundance of 3.33%. *S. enterica* and *F. nucleatum* showed lower relative abundance than *B. fragilis* although were spiked at higher concentrations. The other microorganisms spiked in the salmon were not detected using Mg-RAST and the database RDP.

Table 12: Relative abundance of the bacteria spiked in the tested samples calculated using MG-RAST and the database RDP.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Phylum	Proteobacteria	18.05	37.14	22.38	25.86	8.17
	Firmicutes	12.39	7.97	7.93	9.43	2.09
	Actinobacteria	32.20	30.05	34.60	32.28	1.86
	Bacteroidetes	15.09	16.17	25.30	18.85	4.58
	Fusobacteria	21.77	8.51	8.64	12.97	6.22
Class	Gammaproteobacteria	17.92	19.36	18.54	18.61	0.59
	Bacilli	12.40	8.04	8.43	9.62	1.97
	Actinobacteria	32.25	30.31	36.81	33.12	2.73
	Bacteroidia	15.11	16.31	26.91	19.44	5.30
	Fusobacteria	21.81	8.58	9.19	13.19	6.10
Order	Enterobacterales	4.56	6.44	4.45	5.15	0.91
	Bacillales	12.21	7.99	8.19	9.46	1.94
	Propionibacteriales	32.40	31.09	36.43	33.31	2.27
	Bacteroidales	15.18	16.75	16.64	16.19	0.72

	Fusobacteriales	21.92	8.81	9.10	13.28	6.11
Family	Enterobacteriaceae	4.56	6.44	4.50	5.17	0.90
	Staphylococcaceae	8.27	7.12	7.57	7.65	0.47
	Propionibacteriaceae	30.73	30.07	35.57	32.12	2.45
	Bacteroidaceae	14.71	16.34	16.56	15.87	0.83
	Fusobacteriaceae	12.92	8.81	9.20	10.31	1.85
Genus	*Escherichia*	0.68	0.19	0.70	0.52	0.24
	Staphylococcus	7.95	7.02	7.15	7.37	0.41
	Propionibacterium	30.68	30.27	32.96	31.30	1.18
	Bacteroides	14.76	16.50	16.42	15.89	0.80
	Salmonella	0.32	0.42	0.35	0.36	0.04
	Fusobacterium	9.62	8.68	8.97	9.09	0.39
Species	*E. coli*	0.42	0.98	0.50	0.63	0.25
	S. aureus	4.43	2.25	3.32	3.33	0.89
	P. freudenreichii	18.83	18.33	20.27	19.14	0.82
	B. fragilis	6.82	6.52	7.51	6.95	0.41
	S. enterica	0.14	0.19	0.16	0.16	0.02

Table 13 summarises the relative abundance (%) calculated using MG-RAST and the database Greengenes for the bacteria groups spiked in the salmon samples. The database Greengenes identified all the taxonomic groups belonging to the bacteria species spiked in the salmon. Moreover, *P. freudenreichii* was detected as the most abundant species while *S. aureus* was detected with a relative abundance of 3.10%. Among the bacteria spiked at Log 7 cfu/g *S. enterica*, *E. coli* and *F. nucleatum* showed lower relative abundance than *B. fragilis* although were spiked at higher concentrations. The yeast, parasite and virus spiked in the salmon were not detected using the database Greengenes.

Table 13: Relative abundance of the bacteria spiked in the tested samples calculated using MG-RAST and the database Greengenes.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Phylum	Proteobacteria	27.03	26.61	23.88	25.84	1.40
	Firmicutes	10.35	6.88	7.18	8.14	1.57
	Actinobacteria	33.15	31.75	35.52	33.47	1.56
	Bacteroidetes	14.68	15.02	22.98	17.56	3.84
	Fusobacteria	11.05	9.31	9.41	9.92	0.80
Class	Gammaproteobacteria	17.00	16.19	19.47	17.55	1.40
	Bacilli	10.43	6.99	7.69	8.37	1.48
	Actinobacteria	33.45	32.25	38.08	34.59	2.51
	Bacteroidia	14.81	15.25	24.64	18.23	4.53
	Fusobacteria	24.27	9.27	9.08	9.54	0.52
Order	Enterobacterales	5.80	7.85	5.69	6.45	0.99
	Bacillales	10.42	7.13	7.56	8.37	1.46
	Propionibacteriales	33.42	33.09	37.75	34.75	2.12

	Bacteroidales	14.80	15.65	14.43	14.96	0.51
	Fusobacteriales	9.25	8.98	8.90	9.04	0.15
Family	Enterobacteriaceae	5.84	7.90	5.75	6.50	0.99
	Staphylococcaceae	9.87	6.58	7.13	7.86	1.44
	Propionibacteriaceae	33.25	32.68	37.75	34.56	2.27
	Bacteroidaceae	14.59	15.58	14.38	14.85	0.52
	Fusobacteriaceae	8.42	8.57	8.70	8.56	0.11
Genus	*Escherichia*	2.41	2.95	1.62	2.33	0.55
	Staphylococcus	8.40	6.41	6.62	7.14	0.89
	Propionibacterium	33.11	32.84	34.43	33.46	0.70
	Bacteroides	14.53	15.66	14.18	14.79	0.63
	Salmonella	1.78	1.79	2.40	1.99	0.29
	Fusobacterium	23.87	9.39	5.82	5.92	0.47
Species	*E. coli*	1.66	1.67	1.19	1.51	0.22
	S. aureus	3.90	2.06	3.33	3.10	0.77
	P. freudenreichii	23.32	24.72	24.72	24.25	0.66
	B. fragilis	7.08	5.91	7.78	6.92	0.77
	S. enterica	1.26	1.03	1.73	1.34	0.29
	F. nucleatum	1.98	1.74	2.37	2.03	0.26

Table 14 summarises all the mean relative values of abundance (%) quantified for the species spiked in the salmon samples using the different databases available in MG-RAST. According to Petersen *et al*. (2017), whenever the ratio between the number of reads associated with a specific microorganism and the total number of reads in the sample is > 0.1% the taxonomic classification can be considered reliable.

Table 14: Mean relative values of abundance calculated for the spiked species using MG-RAST with the databases RefSeq, Silva LSU, Silva SSU, RDP and Greengenes.

Species	**RefSeq**	**Silva LSU**	**Silva SSU**	**RDP**	**Greengenes**
Escherichia coli	4.79	2.07	4.32	1.51	0.63
Staphylococcus aureus	10.13	2.32	6.23	3.10	3.33
Propionibacterium freudenreichii	23.08	3.82	14.91	24.25	19.14
Bacteroides fragilis	9.41	1.54	4.49	6.92	6.95
Salmonella enterica	8.72	1.25	1.52	1.34	0.16
Fusobacterium nucleatum	1.62	1.59	1.60	2.03	2.80
Saccharomyces cerevisiae	0.01	<0.01	<0.01	0	0
Cryptosporidium parvum	0.15	0.01	0.13	0	0
Bovine alphaherpesvirus 1	0	0	0	0	0

All in all, the results obtained using the software tool MG-RAST demonstrated that all bacteria species spiked in the salmon samples can be correctly identified by all tested databases available in

the tool. On the contrary, the fungus and the virus were never detected, and the parasite can be detected only using RefSeq and Silva SSU.

Metagenomes analysis using CosmosID

Table 15 summarises the relative abundance calculated using CosmosID and the database GenBook for the bacteria groups spiked in the salmon samples. All bacteria species were identified using this software. Moreover, *P. freudenreichii* and *S. aureus* were identified at the higher abundance values while among the bacteria spiked at Log 7 cfu/g, *B. fragilis* showed a relative abundance much higher than *S. enterica* and *F. nucleatum*. Among the other microorganisms spiked in the salmon (Table 16), CosmosID identified the parasite *C. parvum* with a relative abundance of 88.74% and the virus with a relative abundance of 7.14%. However, the yeast was not detected.

Table 15: Relative abundance of the bacteria spiked in the tested samples calculated using CosmosID and the database GenBooK.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Phylum	Proteobacteria	10.07	10.92	11.53	10.84	0.60
	Firmicutes	11.58	20.26	17.24	16.36	3.60
	Actinobacteria	49.82	44.97	43.79	46.19	2.61
	Bacteroidetes	18.89	19.82	21.35	20.02	1.01
	Fusobacteria	9.00	4.67	6.09	6.59	1.80
Class	Gammaproteobacteria	10.64	10.26	11.51	10.80	0.52
	Bacilli	11.58	20.26	17.24	16.36	3.60
	Actinobacteria	49.82	44.97	43.79	46.19	2.61
	Bacteroidia	18.89	19.82	21.35	20.02	1.01
	Fusobacteria	9.00	4.67	6.09	6.59	1.80
Order	Enterobacterales	10.18	9.16	11.15	10.16	0.81
	Bacillales	11.56	20.2	17.21	16.32	3.58
	Propionibacteriales	49.82	44.97	43.64	46.14	2.66
	Bacteroidales	18.89	19.82	21.35	20.02	1.01
	Fusobacteriales	9.00	4.67	6.09	6.59	1.80
Family	Enterobacteriaceae	10.17	9.12	11.14	10.14	0.83
	Staphylococcaceae	11.56	20.2	17.21	16.32	3.58
	Propionibacteriaceae	49.82	44.97	43.64	46.14	2.66
	Bacteroidaceae	18.89	19.82	21.35	20.02	1.01
	Fusobacteriaceae	9.00	4.67	6.09	6.59	1.80
Genus	*Escherichia*	0.32	0.40	0.44	0.39	0.05
	Staphylococcus	11.56	20.20	17.21	16.32	3.58
	Propionibacterium	49.28	44.48	43.18	45.65	2.62
	Bacteroides	18.89	19.82	21.35	20.02	1.01
	Salmonella	9.83	8.69	10.67	9.73	0.81

	Fusobacterium	9.00	4.67	6.09	6.59	1.80
Species	*E. coli*	0.31	0.38	0.44	0.38	0.05
	S. aureus	18.89	19.81	21.34	20.01	1.01
	P. freudenreichii	49.28	44.48	43.18	45.65	2.62
	B. fragilis	17.56	20.02	17.21	18.26	1.25
	S. enterica	9.83	8.69	10.67	9.73	0.81
	F. nucleatum	9.00	4.67	6.09	6.59	1.80

Table 16: Relative abundance of the yeast, parasite and virus spiked in the tested samples calculated using CosmosID and the database GenBook.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Species	*Saccharomyces cerevisiae*	-	-	-	-	-
	Cryptosporidium parvum	88.44	91.20	86.57	88.74	1.90
	Bovine alphaherpesvirus 1	8.04	7.28	6.09	7.14	0.80

Metagenomes analysis using MGmapper

Table 17 summarises the relative abundance (%) calculated using MGmapper for the bacteria groups spiked in the salmon samples. Among the spiked species, *P. freudenreichii* was quantified as the most abundant species while *S. aureus* was quantified at a much lower abundance. Among the bacteria spiked at Log 7 cfu/g, *E. coli* and *B. fragilis* displayed comparable levels of relative abundance. The other microorganisms spiked in the salmon were all identified by MGmapper but at a very low level of relative abundance (Table 17).

Table 17: Relative abundance of the microorganisms spiked in the tested samples calculated using MGmapper.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Species	*Escherichia coli*	1.24	1.56	0.81	1.20	0.30
	Staphylococcus aureus	0.54	0.61	0.24	0.46	0.16
	Propionibacterium freudenreichii	5.00	5.49	3.35	4.61	0.91
	Bacteroides fragilis	1.09	1.60	0.95	1.22	0.28
Fungus	*Saccharomyces cerevisiae*	0.003	0.003	0.001	0.002	0.001
Parasite	*Cryptosporidium parvum*	0.02	0.02	0.02	0.02	0.002
Virus	Bovine alphaherpesvirus 1	0.01	0.01	0.003	0.01	0.001

Metagenomes analysis using OneCodex

Table 18 summarises the percentage of reads assigned to each taxonomic group by OneCodex. This tool was able to identify all spiked microorganisms. In particular, *P. freudenreichii* was quantified as

the most abundant microorganism (Table 19) while *S. aureus* spiked at the same concentration was quantified at lower abundance. *E. coli*, *S. enterica* and *B. fragilis* all spiked at Log 7 cfu/g were quantified at a comparable level of abundance while *F. nucleatum* at lower levels. Among the other microorganisms spiked in the salmon, OneCodex was able to identify *C. parvum* and Bovine alphaherpesvirus1, while *S. cerevisiae* was not identified.

Table 18: Relative abundance of the bacteria spiked in the tested samples calculated using OneCodex.

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Phylum	Proteobacteria	24.66	30.50	24.31	26.49	2.84
	Firmicutes	5.34	4.82	3.57	4.58	0.74
	Actinobacteria	54.02	47.68	53.32	51.67	2.84
	Bacteroidetes	11.57	13.94	15.05	13.52	1.45
	Fusobacteria	1.85	0.85	1.08	1.26	0.43
Class	Gammaproteobacteria	24.5	30.32	24.09	26.30	2.85
	Bacilli	3.02	2.75	2.05	2.61	0.41
	Actinobacteria	54.02	47.67	53.32	51.67	2.84
	Bacteroidia	11.57	13.93	15.05	13.52	1.45
	Fusobacteria	1.85	0.85	1.08	1.26	0.43
Order	Enterobacterales	20.43	20.65	21.66	20.91	0.54
	Bacillales	3	2.71	2.02	2.58	0.41
	Propionibacteriales	52.45	46.31	51.70	50.15	2.73
	Bacteroidales	11.75	13.93	15.05	13.58	1.37
	Fusobacteriales	1.85	0.85	1.08	1.26	0.43
Family	Enterobacteriaceae	17.87	18.08	19.30	18.42	0.63
	Staphylococcaceae	2.95	2.67	1.93	2.52	0.43
	Propionibacteriaceae	52.43	46.30	51.69	50.14	2.73
	Bacteroidaceae	11.54	13.69	14.77	13.33	1.34
	Fusobacteriaceae	1.85	0.85	1.08	1.26	0.43
Genus	*Escherichia*	3.59	3.70	3.34	3.54	0.15
	Staphylococcus	2.95	2.66	1.93	2.51	0.43
	Propionibacterium	51.43	45.39	50.69	49.17	2.69
	Bacteroides	11.54	13.69	14.77	13.33	1.34
	Salmonella	8.04	8.03	9.82	8.63	0.84
	Fusobacterium	1.84	0.85	1.08	1.26	0.42

Table 19: Relative abundance (%) calculated with OneCodex for the species spiked in the salmon

		Metagenome ID				
		1216	**1217**	**1218**	**Mean**	**SD**
Species	*Escherichia coli*	7.78	8.31	7.32	7.80	0.40
	Staphylococcus aureus	7.42	7.07	5.03	6.51	1.05
	Propionibacterium freudenreichii	64.36	60.12	65.55	63.34	2.33
	Bacteroides fragilis	7.44	9.29	8.79	8.51	0.78
	Salmonella enterica	6.44	6.83	8.18	7.15	0.75

	Fusobacterium nucleatum	3.26	1.60	2.00	2.29	0.71
Fungus	*Saccharomyces cerevisiae*	-	-	-	-	-
Parasite	*Cryptosporidium parvum*	0.08	0.09	0.07	0.08	0.01
Virus	Bovine alphaherpesvirus 1	1.53	1.52	1.25	1.43	0.13

4.2 Study 2

In the study 2, a total of 59 metagenomes were obtained. The output of each metagenome is detailed in Table 20. All metagenomes are deposited in MG- RAST and are available at the following link https://www.mg-rast.org/mgmain.html?mgpage=project&project=mgp89213. According to the results of study 1, the taxonomic composition of these caeca and carcasses was assessed using Mg-RAST and the database RefSeq.

Table 20: Numbers identifying the tested broilers and corresponding metagenomes for the caeca and the carcasses. The output of each metagenome is detailed as number of sequences and bp.

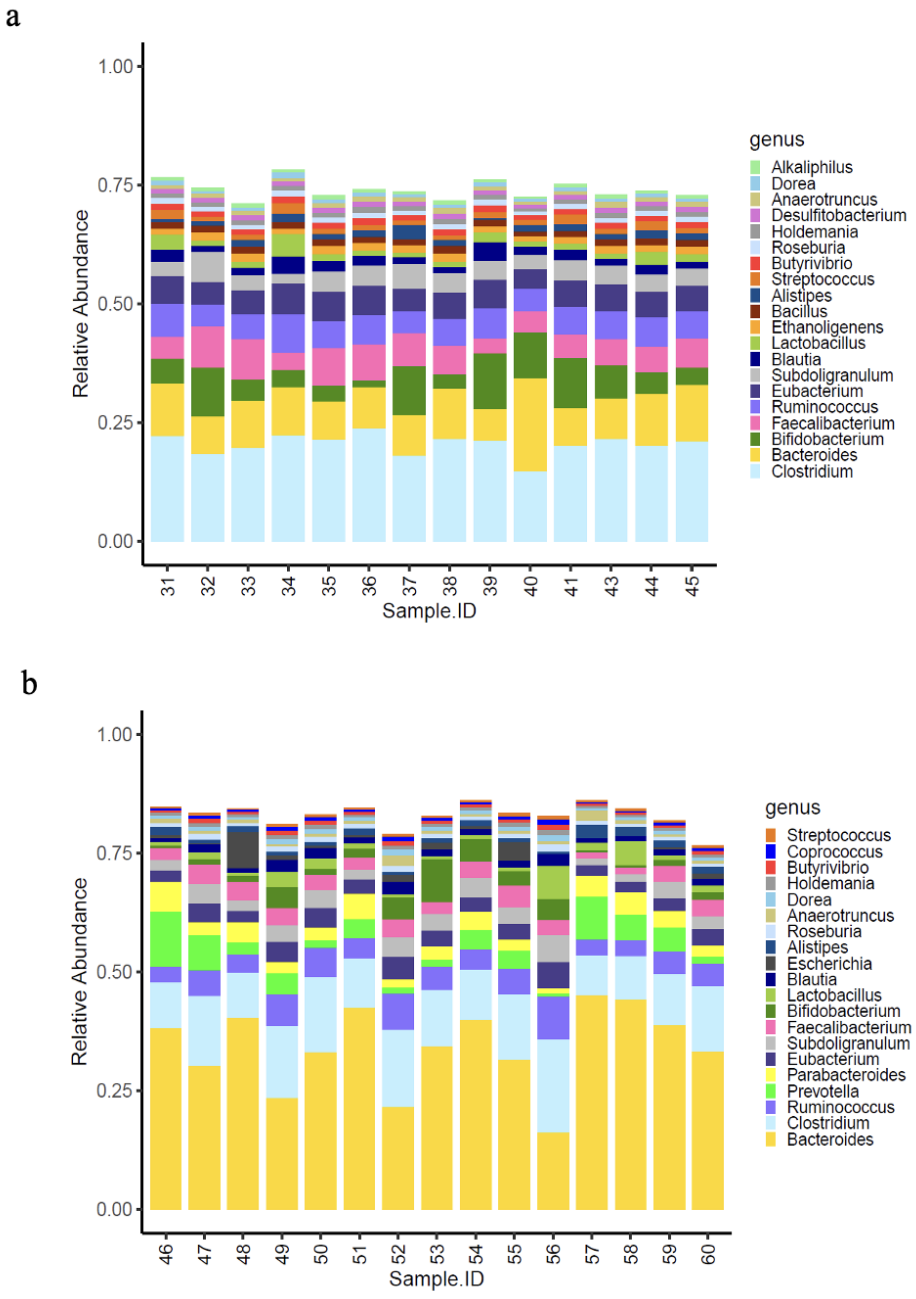
Broiler number	Farm	Metagenome ID caeca	N. of reads	Gbp	Metagenome ID carcass	N. of reads	Gbp
31	Conventional	1343	40662712	5.95	1280	44582923	3.31
32	Conventional	1344	44858014	6.56	1281	48151463	3.59
33	Conventional	1345	40333432	5.81	1282	48778635	3.64
34	Conventional	1346	35639899	4.81	1283	51922499	3.85
35	Conventional	1347	42948477	5.80	1284	76815286	5.77
36	Conventional	1348	48924902	7.14	1285	50869653	3.72
37	Conventional	1349	38427233	5.54	1286	7116609	534.11 (Mb)
38	Conventional	1350	136407195	19.28	1287	62175940	4.60
39	Conventional	1351	38701919	5.44	1288	50420635	3.78
40	Conventional	1352	34100337	4.70	1289	61001877	4.57
41	Conventional	1353	35244995	5.11	1290	49828965	7.28
42	Conventional	-	-	-	1291	43162564	3.24
43	Conventional	1355	67927139	10.00	1292	40665704	3.05
44	Conventional	1356	40043411	5,78	1293	46507602	6.85
45	Conventional	1357	104579787	14.88	1294	36252353	2.72
46	Antibiotic free	1358	52954003	3.98	1295	131516916	9.87
47	Antibiotic free	1359	41827994	3.14	1296	648195	48.73 (Mb)
48	Antibiotic free	1360	51865847	3.90	1297	38071848	2.86
49	Antibiotic free	1361	64448503	4.72	1298	34826163	2.61
50	Antibiotic free	1362	59741307	4.46	1299	69774102	5.23
51	Antibiotic free	1363	45802509	3.42	1300	66262603	4.96
52	Antibiotic free	1364	43535017	3.27	1301	54383524	4.04
53	Antibiotic free	1365	35032973	2.63	1302	38138345	5.49
54	Antibiotic free	1366	45990811	3.45	1303	195441747	14.63
55	Antibiotic free	1367	47668525	3.58	1304	46202148	6.72
56	Antibiotic free	1368	41586126	3.13	1305	58668081	4.40
57	Antibiotic free	1369	46393713	3.49	1306	50889023	3.80
58	Antibiotic free	1370	50371723	3.79	1307	31114983	2.33

59	Antibiotic free	1371	48911269	3.67	1308	46767036	6.74
60	Antibiotic free	1372	34825336	4.86	1309	49686448	3.72

Taxonomic and functional gene composition of the caeca contents

The taxonomic and functional gene composition was investigated in the caeca of birds reared in the conventional and antibiotic free farm. Sixteen of the top 20 most abundant genera identified were shared between the two tested groups (Figure 4). Moreover, *Alkaliphilus*, *Desulfibacterium*, *Bacillus* and *Ethanoligenens* were detected in the caeca of birds from the conventional farm (Figure 4a); *Coprococcus*, *Escherichia*, *Parabacteroides* and *Provotella* were detected in the caeca of birds from the antibiotic free farm (Figure 4b).

Figure 4: Distribution top 20 genera in the bar plots characterizing the caeca of the birds reared in the conventional (a) and antibiotic free (b) farm. Sample 42 is not included among samples in panel a because it was not processed for technical reasons.



The normalized mean counts of *Alkaliphilus*, *Bacillus*, *Desulfitobacterium*, *Ethanoligenens* and *Streptococcus* were significantly higher in the caeca of birds reared in the conventional farm, while those of *Alistipes*, *Anaerotruncus*, *Bacteroides*, *Coprococcus*, *Dorea*, *Escherichia*, *Holdemania*, *Lactobacillus*, *Parabacteroides*, *Prevotella*, *Roseburia*, *Ruminococcus* and *Subdoligranulum* in the caeca of birds reared in the antibiotic free farm (Table 21).

Table 21: Normalized mean values (DESeq2 normalized counts) of the genera belonging in the top 20 genera significantly different ($p<0.05$) in the caeca of the birds reared in the conventional and antibiotic free farm.

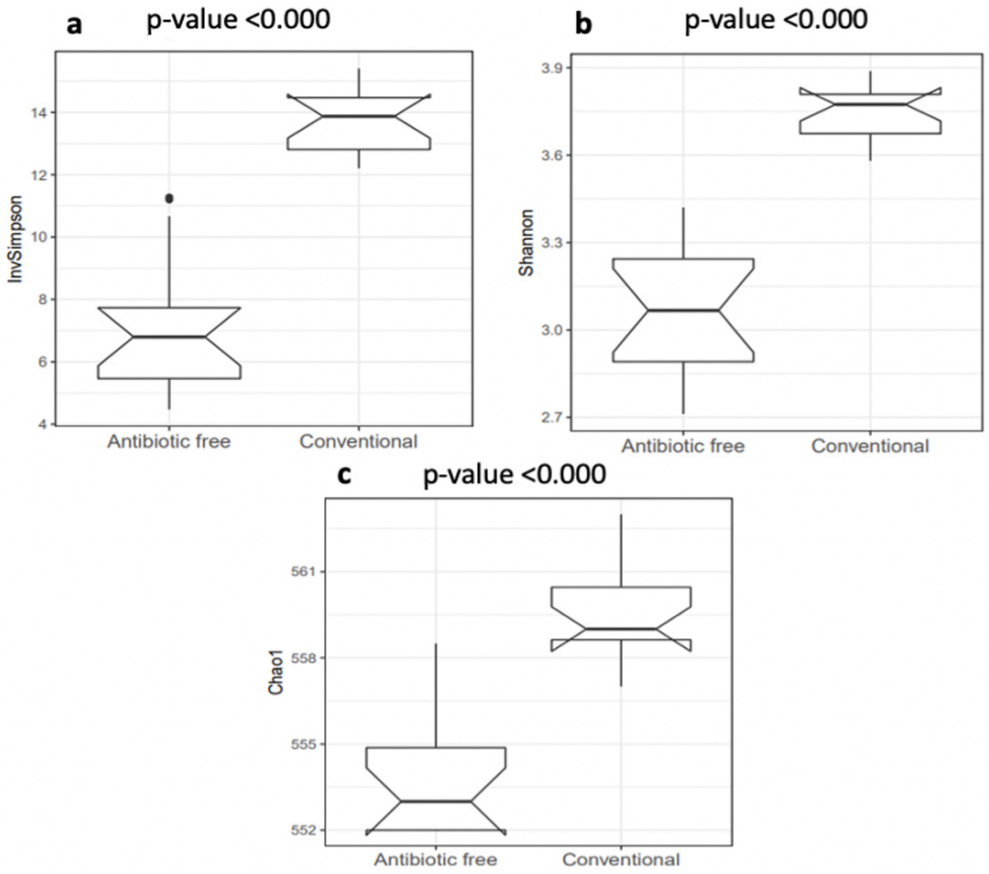
Genus	**Norm_Mean_ Conventional**	**Norm_Mean_ Antibiotic free**	**p-value**
Alistipes	121394.48	201764.20	0.00
Alkaliphilus	65854.66	37406.43	0.00
Anaerotruncus	81830.35	141769.05	0.00
Bacillus	121574.41	80982.30	0.00
Bacteroides	903112.03	5541950.68	0.00
Coprococcus	56576.51	94133.58	0.00
Desulfitobacterium	83821.62	49267.15	0.00
Dorea	69937.78	121675.15	0.00
Escherichia	40613.94	215143.60	0.00
Ethanoligenens	126272.74	80004.04	0.00
Holdemania	87562.75	113518.34	0.00
Lactobacillus	162869.83	297992.16	0.02
Parabacteroides	54047.67	557109.97	0.00
Prevotella	41748.15	697359.16	0.00
Roseburia	95006.08	151647.56	0.00
Ruminococcus	522286.96	806960.48	0.00
Streptococcus	122827.12	90076.97	0.01
Subdoligranulum	342496.63	505311.50	0.00

The genera richness and diversity observed in the caeca of birds in terms of alfa diversity were estimated using the InvSimpson, Shannon and Chao1 indexes and were significantly higher in the caeca of birds from the group of animals reared in the conventional farm in comparison to the antibiotic free farm (Table 22 and Figure 5).

Table 22: Mean values of the InvSimpson, Shannon, and Chao1 indexes quantified for the genera colonizing in the caeca of birds reared in the conventional and antibiotic free farm.

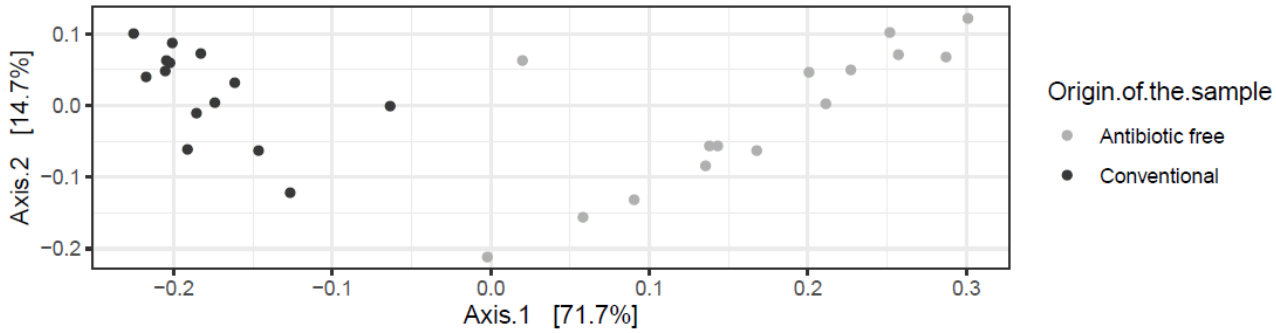
	InvSimpson	**Shannon**	**Chao1**
Mean_Conventional	13.73	3.75	559.63
Mean_Antibiotic free	7.08	3.07	554.05

Figure 5: Box plots of the alpha diversity indexes calculated at genus level: (a) InvSimpson index; (b) Shannon index; (c) Chao1 index. p value <0.001 were considered significantly different.



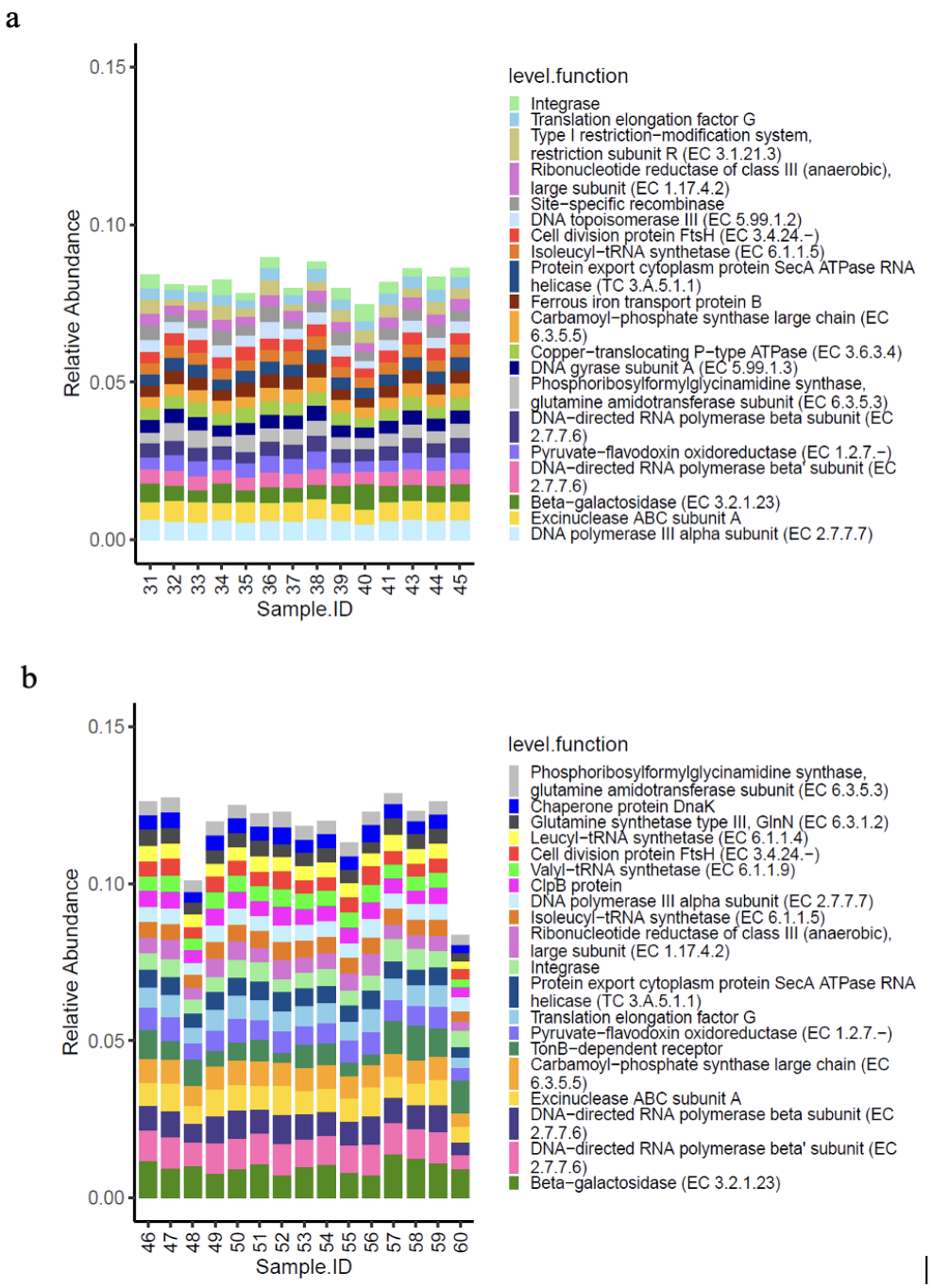
The beta diversity index (generated using the Bray-Curtis distance metric) calculated at the genus level highlighted a clear dissimilarity in the community composition in the caeca sampled in the two groups of broilers tested (adonis2 p-value<0.00001) (Figure 6).

Figure 6: Bray-Curtis dissimilarity plots showing the genera detected in the caeca of birds from the conventional and antibiotic free farm.



Fourteen of the top 20 most abundant functional genes identified within the caeca of birds were shared between the two tested groups (Figure 7). Moreover, functional genes type I restriction−modification system-restriction subunit R (EC 3.1.21.3), site-specific recombinase, DNA topoisomerase III (EC 5.99.1.2), ferrous iron transport protein B, copper−translocating P−type ATPase (EC 3.6.3.4) and DNA gyrase subunit A (EC 5.99.1.3) were listed among the most abundant functional genes in the caeca of birds from the conventional farm (Figure 7a), while chaperone protein DnaK, glutamine synthetase type III, GlnN (EC 6.3.1.2), leucyl−tRNA synthetase (EC 6.1.1.4), valyl−tRNA synthetase (EC 6.1.1.9), clpB protein and tonB−dependent receptor were detected only in the caeca of birds from the antibiotic free farm (Figure 7b).

Figure 7: Distribution top 20 functional genes characterizing the caeca of the birds reared in the conventional (a) and antibiotic free (b) farm. Sample 42 is not included among samples in panel a because it was not processed for technical reasons.



Besides these qualitative differences, the normalized mean values of beta-galactosidase (EC 3.2.1.23) and DNA topoisomerase III EC 5.99.1.2 (in PFGI-1-like cluster) were significantly higher in the caeca of birds from the conventional farm, while those of integrase, translation elongation factor G, ribonucleotide reductase of class III (anaerobic)-large subunit EC 1.17.4.2, excinuclease ABC subunit

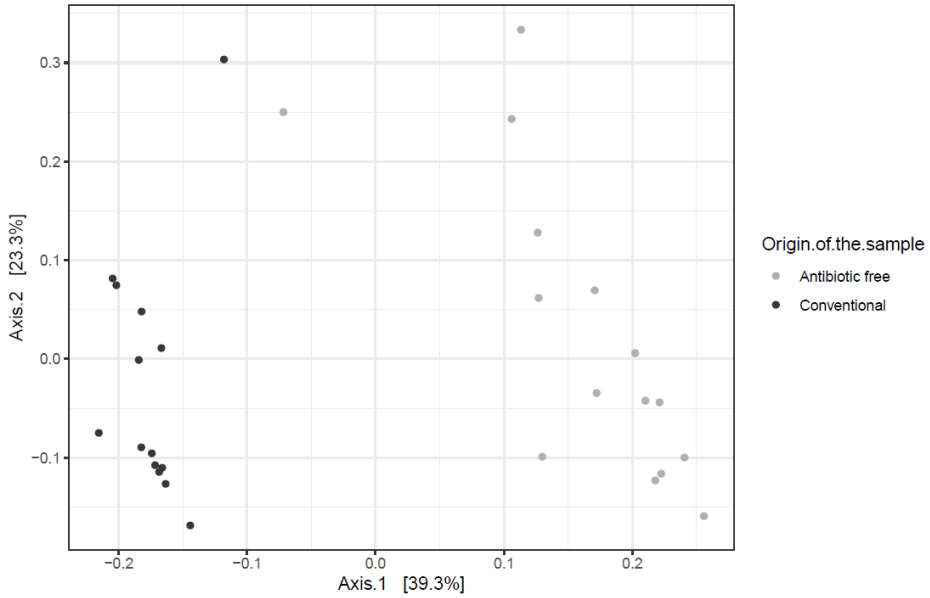
A paralog in greater Bacteroides group and chaperone protein DnaK were significantly higher in the caeca of birds reared in the antibiotic free farm (Table 23).

Table 23: Normalized mean values (DESeq2 normalized counts) of the functional genes belonging in the top 20 functional genes significantly different ($p<0.05$) in the caeca of the birds reared in the conventional and antibiotic free farm.

Level function	Norm_Mean_ Conventional	Norm_Mean_ Antibiotic free	p-value
Beta-galactosidase (EC 3.2.1.23)	7282.34	0	0.00
Chaperone protein DnaK	9293.18	12738.49	0.00
DNA topoisomerase III (EC 5.99.1.2) in PFGI-1-like cluster	2.01	0.21	0.00
Excinuclease ABC subunit A paralog in greater Bacteroides group	410.25	2013.30	0.00
Integrase	9163.52	15684.38	0.00
Ribonucleotide reductase of class III (anaerobic) large subunit (EC 1.17.4.2)	9822.28	14658.28	0.00
Translation elongation factor G	9751.14	18514.22	0.00

The beta diversity index calculated for the functional genes confirmed a clear association of the caeca with their farm of origin (adonis2 p-value<0.00001) (Figure 8).

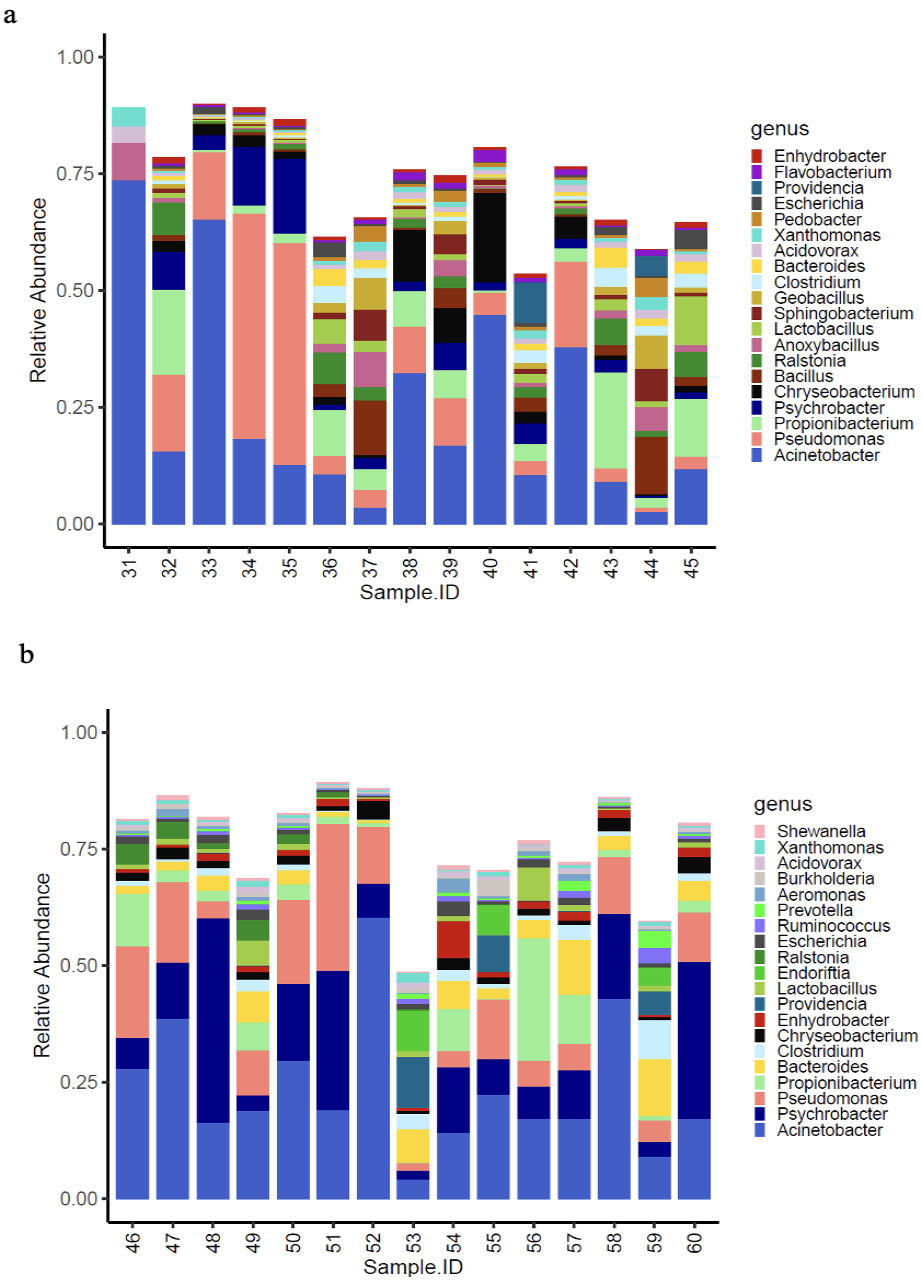
Figure 8: Bray-Curtis dissimilarity plots showing the functional genes detected in caeca of birds reared in the antibiotic free and conventional farm.



Taxonomic and functional gene composition of the carcasses

Overall, 14 of the top 20 most abundant genera identified on carcasses were shared between the two tested groups (Figure 9). Moreover, *Anoxybacillus*, *Bacillus*, *Flavobacterium*, *Pedobacter*, *Geobacillus* and *Sphingobacterium* were only detected on carcasses of birds from the conventional farm (Figure 9a); *Aeromonas*, *Burkholderia*, *Endoriftia*, *Prevotella*, *Ruminococcus* and *Shewanella* on carcasses of birds from the antibiotic free farm (Figure 9b).

Figure 9: Distribution of the top 20 genera in the bar plots that characterizing the carcasses of the birds reared in the conventional (a) and antibiotic free (b) farm.



The normalized mean values of *Anoxybacillus*, *Bacillus*, *Gebacillus*, *Pedobacter* and *Sphingobacterium* were significantly higher on carcasses of birds reared in the conventional farm, while those of *Aeromonas*, *Bacteroides*, *Prevotella* and *Ruminococcus* on carcasses of birds reared in the antibiotic free farm (Table 24).

Table 24: Normalized mean values (DESeq2 normalized counts) of the genera belonging in the top 20 genera significantly different ($p<0.05$) on carcasses of the birds reared in the conventional and antibiotic free farm.

Genus	**Norm_Mean_ Conventional**	**Norm_Mean_ Antibiotic free**	**p-value**
Aeromonas	118.65	325.11	0.00
Anoxybacillus	702.12	14.22	0.00
Bacillus	1032.50	183.12	0.00
Bacteroides	520.73	2011.12	0.00
Geobacillus	616.23	40.34	0.00
Pedobacter	429.52	127.03	0.00
Prevotella	54.76	339.25	0.00
Ruminococcus	142.06	361.58	0.00
Sphingobacterium	684.45	188.17	0.00

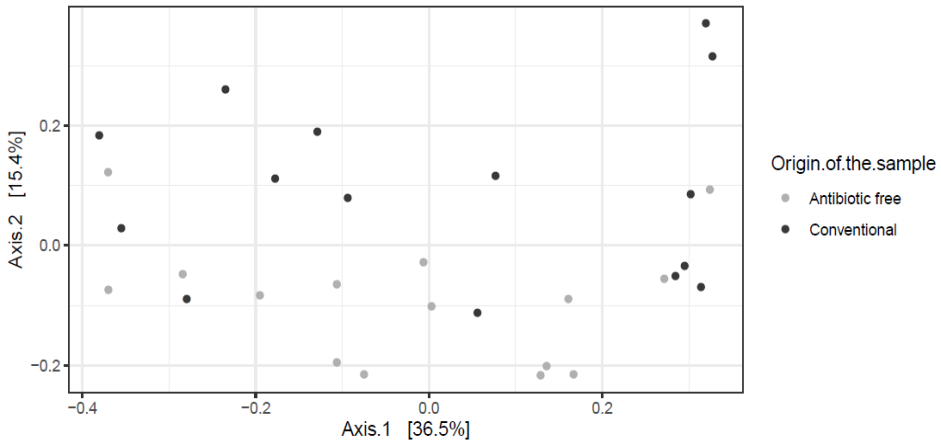
The alpha diversity indexes (i.e., InvSimpson, Shannon and Chao1) calculated for bacteria genera identified on carcasses from conventional and antibiotic free farm did not show any significative difference (Table 25).

Table 25: Mean values and corresponding p values of the InvSimpson. Shannon. Chao1 indexes quantified for the genera identified on carcasses of birds reared in the conventional and antibiotic free farm.

Index	**Mean Conventional**	**Mean Antibiotic free**	**p-value**
InvSimpson	13.312	10.74	0.44
Shannon	3.23	3.19	0.88
Chao1	491.49	503.36	0.75

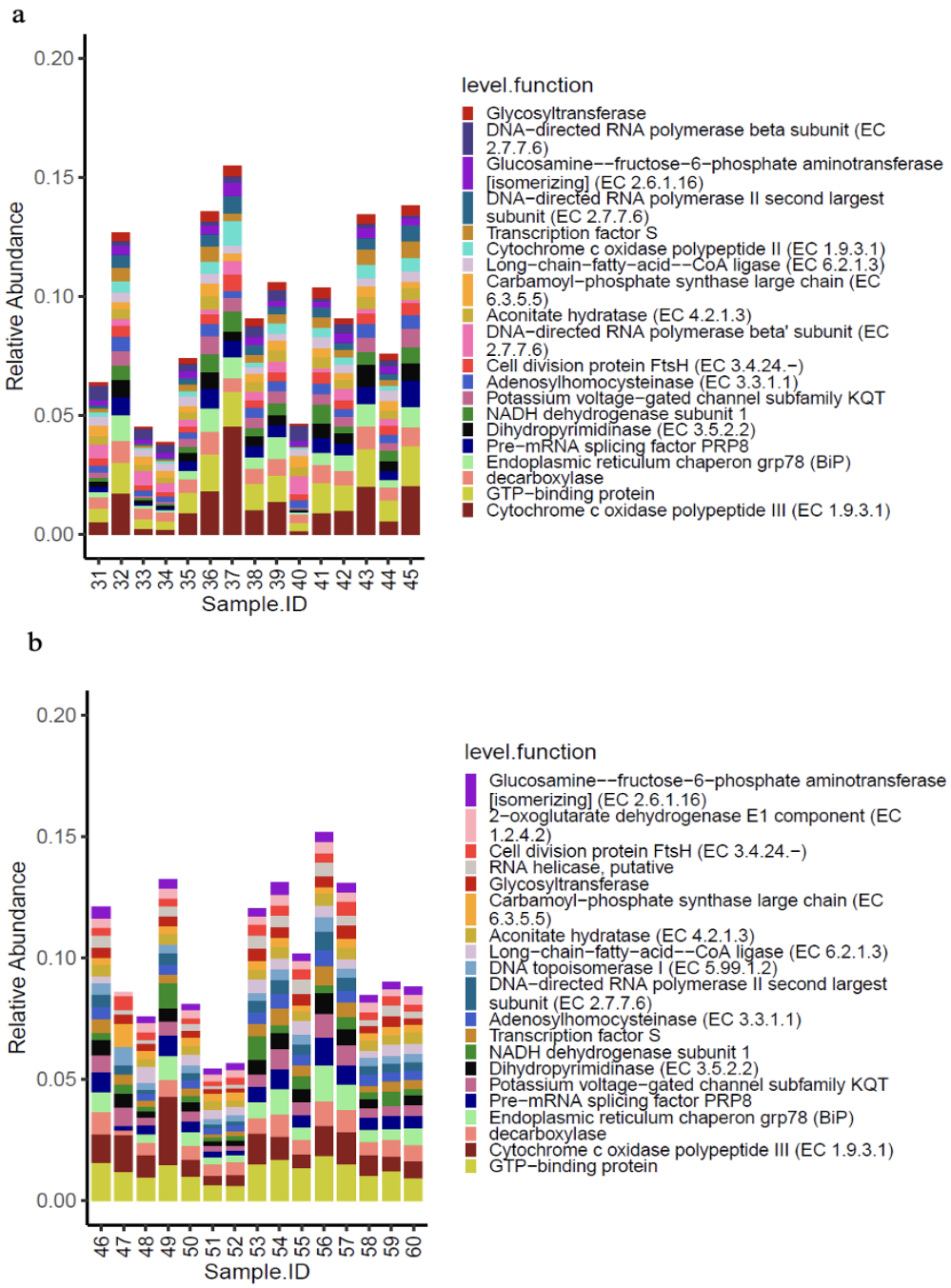
The beta diversity calculated for the functional genes did not group the carcasses according to the farm origin (adonis2 p-value=0.332) (Figure 10).

Figure 10: Bray-Curtis dissimilarity plots showing the genera detected on carcasses from the conventional and antibiotic free farm.



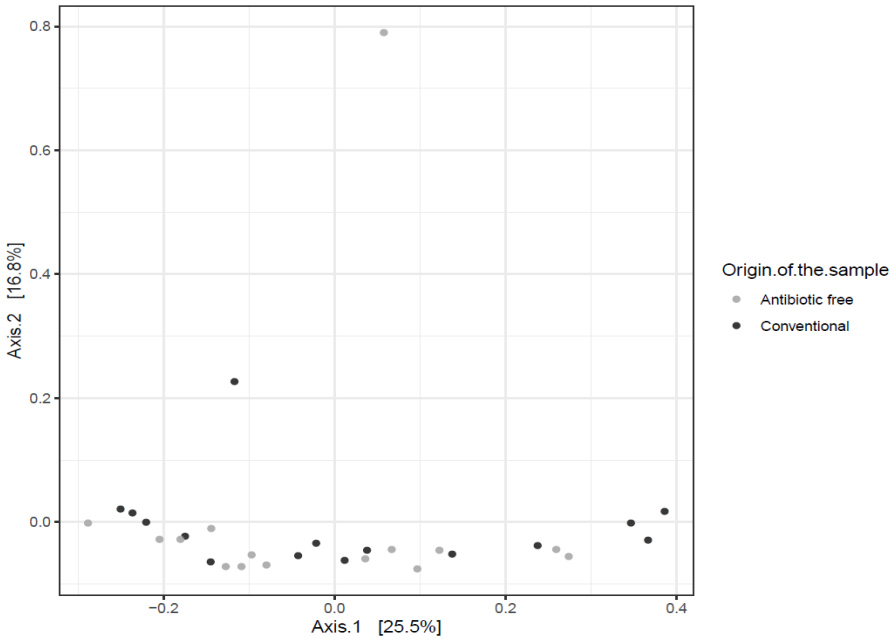
Three of the top 20 most abundant functional genes identified on the carcasses were shared between the two tested groups (Figure 11). Moreover, functional genes cytochrome c oxidase polypeptide II (EC 1.9.3.1) and DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6) were listed among the most abundant functional genes on carcasses of birds from the conventional farm (Figure 11a), while 2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2), RNA helicase putative and DNA topoisomerase I (EC 5.99.1.2) only on carcasses of birds from the antibiotic free farm (Figure 11b).

Figure 11: Distribution top 20 functional genes characterizing the carcasses of the birds reared in the conventional farm (a) and antibiotic free (b).



The functional genes belonging of the top 20 identified in the carcasses did not show a significative difference of normalized mean counts between the two groups of carcasses tested ($p<0.05$). Moreover, the beta diversity generated using the Bray-Curtis distance metric, observed in the functional composition of the two tested groups did not cluster the carcasses according to the farm of origin (adonis2 p-value=0.332) (Figure 12).

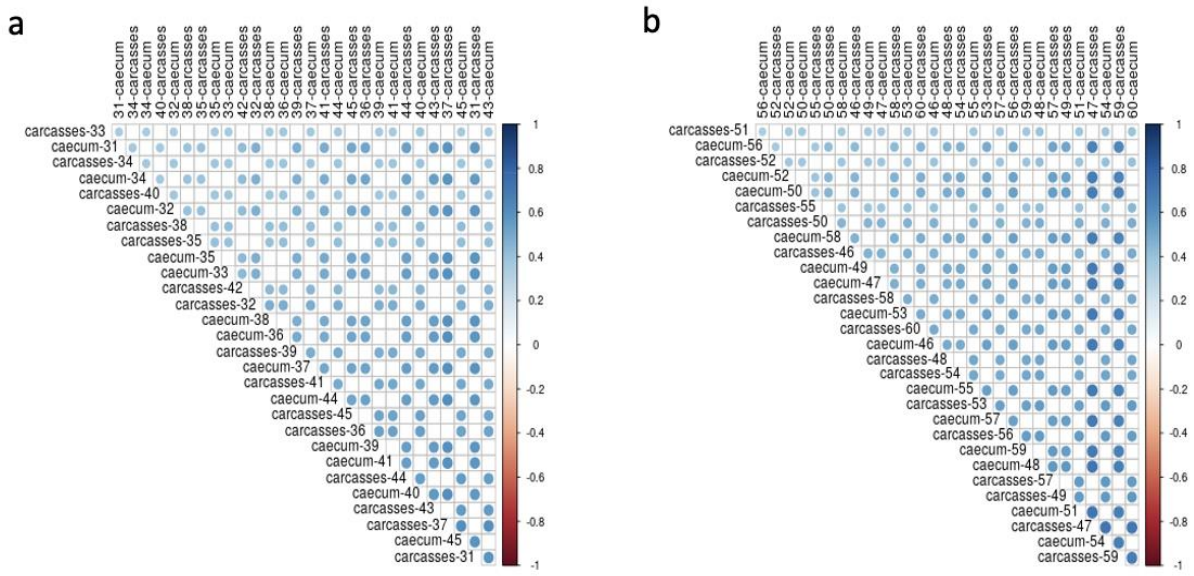
Figure 12: Bray-Curtis plot of beta diversity of functional genes detected on carcasses of birds reared in the antibiotic free and conventional farm.



Longitudinal analysis of caeca and carcass belonging to the same animal

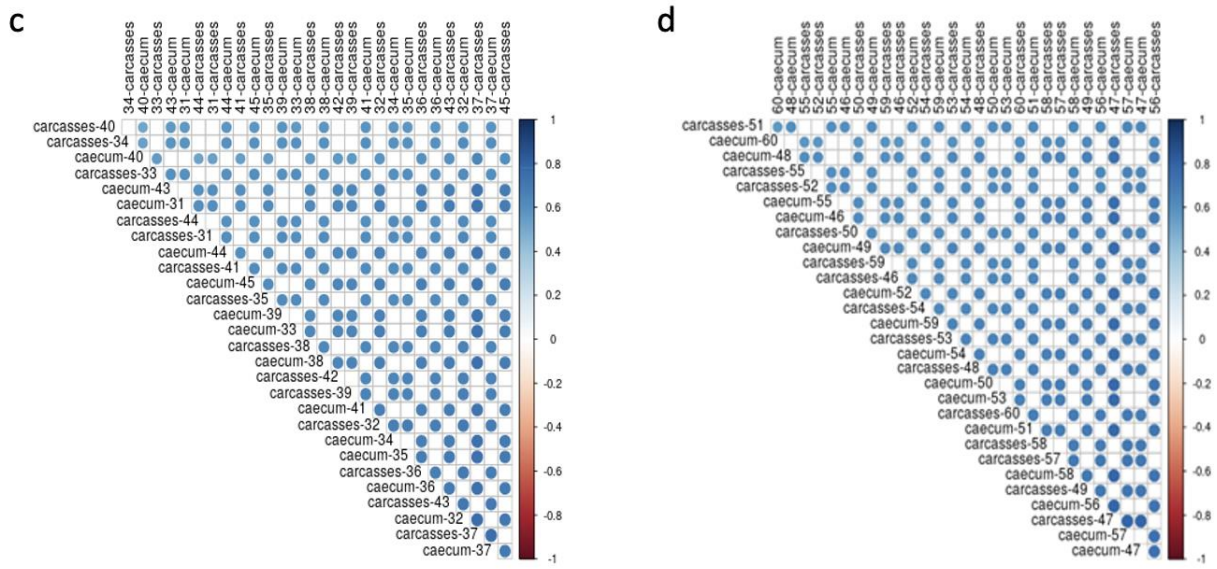
In Figure 13 the correlations of genera identified between caeca samples and carcasses of birds reared in the conventional and antibiotic free farm are shown. Both the dimension of the circle and the colour scale represent the value of the correlation coefficient. Overall, a positive correlation was always detected between the genera colonizing the caeca and corresponding carcass. Nevertheless, such correlation was not stronger for samples collected from the same animal than for animals belonging to the same flock. As an example, in Figure 13a the bacteria genera identified in the caeca of the bird labelled as 31 are well correlated to those of carcass 31 but as well as to carcasses 36, 37, 43, and 45. On the contrary, in Figure 13b the bacteria genera identified in the caeca of the bird labelled as 50 show a higher correlation with the carcasses 47 and 59 in comparison to the carcass of the same animal (i.e., carcass 50).

Figure 13: Kendall correlation plot of genera identified in the caeca and corresponding carcasses of birds reared in the conventional farm (a) and of genera identified in the caeca and corresponding carcasses of birds reared in the antibiotic free farm (b).



A positive correlation was also identified between functional genes (Figure 14) detected in the caeca and corresponding carcass for birds reared in the conventional (Figure 14c) and antibiotic free (Figure 14d) farm. For the functional genes such positive correlation was higher in comparison to that observed for the bacteria genera thus resulting in stronger blue dots.

Figure 14: Kendall correlation plot of functional genes identified in the caeca and corresponding carcasses of birds reared in the conventional farm (c) and functional genes identified in the caeca and corresponding carcasses of birds reared in the antibiotic free farm (d).



Identification of antibiotic resistance genes in the caeca and carcasses

The antibiotic resistance genes (ARGs) were retrieved from functional genes identified in the caeca and in the carcasses of the tested animals and classified as such in level 2 of the biological-function ontology in SEED category (Aziz *et al*., 2008). Among the ARGs with normalized mean values of abundance >1000 in at least one tested group, the regulatory sensor-transducer-BlaR1/MecR1 family, UDP-N-acetylmuramoylalanyl-D-glutamate-2.6-diaminopimelate ligase (EC 6.3.2.13), macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-), multi antimicrobial extrusion protein (Na(+)/drug antiporter)-MATE family of MDR efflux pumps. topoisomerase IV subunit B (EC 5.99.1.-) and vancomycin response regulator VanR were significantly higher in the caeca of birds reared in the conventional farm in comparison to the antibiotic free farm (Table 26). On the contrary, the acriflavin resistance protein and probable RND efflux membrane fusion protein was significantly higher in the caeca of birds reared in the antibiotic free farm (Table 26). As far as the carcasses are concerned, no differences were identified between normalized mean values of abundance of ARGs detected on carcasses sampled in the two tested groups.

Table 26: Genes classified as antibiotic resistant genes in the Level 2 category and showing normalized mean values significantly different ($p<0.05$) in the caeca of birds reared in the conventional and antibiotic free farm. Genes in bold have been identified also on carcasses but with any normalized mean values significantly different in the two tested groups.

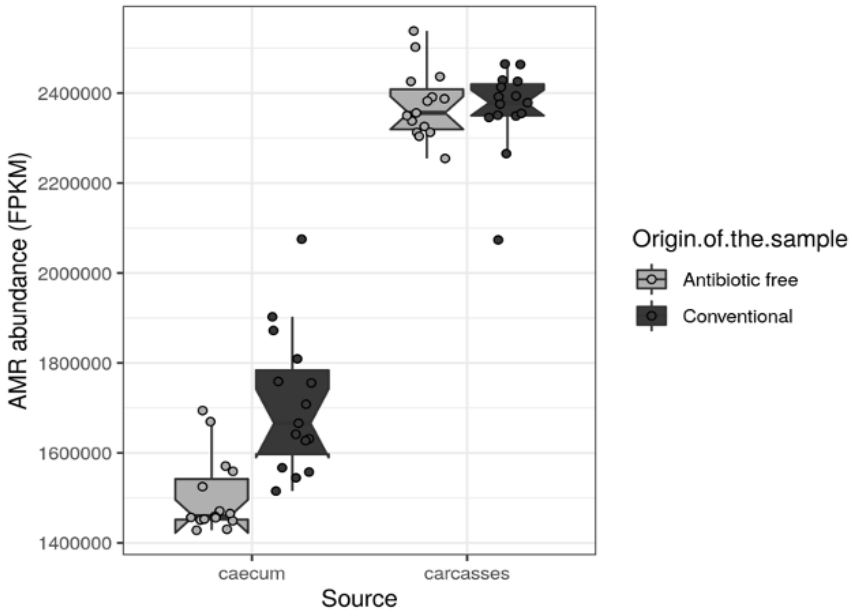
Level 3	Level function	Norm_Mean Conventional	Norm_Mean Antibiotic free	p-value
Aminoglycoside_adenylyltransferases	Aminoglycoside N6'-acetyltransferase (EC 2.3.1.82)	8.09	0.74	0.00
	Spectinomycin 9-O-adenylyltransferase	95.17	19.10	0.00
	Streptomycin 3''-O-adenylyltransferase (EC 2.7.7.47)	115.47	324.20	0.00
Beta-lactamase	Beta-lactamase class A	6.45	0.18	0.02
	Beta-lactamase class C and other penicillin binding proteins	162.99	29.85	0.00
	Beta-lactamase repressor BlaI	395.08	36.98	0.00
	Metal-dependent hydrolases of the beta-lactamase superfamily II	16.52	2.47	0.00
	Negative regulator of beta-lactamase expression	5.18	0.90	0.00
	Regulatory protein BlaR1	23.20	0.51	0.00
BlaR1_Family_Regulatory_Sensor-transducer_Disambiguation	Beta-lactamase repressor BlaI	0	51.57	0.00
	Regulatory sensor-transducer. BlaR1/MecR1 family	1031.43	412.74	0.00
	Transcriptional regulator. MecI family	33.46	83.80	0.00
	Transcriptional repressor. BlaI/MecI family	261.79	462.16	0.00
Erythromycin_resistance	rRNA adenine N-6-methyltransferase (EC 2.1.1.48)	190.67	712.31	0.00
	rRNA adenine N-6-methyltransferase (EC 2.1.1.48) (Macrolide-lincosamide-streptogramin B resistance protein) (Erythromycin resistance protein)	23.96	1.71	0.00
	rRNA adenine N-6-methyltransferase (erm/ksgA)	3.80	16.21	0.00
Methicillin_resistance_in_Staphylococci	FmtA protein involved in methicillin resistance	3.76	0.48	0.02
	Transposase for insertion sequence-like element IS431mec	0.62	4.95	0.01
	UDP-N-acetylmuramoylalanyl-D-glutamate--2.6-diaminopimelate ligase (EC 6.3.2.13)	2541.38	0	0.00
MexE-MexF-OprN_Multidrug_Efflux_System	**Multidrug efflux transporter MexF**	67.03	32.71	0.00
	Transcriptional regulator MexT	1.37	0.11	0.09
	Acriflavin resistance protein	2292.45	7129.56	0.00

Multidrug_Resistance_Efflux_Pumps	**Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)**	1036.87	426.51	0.00
	Macrolide-specific efflux protein MacA	210.92	668.43	0.00
	Multi antimicrobial extrusion protein (Na(+)/drug antiporter). MATE family of MDR efflux pumps	8815.67	3125.98	0.00
	Multidrug efflux pump component MtrF	19.15	2.95	0.00
	Multidrug efflux RND membrane fusion protein MexC	4.11	1.52	0.02
	Multidrug efflux RND transporter MexD	19.08	6.19	0.00
	Multidrug resistance efflux pump PmrA	20.99	3.97	0.00
	Multidrug-efflux transporter. major facilitator superfamily (MFS) (TC 2.A.1)	14.73	47.62	0.00
	Probable outer membrane component of multidrug efflux pump	0.54	4.49	0.01
	Transcription regulator of multidrug efflux pump operon. TetR (AcrR) family	22.02	4.14	0.00
	Transcription repressor of multidrug efflux pump acrAB operon. TetR (AcrR) family	4.97	24.05	0.00
Multiple_Antibiotic_Resistance_MAR_locus	Multiple antibiotic resistance protein MarA	0.31	3.61	0.00
Polymyxin_Synthetase_Gene_Cluster_in_Bacillus	Polymyxin synthetase PmxA	1.95	0.26	0.03
Resistance_to_fluoroquinolones	Efflux pump Lde	25.94	69.77	0.02
	Topoisomerase IV subunit B (EC 5.99.1.-)	1796.57	0	0.00
Resistance_to_Vancomycin	Protein VanZ	1.47	0.09	0.03
	Sensor histidine kinase VanS (EC 2.7.3.-)	847.06	136.85	0.00
	Vancomycin B-type resistance protein VanW	882.44	147.84	0.00
	Vancomycin B-type resistance protein VanX	65.05	29.34	0.00
	Vancomycin resistance protein VanH	11.77	4.74	0.02
	Vancomycin response regulator VanR	1340.46	856.22	0.01
	Vancomycin Teicoplanin A-type resistance protein VanA	44.30	3.05	0.00
Streptococcus_pneumoniae_Vancomycin_Tolerance_Locus	ABC transporter. ATP-binding protein Vex2	84.10	8.61	0.00

Streptothricin_resistance	**Streptothricin acetyltransferase. Streptomyces lavendulae type**	149.21	248.09	0.00
The_mdtABCD_multidrug_resistance_cluster	**Multidrug transporter MdtB**	16.73	79.84	0.01
	Multidrug transporter MdtC	15.58	69.33	0.00
	Multidrug transporter MdtD	7.58	32.85	0.00
	Probable RND efflux membrane fusion protein	709.05	1590.56	0.00
	Response regulator BaeR	53.50	23.84	0.00
	Sensory histidine kinase BaeS	16.45	35.31	0.02

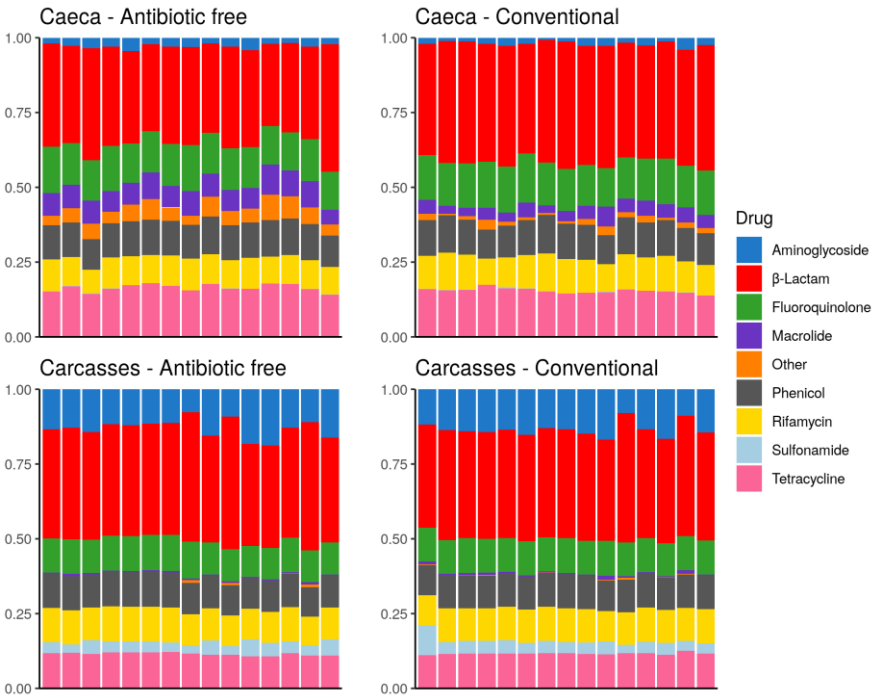
The total antimicrobial resistance (AMR) load per sample in the caeca was significantly higher in the birds reared in the conventional in comparison to the antibiotic free (Wilcoxon rank sum test p-value = 0.00009) (Figure 15). However, in both groups of caeca samples, the total AMR load was lower in comparison to the carcasses which did not show significant differences between the two tested groups (Wilcoxon rank sum test p-value = 0.6).

Figure 15: Box plots showing the total AMR level (FPKM) per sample, stratified by source and origin of the sample. Each sample is also represented by a dot with sideways jitter to minimize overplotting.



The drug classes identified in the broiler caeca and on carcasses are listed in Figure 16. The drug class aminoglycoside was more represented on carcasses in comparison to caeca contents and sulfonamide was identified on carcasses but at very low levels in some caeca. On the contrary, macrolide as well as resistance to other drug classes, including bicyclomicyn, lincosamide, fosfomycin, glycopeptide, pleuromutilin and nitrofuran were mostly identified in the caeca. Figure 16 shows that besides differences in the abundance of specific antibiotic resistance genes described above, in qualitative terms the overall resistome of the caeca of animals reared on the antibiotic free farm overlaps with that of animals reared on the conventional farm, and the same is observed on carcasses. This result can be explained considering that the antibiotic free flock was reared in a farm where antibiotics have been possibly used in the previous flocks, thus supporting the persistence of ARGs in the farm environment over time.

Figure 16: Stacked bar chart of AMR abundance (FPKM) per drug class (colors) per sample (x axis). Each plot refers to one group of samples as detailed in the sub-titles.



As for bacteria genera and functional genes, the ARGs identified in the caeca also clustered separately for the conventional and antibiotic free farms (adonis2 p-value = 0.00001) (Figure 17), while in the carcasses this difference was lost (adonis2 p-value = 0.4278) (Figure 18).

Figure 17: Beta diversity of caeca samples shown as PCoA of Bray–Curtis diversity computed based on the AMR gene family abundances normalized with DESeq2.

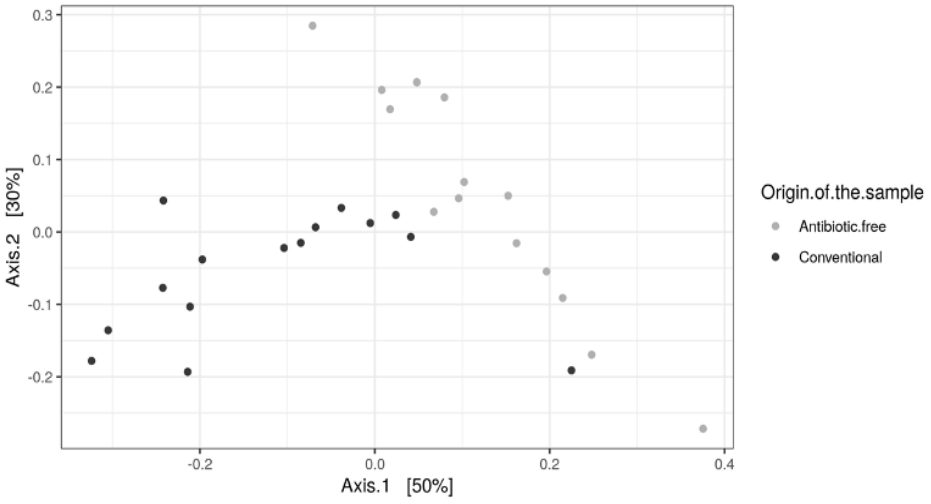
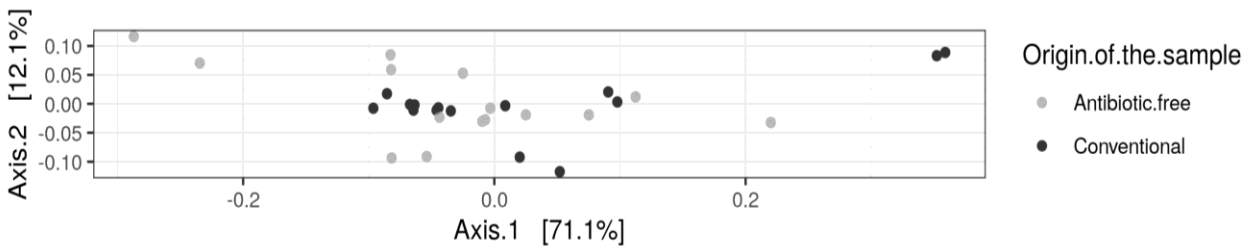


Figure 18: Beta diversity of carcasses samples shown as PCoA of Bray-Curtis diversity computed based on the AMR gene family abundances normalized with DESeq2.



4.3 Study 3

In the study 3, a total of 42 metagenomes were obtained. The output of each metagenome is detailed in Table 27. In this study, the taxonomic composition of the three groups of carcasses was investigated to check differences between carcasses from a conventional farm (Group 1) and two antibiotic free farms (Groups 2 and 3). The carcasses investigated in this study were all slaughtered in the same facility but were not processed as first group of the day. Therefore, their composition was affected by cross contaminations during transport and slaughtering. As far as the age of the animals is concerned, carcasses of Group 1 were obtained from broilers slaughtered at day 35; carcasses of Group 2 were obtained from broilers slaughtered at day 34; carcasses of Group 3 were obtained from broilers slaughtered at day 33. In all tested carcasses, both *Salmonella* and *Campylobacter* were not detected applying both microbiological investigations. According to the results of study 1, the taxonomic composition of these carcasses was assessed using Mg-RAST and the database RefSeq.

Table 27: Numbers identifying the tested metagenomes in each investigated group. The output of each metagenome is detailed as number of sequences and bp.

Metagenomes of Group 1		
Metagenome ID	**N. of reads**	**Gbp**
1790_R	7871004	1.63
1793_R	9316148	1.95
1794	35982973	5.30
1795_R	12727886	2.55
1796	40924658	5.83
1797	37346427	5.45
1798_R	25587789	4.33
1799_R	1705857	345.63 (Mb)
1800	34400324	5.05
1801_R	14446591	3.02
1802	34324700	4.63
1803_R	9576848	2.01

Metagenomes of Group 2		
1820	26892790	3.74
1821	34129544	4.92
1822	34497704	5.00
1823	41672368	5.70
1824	36015129	5.29
1825	44140109	6.12
1826	43116805	5.80
1827	30864299	4.55
1828	35774833	5.28
1829	39316715	5.75
1830_R	6444884	1.37
1831	25004927	3.63
1832	27256510	3.90
1833	24318290	3.48
1834	37470587	5.21
Metagenomes of Group 3		
1850	37413222	5.43
1851_R	8286180	1.71
1852	24592691	3.51
1853	30984826	4.44
1854	25708165	3.62
1855_R	9528735	1.99
1856	28169226	4.06
1857	24735037	3.55
1858_R	8271062	1.74
1859	31279180	4.52
1860	41996917	5.98
1861_R	8288893	1.72
1862	29680011	4.25
1863	38074498	3.45
1864_R	149677	30.20 (Mb)

Taxonomic classification of the carcasses

The taxonomic composition of the tested carcasses was analysed at all taxonomic levels. The most represented phyla are listed in Table 28. Proteobacteria was significantly higher on the carcasses belonging to Group 3 in comparison to Groups 1 and 2. On the contrary, Actinobacteria and Firmicutes showed a significantly lower mean relative frequency of abundance in Group 3 compared to the other two tested Groups. Finally, the phylum Bacteroidetes showed a significantly higher mean relative frequency of abundance in Group 1 compared to both Groups 2 and 3.

Table 28: Phyla with mean values of the relative frequency of abundance (%) ≥1 in at least one group of carcasses and significantly different in the tested groups (p≤0.05).

Phylum	mean±SD			p-value
	Group 1	Group 2	Group 3	
Actinobacteria	4.52±1.95	3.35±2.07	1.75±1.17	0.00
Bacteroidetes	10.62±4.19	4.96±3.19	2.37±1.39	0.00
Firmicutes	28.78±6.12	27.57±17.96	12.61±6.82	0.00
Proteobacteria	53.41±7.42	60.31±22.74	81.94±8.77	0.00

At class level, Gammaproteobacteria was significantly higher in Group 3 in comparison to Groups 1 and 2 while Actinobacteria, Alphaproteobacteria and Negativicutes were significantly lower on carcasses belonging to Group 3 compared to Groups 1 and 2 (Table 29). On the contrary, the mean relative frequency of abundance of Bacilli, Bacteroidia, Betaproteobacteria and Flavobacteria were significantly higher in Group 1 compared to both Groups 2 and 3.

Table 29: Classes with mean values of the relative frequency of abundance (%) ≥ 1 in at least one group of carcasses and significantly different in the tested groups (p≤0.05).

Class	mean±SD			p-value
	Group 1	Group 2	Group 3	
Actinobacteria	4.52±1.95	3.35±2.07	1.75±1.17	0.00
Alphaproteobacteria	1.26±0.41	1.12±0.52	0.76±0.43	0.03
Bacilli	13.96±5.28	9.32±4.59	4.52±4.31	0.00
Bacteroidia	7.06±2.82	4.18±3.08	1.94±1.26	0.00
Betaproteobacteria	4.70±2.28	2.33±1.77	1.31±0.55	0.00
Flavobacteria	2.73±2.34	0.48±0.36	0.29±0.18	0.00
Gammaproteobacteria	46.66±8.72	55.93±23.01	79.38±8.89	0.00
Negativicutes	1.13±0.72	1.87±1.30	0.71±1.43	0.05

At order level (Table 30), Actinomycetales, Bacillales, Bacteroidales, Burkholderiales, Flavobacteriales and Lactobacillales were significantly higher on the carcasses of chickens belonging to Group 1 compared to both Groups 2 and 3, while Aeromonadales, Alteromonadales and Pseudomonadales were significantly higher in Group 3 compared to Groups 1 and 2. Within Group

3, Selenomonadales as well as Xanthomonadales were significantly lower as compared to the other tested groups.

Table 30: Orders with mean values of the relative frequency of abundance (%) ≥ 1 in at least one group of carcasses and significantly different in the tested groups ($p \leq 0.05$).

Order	mean±SD			p-value
	Group 1	Group 2	Group 3	
Actinomycetales	3.73±1.74	2.42±2.01	1.42±1.05	0.00
Aeromonadales	1.25±0.68	1.55±1.31	5.90±4.95	0.00
Alteromonadales	0.65±0.28	0.58±0.29	1.04±0.45	0.00
Bacillales	4.76±4.82	2.09±0.91	1.36±0.66	0.01
Bacteroidales	7.06±2.82	4.18±3.08	1.94±1.26	0.00
Burkholderiales	3.77±1.94	1.91±1.55	0.86±0.43	0.00
Flavobacteriales	2.67±2.27	0.47±0.35	0.29±0.17	0.00
Lactobacillales	9.21±4.04	7.23±4.45	3.16±4.18	0.00
Pseudomonadales	15.77±10.61	5.52±3.79	24.27±22.59	0.01
Selenomonadales	1.13±0.72	1.87±1.30	0.71±1.43	0.05
Xanthomonadales	1.63±0.71	1.30±1.38	0.46±0.65	0.01

At family level (Table 31), Bacteroidaceae, Flavobacteriaceae, Comamonadaceae, Enterococcaceae and Staphylococcaceae were significantly higher in Group 1 than in Groups 2 and 3. On the contrary, Moraxellaceae and Pseudomonadaceae were significantly lower in Group 2 in comparison to the other two groups. The family Aeromonadaceae was significantly higher on the carcass belonging to Group 3 compared to Groups 1 and 2, while Lactobacillaceae, Xanthomonadaceae and Lachnospiraceae were significantly lower in Group 3 than in Groups 1 and 2.

Table 31: Families with mean values of the relative frequency of abundance (%) ≥ 1 in at least one group of carcasses and significantly different in the tested groups (p≤0.05).

Family	mean±SD			p-value
	Group 1	Group 2	Group 3	
Moraxellaceae	14.05±9.68	5.13±3.59	22.30±21.58	0.01
Lactobacillaceae	6.49±3.42	5.39±3.75	2.01±2.47	0.00
Bacteroidaceae	6.06±2.51	3.24±2.68	1.54±1.05	0.00
Bacillaceae	3.15±4.34	1.04±0.50	0.61±0.36	0.03
Flavobacteriaceae	2.49±2.14	0.46±0.34	0.28±0.17	0.00
Comamonadaceae	2.15±1.29	0.63±0.79	0.23±0.10	0.00
Pseudomonadaceae	1.72±1.04	0.39±0.25	1.97±1.63	0.00
Xanthomonadaceae	1.63±0.71	1.30±1.38	0.46±0.65	0.01
Enterococcaceae	1.58±0.85	0.67±0.39	0.58±1.15	0.01
Lachnospiraceae	1.27±0.52	1.72±1.64	0.67±0.46	0.04
Aeromonadaceae	1.25±0.68	1.55±1.31	5.90±4.95	0.00
Staphylococcaceae	1.03±0.66	0.60±0.45	0.49±0.39	0.03

In relation to the bacterial genus with abundances ≥ 0.5% in at least one tested group and significantly different among the tested groups (Table 32), the mean relative frequency of abundance of *Bacillus*, *Bacteroides, Enterococcus, Flavobacterium, Lysinibacillus,* and *Staphylococcus* were significantly higher on carcass belonging to Group 1 compared to both Groups 2 and 3, while *Pseudomonas* showed significantly lower abundance on carcass of broilers belonging to Group 2 in comparison to Groups 1 and 3. The mean relative frequency of abundance of *Acinetobacter, Aeromonas*, *Psychrobacter* and *Shewanella* were significantly higher in Group 3 than in Groups 1 and 2, while *Chlorobium*, *Lactobacillus* and *Xanthomonas* were significantly lower in Group 3 compared to Groups 1 and 2.

Table 32: Genera with mean values of the relative frequency of abundance (%) ≥ 0.5 in at least one group of carcasses and significantly different in the tested groups (p≤0.05).

Genus	mean±SD			p-value
	Group 1	Group 2	Group 3	
Acidovorax	0.76±0.52	0.24±0.31	0.06±0.03	0.00
Acinetobacter	12.61±9.32	3.53±2.58	15.17±18.24	0.04

Aeromonas	1.22±0.68	1.52±1.31	5.79±4.91	0.00
Bacillus	1.75±1.71	0.84±0.40	0.48±0.32	0.01
Bacteroides	6.06±2.51	3.24±2.68	1.54±1.05	0.00
Chlorobium	0.62±0.41	0.81±0.93	0.17±0.30	0.03
Enterococcus	1.57±0.85	0.67±0.39	0.58±1.15	0.01
Flavobacterium	0.62±0.51	0.08±0.06	0.07±0.06	0.00
Lactobacillus	6.36±3.32	5.34±3.73	1.97±2.41	0.00
Pseudomonas	1.66±1.02	0.37±0.23	1.93±1.63	0.00
Psychrobacter	1.12±0.91	1.36±1.63	6.39±5.70	0.00
Shewanella	0.41±0.27	0.35±0.22	0.64±0.31	0.02
Staphylococcus	0.98±0.65	0.58±0.44	0.45±0.38	0.03
Xanthomonas	1.20±0.60	0.97±0.99	0.37±0.57	0.02

In relation to the bacterial species with abundances ≥0.5% in at least one tested group and significantly different among the tested groups (Table 33), the mean relative frequency of abundance of *Bacteroides dorei*, *Bacteroides uniformis*, *Bacteroides vulgatus*, *Enterococcus faecalis*, *Enterococcus faecium*, *Lactobacillus reuteri* and *Lactobacillus salivarius* were significantly higher in Group 1 in comparison to Groups 2 and 3. The species *Chlorobium phaeobacteroides*, *Lactobacillus acidophilus* and *Lactobacillus crispatus* showed a significantly higher relative frequency of abundance in Group 2 rather than in Groups 1 and 3. On the contrary, *Acinetobacter* sp. DR1 showed a significantly lower relative frequency of abundance in Group 2 compared to the other tested groups. The species *Acinetobacter johnsonii*, *Aeromonas hydrophila*, *Aeromonas salmonicida*, *Bacteroides fragilis*, *Pseudomonas fluorescens*, and *Psychrobacter* sp. PRwf-1 showed a significantly higher relative frequency of abundance in Group 3 compared to both Groups 1 and 2.

Table 33: Species with mean values of the relative frequency of abundance (%) ≥ 0.5 in at least one group of carcasses and significantly different in the tested groups (p≤0.05).

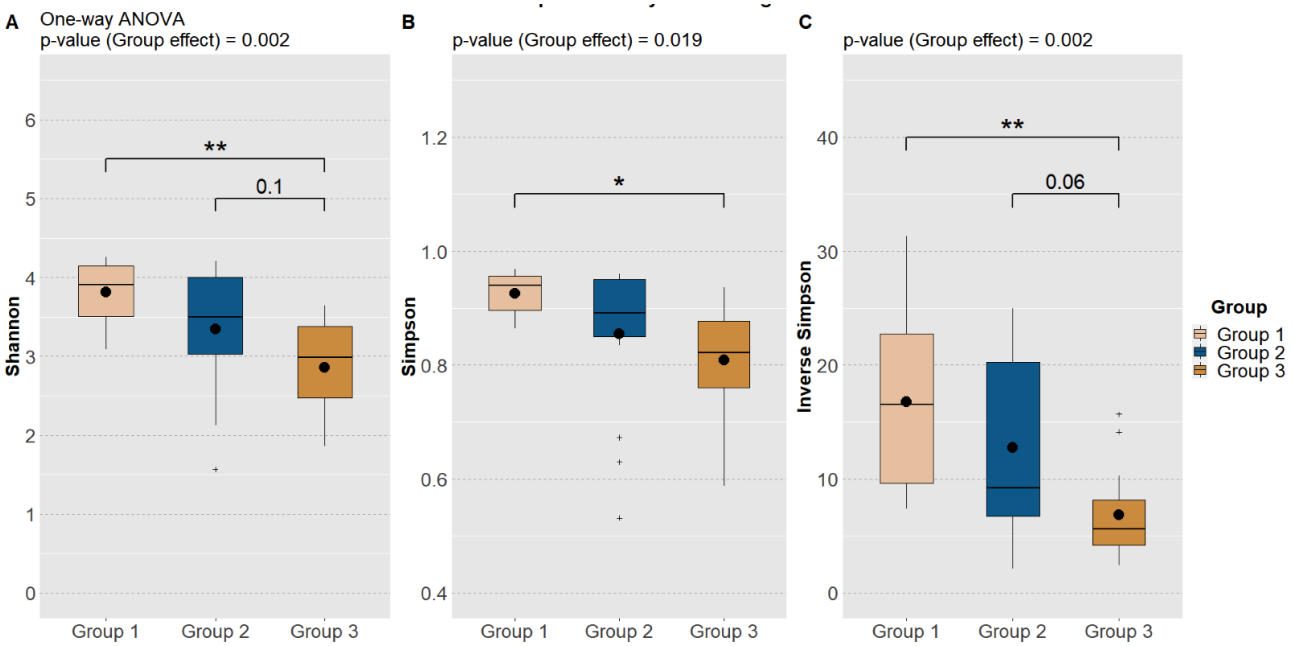
Specie	**mean±SD**			**p-value**
	Group 1	**Group 2**	**Group 3**	
Acinetobacter johnsonii	2.57±1.95	0.77±0.57	5.12±6.14	0.02
Acinetobacter sp. DR1	0.85±0.71	0.18±0.14	0.81±0.98	0.03
Aeromonas hydrophila	0.70±0.49	0.81±0.68	3.09±2.63	0.00
Aeromonas salmonicida	0.51±0.21	0.70±0.63	2.61±2.31	0.00

Bacteroides dorei	0.74±0.50	0.12±0.14	0.08±0.06	0.00
Bacteroides fragilis	0.59±0.23	0.50±0.39	0.28±0.20	0.03
Bacteroides uniformis	0.51±0.20	0.13±0.19	0.02±0.02	0.00
Bacteroides vulgatus	0.91±0.60	0.24±0.25	0.14±0.11	0.00
Chlorobium phaeobacteroides	0.59±0.40	0.79±0.93	0.16±0.30	0.03
Enterococcus faecalis	0.67±0.51	0.25±0.14	0.23±0.40	0.01
Enterococcus faecium	0.51±0.25	0.27±0.19	0.18±0.34	0.01
Lactobacillus acidophilus	0.47±0.15	0.87±0.72	0.21±0.19	0.00
Lactobacillus crispatus	0.83±0.36	1.79±1.51	0.29±0.25	0.00
Lactobacillus reuteri	0.63±0.30	0.29±0.29	0.15±0.24	0.00
Lactobacillus salivarius	2.29±2.30	0.81±1.18	0.52±1.34	0.03
Pseudomonas fluorescens	0.66±0.54	0.09±0.06	1.16±1.03	0.00
Psychrobacter sp. PRwf-1	0.39±0.20	0.21±0.21	3.94±3.12	0.00

Alpha and beta diversity calculated for the bacteria identified on the tested carcasses

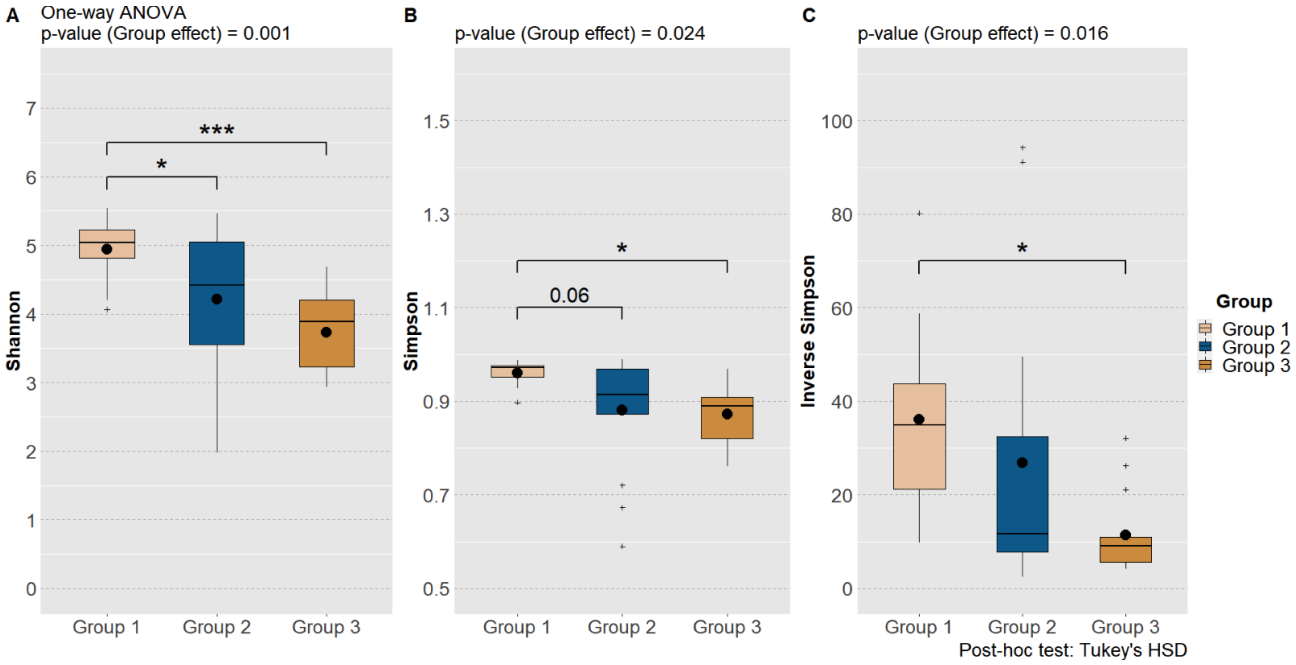
The alpha diversity calculated for the genera identified on the carcasses of broilers tested in each group at the end of the refrigeration tunnel was calculated with the Shannon, Simpson and Inverse Simpson indexes (Figure 19). The results demonstrated that the genera belonging to Group 1 were significantly different in comparison to those colonising carcasses of group 3 using all alpha diversity indexes, i.e., Shannon (Figure 19 A), Simpson (Figure 19 B) and Inverse Simpson (Figure 19 C). On the contrary, the Shannon and Inverse Simpson indexes calculated for the genera colonising the carcasses belonging to Group 2 were not significantly different than those calculated for the genera colonising the carcasses belonging to Group 3 (Figure 19 A and C).

Figure 19: Box plot of the Shannon (A), Simpson (B) and Inverse Simpson (C) indexes calculated for the genera identified on the carcasses belonging to Groups 1-3.



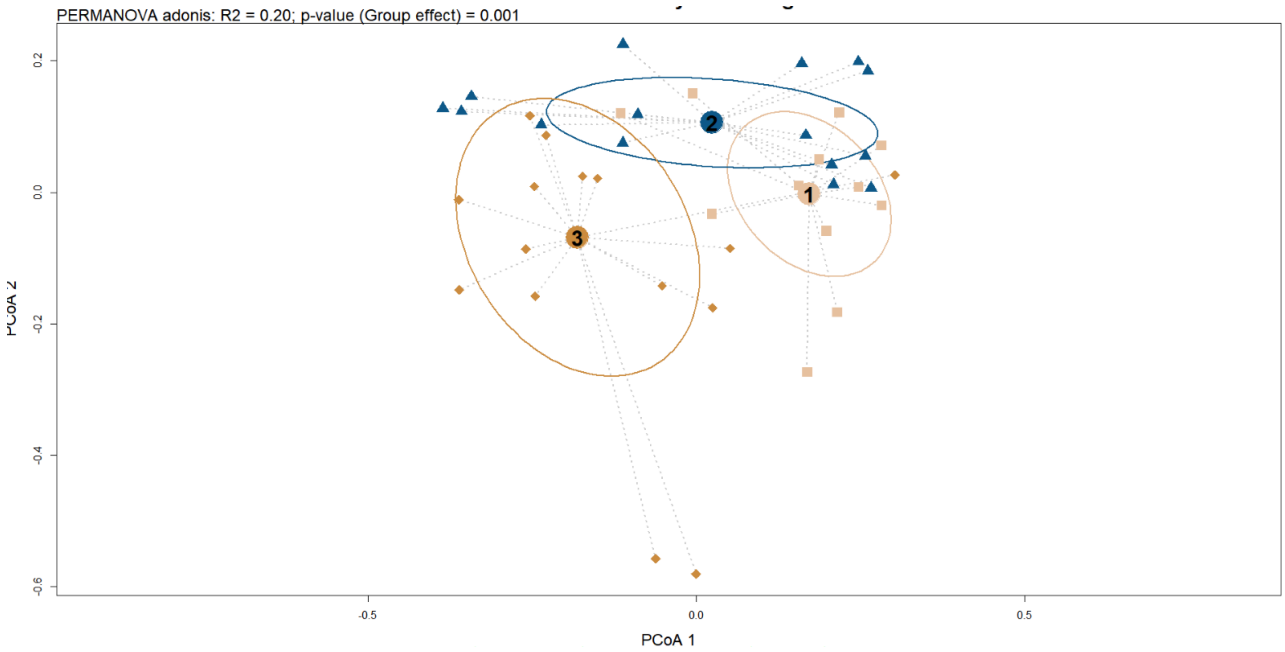
The alpha diversity calculated for the species identified on the carcasses of broilers tested was calculated with the same indexes as above (Figure 20). Overall, the alpha diversity indexes confirm that the species associated to the carcasses belonging to Group 1 were significantly different in comparison to those colonising the carcasses of Group 3 in all indexes, i.e., Shannon (Figure 20 A), Simpson (Figure 20 B) and Inverse Simpson (Figure 20 C). Furthermore, the alpha diversity associated to the species colonising the carcasses belonging to Group 1 calculated with the Shannon index was significantly different in comparison to that calculated for the species identified on carcasses of Group 2 (Figure 20 A). Moreover, the alpha diversity associated to the species on carcasses belonging to Group 1 calculated with the Simpson index did not significantly different than those of species associated to carcasses of Group 2 (p=0.06) (Figure 20 B).

Figure 20: Box plot of the Shannon (A), Simpson (B), and Inverse Simpson (C) indexes calculated for the genera identified on the carcasses belonging to Groups 1-3.



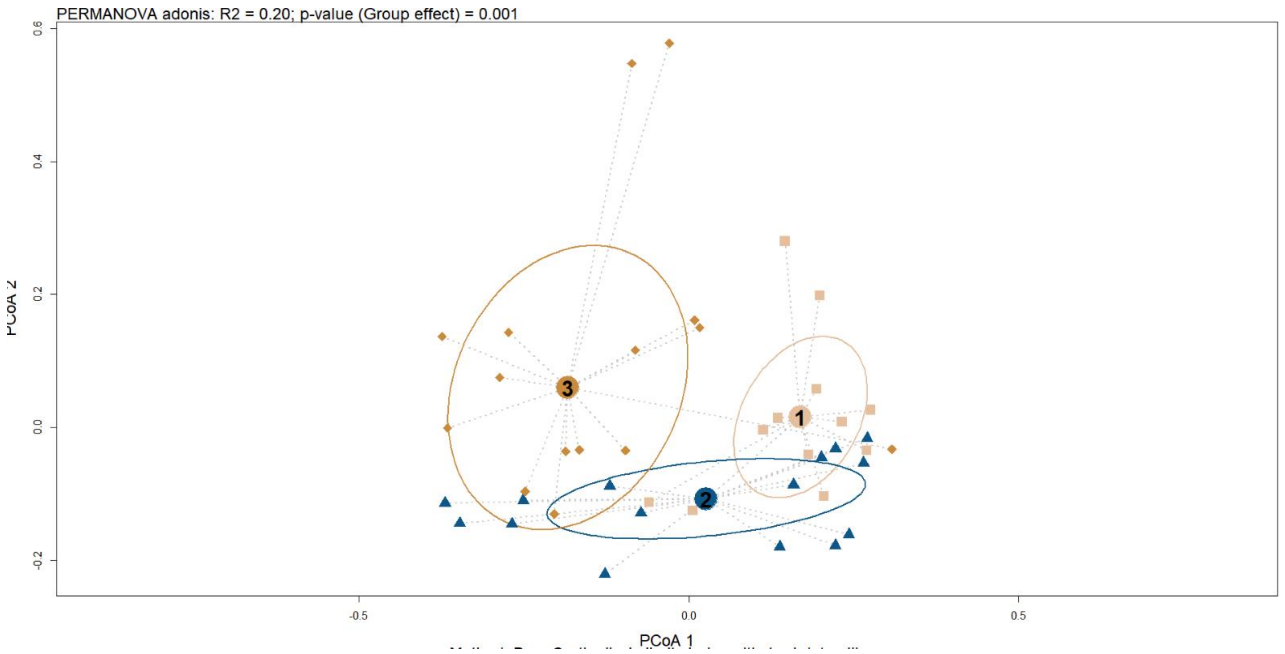
The PCoA analysis of the beta diversity calculated at genus level with the Bray-Curtis distance matrix was used to investigate how genera colonising the carcasses belonging to the different groups cluster one with the other. The results (Figure 21) showed a clear and significant separation between genera colonising carcasses of Group 1 compared to genera colonising carcasses of Groups 2 and 3 (R2=0.20, p = 0.001).

Figure 21: PCoA analysis of the beta diversity calculated for the genera identified on the carcasses belonging to Groups 1-3.



The PCoA analysis of the beta diversity was also calculated at species level using the same approach and the results overlapped with those obtained at genus level (Figure 22). Overall, the beta diversity index confirmed that the species colonising the carcasses belonging to Group 1 were significantly different in comparison to those colonising carcasses of Groups 2 and 3 (R2=0.20, p=0.001).

Figure 22: PCoA analysis of the beta diversity calculated for the species identified on the carcasses belonging to Groups 1-3.



Functional genes identified on the carcasses

In addition to the taxonomic composition, shotgun metagenome sequencing was performed on carcasses of Groups 1 to 3 to identify the functional genes categories analysed up to the functional level. Significantly different abundances were associated to the functional genes identified on the carcasses belonging to all tested groups (Table 34). Genes coding for cytochrome c oxidase polypeptide III (EC 1.9.3.1), dihydropyrimidinase (EC 3.5.2.2), NADH dehydrogenase subunit 1, phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1), RNA helicase-putative and transcription factor S were significantly higher on carcasses of Group 1 in comparison to carcasses of Groups 2 and 3, while gene coding for GTP-binding protein was significantly lower abundant on carcasses of Group 3 in comparison to those of Groups 1 and 2.

Table 34: Genes belonging to the function level showing mean relative frequency of abundance (%) ≥0.5 in at least one group of carcasses and significantly different in the tested groups (p≤0.05).

Function	mean±SD			p-value
	Group 1	Group 2	Group 3	
Cytochrome c oxidase polypeptide III (EC 1.9.3.1)	0.85±0.67	0.46±0.26	0.22±0.30	0.00
Dihydropyrimidinase (EC 3.5.2.2)	0.64±0.25	0.50±0.24	0.21±0.21	0.00
GTP-binding protein	1.18±0.33	1.14±0.43	0.56±0.37	0.00
NADH dehydrogenase subunit 1	0.73±0.47	0.42±0.30	0.19±0.26	0.00
Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)	0.51±0.25	0.34±0.14	0.19±0.18	0.00
RNA helicase. putative	0.67±0.46	0.46±0.20	0.19±0.17	0.00
Transcription factor S	0.60±0.34	0.44±0.27	0.19±0.19	0.00

Antibiotic Resistance Genes on carcasses

ARGs associated to the tested carcasses were retrieved from the functional genes based on their classification in the level 2 of the biological-function ontology in SEED category (Aziz *et al*., 2008). Overall, at level 3 (Table 35) the genes coding for cobalt-zinc-cadmium resistance showed significantly lower abundance on the carcasses of Group 2 in comparison to those of Groups 1 and 3. On the contrary, the genes coding for resistance to fluoroquinolones were significantly higher on carcasses of Group 3 in comparison to carcasses of Groups 1 and 2. Finally, genes coding for multiple antibiotic resistance (MAR) locus, the mdtABCD multidrug resistance cluster and zinc resistance

were significantly lower abundant on carcasses of Group 1 in comparison to those belonging to Groups 2 and 3.

Table 35: Antibiotic resistance genes belonging to the level 3 SEED category showing means of relative frequency of abundance (%) ≥0.01 in at least one group of carcasses and significantly different in the tested groups ($p \leq 0.05$).

Level 3	mean±SD			p-value
	Group 1	Group 2	Group 3	
Cobalt-zinc-cadmium_resistance	0.21±0.12	0.14±0.07	0.29±0.18	0.02
Multiple_Antibiotic_Resistance_MAR_locus	0.00±0.00	0.01±0.01	0.01±0.01	0.02
Resistance_to_fluoroquinolones	0.12±0.08	0.14±0.07	0.21±0.12	0.03
The_mdtABCD_multidrug_resistance_cluster	0.03±0.02	0.07±0.05	0.10±0.08	0.02
Zinc_resistance	0.01±0.01	0.03±0.03	0.04±0.03	0.04

At functional level (Table 36), the genes coding for RND efflux system-inner membrane transporter CmeB showed significantly lower abundance on carcasses of Group 1 in comparison to carcasses of Groups 2 and 3, while vesicular neurotransmitter transporter showed significantly higher abundance on carcasses of Group 1 compared to the other tested groups. Lastly, the genes coding for DNA gyrase subunit B (EC 5.99.1.3), macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-), macrolide specific efflux protein MacA and membrane fusion protein of RND family multidrug efflux pump showed significantly higher abundance on carcasses of Group 3 in comparison to those belonging to Groups 1 and 2.

Table 36: Antibiotic resistance genes belonging to the level function category showing means of relative frequency of abundance (%) ≥0.02 in at least one group of carcasses and significantly different in the tested groups (p≤0.05).

Function	**mean±SD**			**p-value**
	Group 1	**Group 2**	**Group 3**	
DNA gyrase subunit B (EC 5.99.1.3)	0.07±0.05	0.08±0.05	0.14±0.11	0.05
Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-)	0.02±0.02	0.02±0.01	0.08±0.06	0.00
Macrolide-specific efflux protein MacA	0.01±0.01	0.01±0.01	0.03±0.02	0.00
Membrane fusion protein of RND family multidrug efflux pump	0.00±0.01	0.01±0.01	0.03±0.04	0.05
RND efflux system-inner membrane transporter CmeB	0.06±0.03	0.11±0.08	0.18±0.09	0.00
Vesicular neurotransmitter transporter	0.32±0.23	0.20±0.13	0.08±0.07	0.00

4.4 Study 4

In the study 4, three aliquots of a salame homogenate were tested as part of the activities scheduled in the PRIMA project named ArtiSaneFood to verify if a single aliquot of homogenate is representative of the whole sample. The output of each metagenome is shown in Table 37. In the Baird Parker plates obtained from the aliquots of homogenate, coagulase-positive staphylococci were not detected. The metagenomes were analysed using MG-RAST and the database RefSeq.

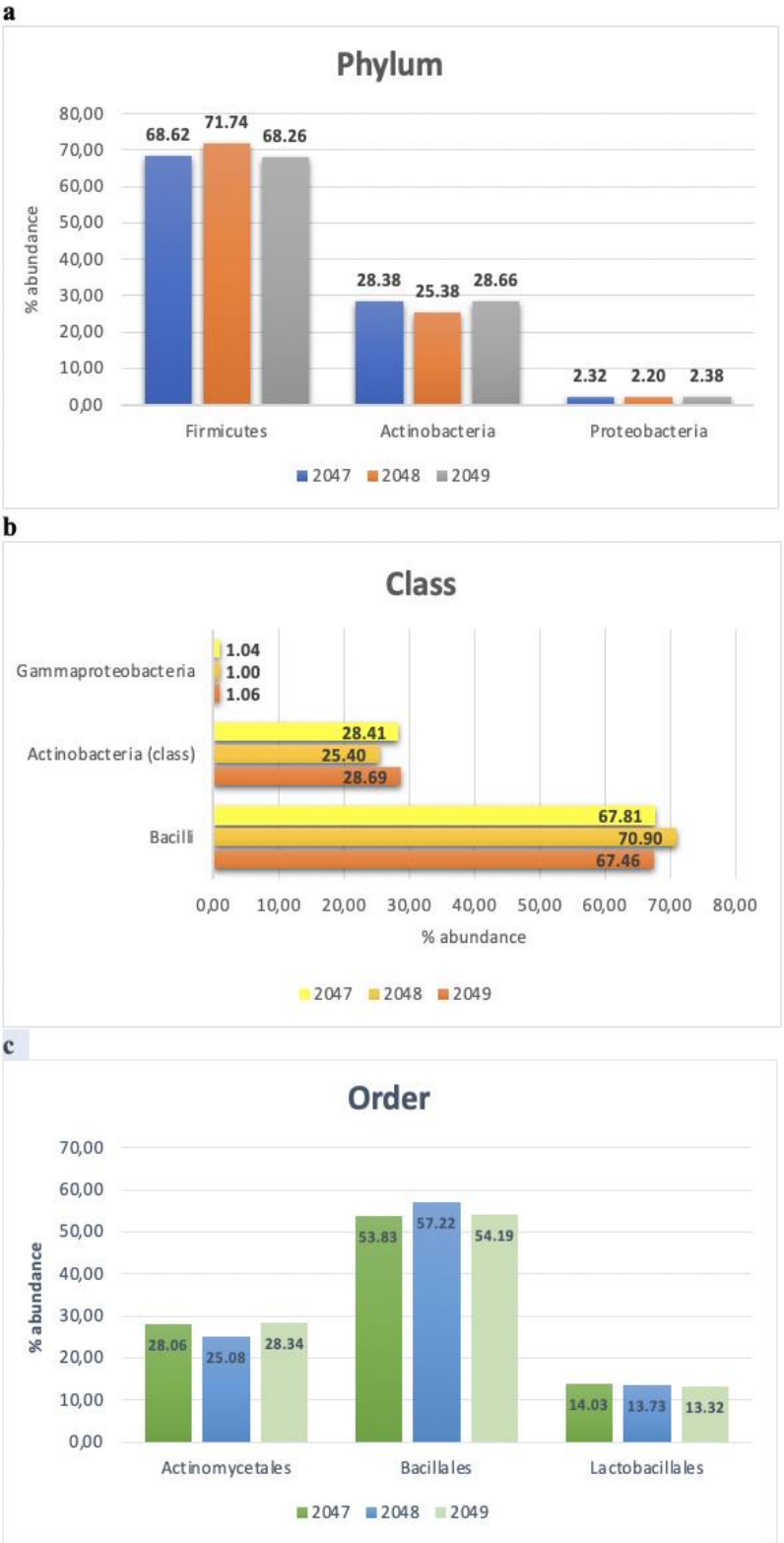
Table 37: Labels and outputs of the tested metagenomes

Food matrix	**Aliquot label**	**Metagenome ID**	**N. of reads**	**Gbp**
Salame	Aliquot 1	2047	21537193	3.07
	Aliquot 2	2048	21733486	3.00
	Aliquot 3	2049	39474471	7.28

Taxonomic groups identified in the investigated metagenomes

At phylum level, in all samples, Firmicutes was the dominant taxonomic group followed by Actinobacteria and Proteobacteria (Figure 23a). At class level, Bacilli was the most represented class followed by Actinobacteria and Gammaproteobacteria (Figure 23a). Finally, at order level, Bacillales was dominant followed by Actinomycetales and Lactobacillales (Figure 23c).

Figure 23: Relative abundance (%) of the dominant bacterial phyla (a), classes (b) and orders (c) identified in the salame samples. Only taxa with a relative abundance >1% in at least one sample are shown.



At family level Staphylococcaceae, Brevibacteriaceae and Lactobacillaceae were the most abundant (Table 38) and this result was confirmed at genus level (Table 39).

Table 38: Bacteria families with a relative abundance (%) ≥1 in at least one sample identified in the salame samples.

Family	**Relative abundance**		
	Metagenome ID 2047	**Metagenome ID 2048**	**Metagenome ID 2049**
Bacillaceae	2.92	3.04	2.92
Brevibacteriaceae	14.14	12.61	14.41
Corynebacteriaceae	1.15	1.04	1.16
Lactobacillaceae	9.54	9.21	8.75
Leuconostocaceae	2.80	2.85	2.97
Micrococcaceae	3.99	3.60	3.91
Nocardiaceae	0.99	0.89	1.02
Staphylococcaceae	50.29	53.53	50.65
Streptomycetaceae	1.21	1.07	1.22

Table 39: Bacteria genera with a relative abundance (%) ≥0.5 in at least one sample identified in the salame samples.

Genus	**Relative abundance**		
	Metagenome ID 2047	**Metagenome ID 2048**	**Metagenome ID 2049**
Arthrobacter	2.27	2.05	2.26
Bacillus	2.29	2.39	2.29
Brevibacterium	14.10	12.58	14.37
Corynebacterium	1.15	1.04	1.15
Enterococcus	0.93	0.91	0.86
Kocuria	0.57	0.51	0.54
Lactobacillus	9.04	8.73	8.29
Leuconostoc	2.60	2.65	2.78
Mycobacterium	0.87	0.77	0.87
Rhodococcus	0.80	0.72	0.82
Rothia	0.54	0.49	0.50
Staphylococcus	49.95	53.17	50.31

Streptomyces	1.21	1.07	1.22

At species level, *Staphylococcus saprophyticus* was the most abundant species among the species of the phylum Firmicutes (Table 40), followed by *Staphylococcus epidermidis* and *Staphylococcus aureus*. Nevertheless, *Brevibacterium linens* was the most abundant species among those within the phylum Actinobacteria (Table 40).

Table 40: Bacteria species with a relative abundance (%) ≥0.5 in at least one sample identified in the salame samples.

Species	**Relative Abundance**		
	Metagenome ID 2047	**Metagenome ID 2048**	**Metagenome ID 2049**
Arthrobacter aurescens	0.79	0.71	0.78
Arthrobacter sp. FB24	0.81	0.73	0.79
Brevibacterium linens	11.98	10.69	12.22
Brevibacterium mcbrellneri	2.09	1.86	2.13
Kocuria rhizophila	0.57	0.51	0.54
Lactobacillus casei	0.79	0.76	0.76
Lactobacillus paracasei	0.68	0.65	0.64
Lactobacillus plantarum	0.79	0.76	0.72
Lactobacillus sakei	4.66	4.49	4.24
Leuconostoc kimchi	0.61	0.63	0.66
Leuconostoc mesenteroides	1.17	1.19	1.25
Staphylococcus aureus	6.24	6.66	6.31
Staphylococcus capitis	1.50	1.56	1.49
Staphylococcus caprae	1.19	1.27	1.21
Staphylococcus carnosus	1.53	1.62	1.53
Staphylococcus epidermidis	6.63	7.07	6.68
Staphylococcus haemolyticus	1.94	2.05	1.94
Staphylococcus hominis	2.57	2.74	2.59
Staphylococcus lugdunensis	2.75	2.90	2.75
Staphylococcus saprophyticus	22.14	23.59	22.28
Staphylococcus warneri	3.15	3.36	3.18

Functional genes identified in salame samples

The only functional gene identified in salame samples with a relative abundance > 0.5% (Table 41) in at least one sample was represented by Aldehyde dehydrogenase (EC 1.2.1.3) involved in a wide range of metabolic pathways including Glycolysis / Gluconeogenesis, Ascorbate and aldarate metabolism, Fatty acid degradation, Valine, leucine and isoleucine degradation, Pyruvate metabolism, etc.

Table 41: Functional gene with a relative abundance (%) ≥0.5 in at least one sample identified in the salame samples.

Level 1	**Function**	**Relative abundance**		
		M ID 2047	**M ID 2048**	**M ID 2049**
Carbohydrates	Aldehyde dehydrogenase (EC 1.2.1.3)	0.53	0.51	0.53

5. Discussion

Metagenomic refers to a non-culture based approach providing insight into the overall functional repertoire of a microbial community, including information on its metabolic capabilities and the potential functional interactions among its members. Metagenomics has become an emerging field of microbial ecology since the first use of the concept in 1998 in relation to soil microorganisms (Handelsman *et al*., 1998). Originally, metagenomics studies depended on the cultivation of clonal cultures, followed by functional expression screening (Handelsman *et al*., 1998). However, such cultures are not able to represent the total community profile and may overlook the vast majority of the microbial biodiversity. Thus, natural microbial communities typically contain a wide diversity of organisms, viruses and other chromosomal and extra-chromosomal genetic elements. The advent of next-generation sequencing (NGS) technologies has made it possible to perform sequencing-based metagenomic analyses. The **main objective** of this research project has been to apply shotgun metagenomic sequencing to investigate both microbiome and resistome of foods of animal origin in order to assess advantages and disadvantages of shotgun metagenomic sequencing in comparison to the ISO cultural methods used to verify the compliance of food lots to the microbiological criteria.

One of the main advantages of shotgun metagenomics is the possibility to detect in the same analysis microorganisms belonging to different domains and to assess those microorganisms as part of a whole ecosystem. The analysis of such ecosystem can provide insight to understand why a pathogen can survive in a food system and eventually increase up to concentrations able to cause human diseases. A further advantage of metagenomics is that beside the taxonomic composition of a food sample it provides a full characterisation of the functional genes characterising a food system including antibiotic resistance genes. The characterisation of the resistome associated to food systems is of increasing interest in food safety in order to map all possible sources of antibiotic resistance which are often misinterpreted and underassessed.

The main disadvantage of shotgun metagenomics is that there is no clue on the relationship between number of reads in a food sample and number of cfu and when reads for specific microorganisms are detected is always challenging to assess if they belong to live or dead cells. Moreover, there is no consensus on the threshold of sequencing depth which should be applied to collect representative data. Ni *et al*., (2013) suggested to consider that prokaryotic genomes have 1 to 15 small subunit (SSU) ribosomal (r) RNA gene copies, that range from 139 kb to 13.034 kb. The diversity of both SSU rDNA copies and prokaryotic genome sizes could significantly disturb the accurate estimation of the depth for metagenomic sequencing. Retrieving the 2339 SSU rDNA sequence of three human faecal specimens from the study by Eckburg *et al*., (2005), Ni *et al*., (2013) suggested a computational

approach demonstrating that the estimated amounts for sequencing specimens were 7.00 Gb at the species level, 6.93 Gb at the genus level, 7.10 Gb at the family level and 6.54 Gb at the order level. These results imply that at least 7 Gb is required for sequencing to enumerate the gene contents of prokaryotes with relative abundance of more than 1% in the human faecal microbiome (Ni *et al*., 2013). This sequencing depth results in a sequencing cost around 300 Euro/sample. This high cost is likely to limit the implementation of shotgun metagenomics outside research projects.

When metagenomics results are available, the next step is data analysis and results of study 1 in relation to **SO1** demonstrated that all bioinformatic tools tested were able to detect the bacteria experimentally spiked in the salmon with mean relative values of abundance > 0.1% which, according to Petersen *et al*., (2017), can be considered a threshold to consider a taxonomic classification as a reliable one. The parasite *C. parvum* has been detected by MG-RAST and CosmosID, while the *Bovine alphaherpesvirus 1* CosmosID and OneCodex. Finally, the fungus *S. cerevisiae* has not been detected displaying a relative abundance always <0.1%. In relation to **SO2**, among the tested bioinformatic tools, OneCodex and CosmosID are the most user friendly in terms of sequence upload and data interpretation. The CosmosID databases are organized phylogenetically and contain hundreds of millions of marker gene sequences. The markers represent both coding and non-coding sequences uniquely identified by taxon and/or distinct nodes of phylogenetic trees. This means that the tree structure was created based on genomic relatedness of organisms rather than predetermined taxonomy based on phenotype. This allows CosmosID to have a high degree of accuracy in identifying microorganisms based on their DNA in metagenomic samples. It also helps identify the closest match to genomes that do not have strain level references in the database (if, for example, they have never been sequenced before). However, as far as quantification results are concern, the high percentage abundances detected using CosmosID for the microorganisms of the mock community are due to the fact that the abundance analysis is done for each domain separately. Therefore, an abundance of 88.74% for *C. parvum* it does not mean that the parasite reads represent the majority of the reads of the metagenome, but it represents the majority of the reads assigned to eukaryotes. One Codex identifies microbial sequences using a "k-mer based" taxonomic classification algorithm as CosmosID and MG-RAST, but it is built on a web-based data platform, using a reference database that currently includes approximately 40,000 bacterial, viral, fungal, and protozoan genomes. Quantitative evaluation of several published microbial detection methods shows that One Codex has the highest degree of sensitivity and specificity (AUC = 0.97, compared to 0.82-0.88 for other methods), both when detecting well-characterized species as well as newly sequenced, "taxonomically novel" organisms (Minot *et al*., 2015). Besides the facility of use and also speed of analysis of both CosmosID and OneCodex, MG-RAST includes data analysis options not available

for the other software. Moreover, in study 1 MG-RAST was able to detect *Saccharomyces cerevisiae* although the DNA virus was neither detected nor quantified.

Using MG-RAST the RefSeq provided the best results. The NCBI's Reference Sequence (RefSeq) collection is a freely accessible database of naturally occurring DNA, RNA, and protein sequences. It is a unique resource because it provides a large, multi-species, curated sequence database representing separate but explicitly linked records from genomes to transcripts and translation products (Pruitt *et al*., 2012). Unlike the sequence redundancy found in the public sequence repositories, the RefSeq collection aims to provide, for each included species, a complete set of non-redundant, extensively cross-linked, and richly annotated nucleic acid and protein records (Pruitt *et al*., 2012). Even though current computational analysis strategies for metagenomic data rely largely on comparisons to reference genomes, they represent only a fraction of what we know and therefore limit our ability to segregate metagenomic data into coherent biological entities and fail to describe previously unknown species, phages and modules of genetic variation within microbial species (Nielsen *et al*., 2014). A possible alternative is the *de novo* assembly (i.e., assembly without a reference) of genomes from complex metagenomic data, although it is inherently difficult due to many sequence ambiguities that confuse the assembly process. Hence, a typical metagenomic assembly will result in a large set of independent contigs that are not easily aggregated into biological entities. Yang *et al*., (2016) acknowledge that given appropriate sequencing depth, shotgun metagenomics has great utility for investigating the ecology of food-borne pathogens. Nevertheless, it cannot currently be used for identification and quantification of pathogens for regulatory purposes due to limitations of the available technology and the incompleteness of bacterial genome databases. Specifically, the misclassification, that is inherent to the read length, the inability to get deep coverage of the pathogenic organisms in the sample due to the existence of other prokaryote and eukaryote DNA within the sample, and the impossibility of obtaining a comprehensive database containing all possible pathogenic organisms of interest invalidates the use of this approach for regulatory purposes. All in all, in relation to **SO2**, the results demonstrate that MG-RAST with the databases RefSeq, OneCodex and CosmosID can be used as data analysis tools to detect microorganisms belonging to different domains experimentally spiked in smoked salmon analysed by shotgun metagenomics sequencing. Nevertheless, a direct correlation between cell concentration of each spiked microorganism and number of corresponding reads is still not possible, although bacteria were identified with higher abundances than *C. parvum, S. cerevisiae* and *Bovine alphaherpesvirus*. The number of reads selected as detection cut-off level must be clearly defined to use shotgun metagenomic sequencing in food microbiome studies, and in food safety risk assessment. There are many valuable papers on the application of shotgun metagenomics, but the lack of transparent

information on the technical details of both the wet-lab and bioinformatic procedures are delaying the full implementation of this powerful sequencing approach (Sala *et al*., 2020).

In relation to studies 2 and 3 they were planned because nowadays conventionally raised poultry continues to dominate the EU poultry industry. However, there is an increasing consumer demand for meat obtained in antibiotic free rearing cycles. Moreover, in January 2022, the new EU Regulations 2019/6 and 2019/4 will enter in force, further limiting the use of veterinary medical products and medicated feed in animal productions. Therefore, I decided to investigate whether the efforts of raising chickens without the use antibiotics would make any difference in the microbiome of poultry meat eaten by consumers (**SO3**). The results of study 2 demonstrated three key findings. The first one is a clear separation between the taxonomic, functional and antibiotic resistant genes in the caeca of the birds reared on the conventional and antibiotic free farm. This result is due to the fact that each poultry farm has an associated ecosystem due to the geographical and specific environmental conditions, to what chickens eat and drink, to the litter type, to the workers, and certainly to the medications they receive or not. That separation was completely lost on carcasses belonging to the two groups, which did not mirror whatever positive or negative impact the farm ecosystem and rearing condition had on the chicken caeca. As for the caeca, and also for the carcass microbiomes, there are many contributing factors besides the possible cross contamination during the evisceration. Indeed, the ecosystems interacting with the animals during transport and then during each slaughtering step, including the final refrigeration tunnel, all contribute to the final carcass microbiomes.

The second key finding is that the antibiotic free production resulted in statistically significant lower antimicrobial resistance load in the caeca of chickens in comparison to the conventional production, thus confirming that besides external sources of ARGs, when antimicrobials are not administered to the animals in the caeca of that flock there is a lower antimicrobial resistance load. In relation to the short- and long-term effects of the use of antimicrobials on antimicrobial resistance, Mughini-Gras *et al*., (2021) showed that the antimicrobial use at flock level is more relevant for antimicrobial resistance in *Escherichia coli* than the historical use of antimicrobials at the farm level. Overall, these observations demonstrate that the effort of reducing antimicrobial use by means of rearing antibiotic free flocks should be associated with a better understanding of the antimicrobial resistance persistence in the farm environment in the absence of direct antimicrobial use. Further insights into the antimicrobial resistance persistence in the farm environment might help us to understand why, for instance, the relative abundance of acriflavine resistance protein genes was higher in the caeca of antibiotic free animals compared to conventional ones.

The last main finding is that the antimicrobial resistance load on carcasses was much higher than in the caeca, without any significative difference between carcasses coming from the two types of farming. As described above, this result demonstrated that all post- harvest steps, including transport and slaughtering, but also the loading and unloading of the animals contributes not only to the microbiome colonizing the final carcass reaching the consumers, but also to its antimicrobial resistance load. Therefore, although the most important antimicrobial resistant risk factors and possible mitigation measures are still under investigation at farm level (Davies and Wales, 2019), the implementation of past and future EU regulations aimed at reducing antimicrobial use for food production animals has been ensuring a significant reduction of antimicrobial resistance load at farm level. Therefore, the same effort made for the identification of relevant sources of pathogen and spoilage microorganisms and ARGs should be now devoted to the post-harvest steps (EFSA BIOHAZ Panel, 2021). Little currently available data demonstrate that both transport trucks and cages can contaminate the birds with bacteria and ARGs (Althaus *et al*., 2017) and contribute to the cross contamination between the slaughterhouse and the farms (Buess *et al*., 2019). Moreover, when the animals reach the slaughtering line, scalding, defeathering and evisceration can spread both microorganisms and ARGs from the animal to the environment, although some tentative steps toward reducing these cross contaminations using innovative technologies are in place (Rasschaert *et al*., 2020). Additional sources of both microorganisms and ARGs are workers, equipment, air, process water and wastewater from slaughtering (Savin *et al*., 2020). All these sources together contribute to the carcass microbiome, and our results showed that at the end of the refrigeration tunnel the microbiome of carcasses from animals reared in the conventional farm overlaps with that of carcasses from birds reared in an antibiotic free cycle. My results are consistent with those of Li *et al*., (2020), who investigated chicken breast microbiomes at the retail level, accounting also for the effect of the processing environment and packaging conditions. Their results confirmed that the microbiome of the chicken breast is affected by packaging in air versus under vacuum and by the processing plant where the chicken breast is processed. On the contrary, both the use of antimicrobial at the farm level as well as seasonality affected neither the composition nor the diversity of chicken breast microbiomes in terms of both alpha and beta diversities.

The alpha diversity calculated in this study at genus level using the indicators of richness (Chao1), evenness (InvSimpson) and diversity (Shannon) within the caeca samples show values significantly higher in the caeca of birds reared on the conventional farm in comparison to the antibiotic free farm. The bacteria biodiversity within the GI tract is considered an indicator of good health, and it was expected to be higher in the caeca of chickens not treated with Amoxicillin and Sulfadimethoxine/Trimethoprim. However, these antibiotics are only partially absorbed in the gut

(Anadón *et al*., 1996; Spielmeyer *et al*., 2014) and this might explain why their administration did not reduce the overall bacteria richness in the caeca. Moreover, the intestinal microbiota biodiversity is the result of different factors such as management protocols applied on the farm, animal characteristics and administered diets (Feye *et al*., 2020) which were possibly different in the conventional and antibiotic free farm investigated in this research.

Among the most represented genera detected in the caeca, *Alkaliphilus*, *Desulfibacterium*, *Bacillus* and *Ethanoligenens* were identified as signature genera in the birds from the conventional farm, while butyrate-producing microorganisms as *Coprococcus*, *Roseburia* and *Subdoligranulum* were identified in the caeca of birds reared on the antibiotic free farm. This result highlighted that beside the higher bacteria biodiversity identified in the caeca of birds from the conventional farm, the signature genera colonizing the caeca of the birds reared on the antibiotic free farm belonged to microbial groups supporting animal health. Indeed, butyrate fights against pathogen colonization in poultry (Fernández-Rubio *et al*., 2009) and is involved in several intestinal functions, being an energy source stimulating epithelial cell proliferation and differentiation, other than exerting an antimicrobial effect by promoting the production of peptides and stimulating the production of tight junction proteins (Dalmasso *et al*., 2008).

In this pilot study, we investigated for the first time the microbiome of the caeca of a bird and that of the corresponding carcass. The results showed that the caeca and carcasses of the same flock positively correlate one with the other. However, the correlation between the microbiome of the caeca and the carcass of the same bird was not stronger than that with other caeca and carcasses of the same flock. Therefore, the target analysis of caeca and carcass of the same animal does not provide any added value in comparison to the microbiome analysis at flock level. It is also clear from Figure 8 that the correlation between the functional genes was higher than for bacteria genera, possibly because the same functional gene can be shared between different bacteria genera.

Besides the qualitative and quantitative differences in the most represented functional genes identified in the caeca and on carcasses from the animals reared in the conventional and antibiotic free farm, the most relevant result concerns the antibiotic resistant genes and the total antimicrobial resistance load. In relation to the ARGs, the multi antimicrobial extrusion protein (Na(+)/drug antiporter)-MATE family of MDR efflux pumps was significantly higher in the caeca of birds reared on the conventional farm in comparison to the antibiotic free farm, along with few other ARGs. The MATE gene family is widely distributed in both Gram-positive and Gram-negative bacteria and contributes to the intrinsic, acquired, and phenotypic resistance of bacterial pathogens (Blanco *et al*., 2016). Moreover, it can confer resistance to a specific class of antibiotics or to many drugs, thus conferring a multi-drug resistance (MDR) phenotype to bacteria (Marquez, 2016). In contrast, the abundance of

genes encoding acriflavin resistance protein was significantly higher in the caeca of birds reared in the antibiotic free farm. The acriflavine resistance protein is among multidrug resistance efflux transporter proteins that belongs to the resistance modulation division superfamily (RND), conferring broad spectrum resistance to Gram-negative bacteria (Seeger *et al*., 2006).

For both caeca and carcasses, the overall antimicrobial resistance abundance per drug class did not show significative difference between the birds collected in the two tested farms, while specific differences were observed between drug classes associated to caeca and carcasses. The drug classes identified in both caeca and carcasses largely overlap with those identified by Munk *et al*., (2018) in the faecal resistome investigated in European poultry farms, including Italian farms. In both studies, aminoglycoside, β-lactam, tetracycline and macrolide are widely represented although we identified a larger proportion of β-lactam as well as rifamycin not reported by Munk *et al*., (2018). Our results confirm what was observed by Li *et al*., (2020) in relation to the absence of difference between the resistome associated to chicken breast from birds reared in conventional and antibiotic free farms. On the contrary, the results on the antimicrobial resistance load are the opposite because we calculated a higher antimicrobial resistance load on carcasses while Li *et al*., (2020) discovered a low risk of ARG accumulation on chicken breast. This result is possibly because in the US, poultry carcasses can be disinfected using chlorinated water or organic acids, while in the European Union the use of substances intended to remove microbial surface contamination is only permitted after a full risk analysis taking into account the results of a risk assessment based on the available scientific evidence (EFSA BIOHAZ Panel, 2014).

Genes coding for resistance to vancomycin were identified among the ARGs with normalized mean values of abundance >1000 in at least one tested group. In accordance with other authors (Savin *et al*., 2020; Di Fracesco *et al*., 2021), vancomycin resistance genes can be identified in poultry flocks, although avoparcin has been banned by the EU since 1997. The relative abundance we estimated for the vancomycin resistance genes constitutes a body of evidence of their persistence, while Savin *et al*., (2020) reported a declining trend.

Overall, the results of study 2 and the scientific literature demonstrate that each intervention in whatever processing step that the chicken and poultry meat is at, as with other food productions, is affected by the existing microbiome and resistome shifting and changing from farm to fork. Therefore, building robust, comparable, and representative databases of animal-, farm-, food- and production environment-associated microbiomes and resistome from farm to fork, as it is done for individual foodborne isolates and indicator microorganisms (EFSA and ECDC, 2021), would certainly help to predict the effect of control strategies to reduce food contamination by foodborne pathogens as well as antimicrobial resistance genes in a systemic way.

All in all, the results of study 2 in relation to **SO3** indicate that post-harvest steps withdraw the positive effects of antibiotic free rearing on carcass microbiomes. Therefore, it is crucial to assess the contribution of both transport and slaughter on carcass contamination and spreading of ARGs to identify possible mitigation options addressing consumer concerns on antimicrobial resistance and enhancing the positive impact of the European legislation as well as the economic and management efforts of producers to rear antibiotic free chickens.

In relation to **SO4**, in the study 2 I investigated for the first time the microbiome of the caeca of a bird and that of the corresponding carcass. The results showed that caeca and carcasses of the same flock positive correlate one with the other. However, the correlation between the microbiome of the caeca and the carcass of the same bird was not stronger than that with other caeca and carcasses of the same flock. Therefore, the conclusion is that the target analysis of caeca and carcass of the same animal does not provide any added value in comparison to the microbiome analysis at flock level.

Study 3 was performed as part of a biggest study within the CIRCLES project which has the overall aim to investigate how the microbiomes associated to poultry food system interact and impact one with the other. The carcasses I tested are only part of all the samples which have been collected and investigated in association with metadata associated to the farms and flocks tested. Unfortunately, the overall analysis has not been finalised before the end of my PhD project and few considerations can be done on the tested carcasses as self-standing samples. Those carcasses were all obtained by female broilers reared in farms which in the case of group 1 administered antibiotics for therapeutic treatment during the rearing cycle, while for groups 2 and 3 antibiotics were never administered. As in study 2, in study 3 all animals were slaughtered in the same slaughterhouse but at different days. Moreover, in this case, it was not possible to schedule the slaughtering of the investigated groups as first group of the day. Since in this study both the alpha and beta diversity, calculated at genus and species level, highlighted a significative difference between groups 1 and 3 an impact of the administration of the antibiotic in group 1 might result in an impact on the carcass microbiome. On the other hand, it is also possible that the detected differences are not related to the rearing conditions but to the cross contamination between the transport and slaughterhouse environment. In all tested carcasses, both *Salmonella* and *Campylobacter* were not detected applying both microbiological investigation and metagenomics analysis. In relation to identified functional genes in the three groups tested, cytochrome c oxidase polypeptide III (EC 1.9.3.1), dihydropyrimidinase (EC 3.5.2.2), GTP-binding protein, NADH dehydrogenase subunit 1 and transcription factor S were the same functional genes identified, (but without significative difference) on the carcases in the study 2. In relation to the genes coding for resistance to antibiotics as the macrolide export ATP-binding/permease protein MacB (EC 3.6.3.-) and macrolide-specific efflux protein MacA were significantly higher abundance on carcasses

of Group 3 in comparison to those belonging to Groups 1 and 2 and were also the two antibiotic resistance genes identified on the conventional carcasses of study 2. MacA and MacB together TolC is an ABC type tripartite efflux pump responsible for conferring resistance in Gram negative bacteria (i.e., *Enterobacteriaceae*) to several antibiotics, actively expels macrolide antibiotics or drug, fluroquiolones, penicillins, and solithromycin out of the cell (Fitzpatrick *et al*., 2017). It also regulates the process of colonization and virulence in clinical isolates (Bogomolnaya *et al*., 2013; Sun *et al*., 2014). The macA genes encode for the periplasmic protein that binds the outer membrane protein (TolC), to the inner membrane protein encoded by macB gene (Phan *et al*., 2015; Yewale *et al*., 2020). When all CIRCLES data will be available it will be hopefully clear how metagenomes associated to the three tested poultry farms and investigated flock affect one the other and to which extend microbiome associated to the animals at farm level reflect those on carcasses at the end of the refrigeration tunnel.

In the last part of my PhD project, I contributed to the activities of the PRIMA project ArtiSaneFood. I started to apply shotgun metagenomic to fermented meat and the question I tried to answer was if testing one aliquot of artisanal food homogenate is representative of the whole homogenate (**SO5**). To do so I compared the metagenomes obtained from three aliquots of homogenate obtained diluting 25 g of an artisanal salame in 225 ml of Buffered Peptone Water. The results clearly showed that the metagenomes obtained from three aliquots of the homogenate displayed overlapping taxonomic and functional composition. Therefore, shotgun metagenomics of a single aliquot of an artisanal fermented food is representative of the whole homogenate. The results on the taxonomic genera associated to the tested salame highlighted a very high abundance of the genus *Staphylococcus*, displaying a relative abundance around 50% in all three tested aliquots. Within the group of cocci gram positive catalase positive (GPCP), the Staphylococcus are the predominant genera present in the GPCP community (Talon and Leroy, 2011), contributing to the development and stability of the colour. They prevent the rancidity of salame due to their antioxidant activities and they also enhance the flavour of fermented salame, mainly through amino and fatty acid degradation (Leroy *et al*., 2010; Pisacane *et al*., 2015). In the Baird Parker plates obtained from the aliquots of homogenate, then pelleted before making the extraction of the total DNA submitted to shotgun metagenomics, colonies of *Staphylococcus* were identified but without an opaque zone, due to an egg yolk-lecithinase reaction, around the colonies, generally associated only to coagulase positive *Staphylococcus* as *Staphylococcus aureus*. This result was confirmed by the metagenomic data because the most abundant *Staphylococcus* species identified in the three metagenomes investigated with a relative abundance always >22% was represented by *Staphylococcus saprophyticus*, thus representing a coagulase negative species. *Staphylococcus epidermidis*, also representing a coagulase negative

Staphylococcus was identified as well, with relative abundance between 6.63 and 7.07 in the investigated metagenomes. Both these species might have reduced the possibility of *Staphylococcus aureus* to growth on Baird Parker media because according to the metagenomic results it was detected in the metagenomes investigated with relative abundance between 6.24 and 6.66%.

6. Conclusion

There are many valuable papers on the application of shotgun metagenomics, but the lack of transparent information on the technical details of both the wet-lab and bioinformatic procedures are delaying the full implementation of this powerful sequencing approach in different sectors including food inspection. The investments in research on shotgun metagenomics are justified by the fact that the results achieved in the food system can result in the identification of both spoilage and pathogenic microorganism in the ecosystems where those microorganisms are in real life, and as a matter of fact, the interaction between microorganisms and their ecosystems impacts on both pathogen survival and multiplication ability (Yang *et al*., 2016; Escobar-Zepeda *et al*., 2016). Whenever a sample processing is designed to be as non-specific as possible to capture all nucleic acids regardless of their source, shotgun metagenomics is applicable simultaneously for viruses, bacteria, and parasites (Sala *et al*., 2020; Wylezich *et al*., 2018). Since in each annual EFSA-ECDC report on foodborne and waterborne outbreaks occurring in the EU, there is always a high percentage of outbreaks for which the causative agent is described as ‘unknown’ or ‘unspecified’ and these unknowns are likely to be uncultivable or difficult to culture microorganisms, they could possibly be detected using shotgun sequencing. However, for diagnostic metagenomics to become truly useful, the method must provide robust and reproducible outputs (Andersen *et al*., 2018).

7. References

Alcock, B.P., Raphenya, A.R., Lau, T.T., Tsang, K.K., Bouchard, M., Edalatmand, A., Huynh,W., Nguyen, A.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A., Tran, H.T., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M., Hernandez-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L., Dooley, D., Griffiths, E., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.V., McArthur, A.G. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research*, 48(D1), D517-D525.

Allen, V.M., Corry, J.E., Burton, C.H., Whyte, R.T., Mead, G.C. (2000). Hygiene aspects of modern poultry chilling. *International Journal of Food Microbiology*, 58(1-2), 39-48.

Althaus, D., Zweifel, C., Stephan, R. (2017). Analysis of a poultry slaughter process: Influence of process stages on the microbiological contamination of broiler carcasses. *Italian Journal of Food Safety*, 6(4).

Althaus, D., Zweifel, C., Stephan, R. (2017). Analysis of a poultry slaughter process: Influence of process stages on the microbiological contamination of broiler carcasses. *Italian Journal of Food Safety*, 6(4).

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410.

Amann, R.I., Ludwig, W., Schleifer, K.H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews*, 59(1), 143-169.

Aminov, R.I. (2010). A brief history of the antibiotic era: lessons learned and challenges for the future. *Frontiers Microbiology*, 1, 134.

Anadón, A., Martinez-Larrañaga, M. R., Diaz, M. J., Bringas, P., Fernandez, M. C., Martinez, M. A., Fernandez-Cruz, M. L. (1996). Pharmacokinetics of amoxicillin in broiler chickens. *Avian Pathology*, 25(3), 449-458.

Andersen, S.C., Hoorfar, J. (2018). Surveillance of foodborne pathogens: Towards diagnostic metagenomics of fecal samples. *Genes*, 9, 14.

APHA/Defra (Animal & Plant Health Agency), 2020. Salmonella in livestock production in Great Britain. Available online: https://www.gov.uk/government/publications/salmonella-in-livestock-production-in-great-britain.

Arango-Argoty, G.A., Guron, G.K.P., Garner, E., Riquelme, M.V., Heath, L.S., Pruden, A., Vikesland, P.J., Zhang, L. (2020). ARGminer: a web platform for crowdsourcing-based curation of antibiotic resistance genes. *Bioinformatics*, 36(9), 2966-2973.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, 9, 75.

Barwell, L.J., Isaac, N.J.B., Kunin, W.E. (2015). Measuring β-diversity with species abundance data. *Journal of Animal Ecology*, 84(4), 1112-1122.

Battersby, T., Walsh, D., Whyte, P., Bolton, D. (2017). Evaluating and improving terminal hygiene practices on broiler farms to prevent *Campylobacter* cross-contamination between flocks. *Food Microbiology*, 64, 1-6.

Baym, M., Kryazhimskiy, S., Lieberman, T.D., Chung, H., Desai, M.M., Kishony, R. (2015). Inexpensive multiplexed library preparation for megabase-sized genomes. *PloS One*, 10(5), e0128036.

Bengtsson-Palme, J. (2018). The diversity of uncharacterized antibiotic resistance genes can be predicted from known gene variants-but not always. *Microbiome*, 6, 1-12.

Bengtsson-Palme, J., Boulund, F., Edström, R., Feizi, A., Johnning, A., Jonsson, V.A., Karlsson, F.H., Pal. C., Pereira, M.B., Rehammar, A., Sanchez, J., Sanli, K., Thorell, K. (2016). Strategies to improve usability and preserve accuracy in biological sequence databases. *Proteomics*, 16(18), 2454-60.

Bengtsson-Palme, J., Larsson, D.J., Kristiansson, E. (2017). Using metagenomics to investigate human and environmental resistomes. *Journal of Antimicrobial Chemotherapy*, 72(10), 2690-2703.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological*), 57(1), 289-300.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Boutell, J.M. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218), 53.

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Champomier Vergès, M.C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Herrero Corral, G., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I., Smidt, H., Schelkle, B., Roume, H., Seghal Kiran, G., Selvin, J., Correa de Souza, R.S., Van Overbeek, L., Singh, B. K., Wagner, M., Walsh, A., Sessitsch, A., Schloter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, 8, 103.

Blanco, P., Hernando-Amado, S., Reales-Calderon, J. A., Corona, F., Lira, F., Alcalde-Rico, M., Bernardini, A., Sanchez, M.B., Martinez, J.L. (2016). Bacterial multidrug efflux pumps: much more than antibiotic resistance determinants. *Microorganisms*, 4(1), 14.

Bogomolnaya, L.M., Andrews, K.D., Talamantes, M., Maple, A., Ragoza, Y., Vazquez-Torres, A., Andrews-Polymenis, H. (2013). The ABC-type efflux pump MacAB protects Salmonella enterica serovar typhimurium from oxidative stress. *MBio*, 4(6), e00630-13.

Bokulich, N.A., Lewis, Z.T., Boundy-Mills, K., Mills, D.A. (2016). A new perspective on microbial landscapes within food production. *Current Opinion in Biotechnology*, 37, 182-189.

Bolder, N.M. (2007). Microbial challenges of poultry meat production. *World's Poultry Science Journal*, 63(3), 401-411.

Bolton, F.J., Hutchinson, D.N., Coates, D. (1984). Blood-free selective medium for the isolation of *Campylobacter jejuni* from faeces. *Journal of Clinical Microbiology*, 19(2), 169-171.

Boolchandani, M., D'Souza A.W., Dantas, G. (2019). Sequencing-based methods and resources to study antimicrobial resistance. *Nature Reviews Genetics*, 20, 356-70.

Bray, J.R., & Curtis, J.T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs*, 27(4), 326-349.

Breitwieser, F.P., Lu, J., Salzberg, S.L. (2019). A review of methods and databases for metagenomic classification and assembly. *Briefings in Bioinformatics*, 20(4), 1125-1136.

Brooks, J.P., McLaughlin, M.R., Scheffler, B., Miles, D.M. (2010). Microbial and antibiotic resistant constituents associated with biological aerosols and poultry litter within a commercial poultry house. *Science of the Total Environment*, 408(20), 4770-4777.

Brugaletta, G., De Cesare, A., Zampiga, M., Laghi, L., Oliveri, C., Zhu, C., Manfreda, G., Syed, B., Valenzuela, L., Sirri, F. (2020). Effects of Alternative Administration Programs of a Synbiotic Supplement on Broiler Performance, Foot Pad Dermatitis, Caecal Microbiota, and Blood Metabolites. *Animals*, 10(3), 522.

Brul, S., Kallemeijn, W., Smits, G. (2008). Functional genomics for food microbiology: molecular mechanisms of weak organic acid preservative adaptation in yeast. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 3, 1-14.

Buchfink, B., Xie, C., Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59-60.

Buess, S., Zurfluh, K., Stephan, R., Guldimann, C. (2019). Quantitative microbiological slaughter process analysis in a large-scale Swiss poultry abattoir. *Food Control*, 105, 86-93.

Buttigieg, P.L., Hankeln, W., Kostadinov, I., Kottmann, R., Yilmaz, P., Duhaime, M.B., Glöckner, F.O. (2013). Ecogenomic perspectives on domains of unknown function: correlation-based exploration of marine metagenomes. *PloS One*, 8(3), e50869.

Calle, M.L. (2019). Statistical analysis of metagenomics data. *Genomics & Informatics*, 17(1).

Cao, J., Hu, Y., Liu, F., Wang, Y., Bi, Y., Lv, N., Li, J., Zhu, B., Gao, G.F. (2020). Metagenomic analysis reveals the microbiome and resistome in migratory birds. *Microbiome*, 8(1), 1-18.

Cao, Y., Fanning, S., Proos, S., Jordan, K., Srikumar, S. (2017). A review on the applications of next generation sequencing technologies as applied to food-related microbiome studies. *Frontiers in Microbiology*, 8, 1829.

Capita, R., & Alonso-Calleja, C. Antibiotic-Resistant Bacteria: A Challenge for the Food Industry. *Critical Reviews in Food Science and Nutrition*, 2013, 53, 11-48.

Cassini, A., Högberg, L.D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G.S., Colomb-Cotinat, M., Kretzschmar, M.E., Devleesschauwer, B., Cecchini, M., Ouakrim, D.A., Oliveira, T.C., Struelens, M.J., Suetens, C., Monnet, D.L., Hopkins, S. (2019). Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *The Lancet infectious diseases*, 19(1), 56-66.

Chaillou, S., Chaulot-Talmon, A., Caekebeke, H., Cardinal, M., Christieans, S., Denis, C., Desmonts, M.-H., Dousset X., Feurer, C., Hamon, E., Joffraud, J.-J., La Carbona, S., Leroi, F., Leroy, S., Lorre, S., Macé, S., Pilet, M.-F., Prévost, H., Rivollier, M., Roux, D., Talon, R., Zagorec, M., Champomier-Vergès, M.C. (2015). Origin and ecological selection of core and food-specific bacterial communities associated with meat and seafood spoilage. *The ISME Journal*, 9(5), 1105-1118.

Chan, C.K., Hsu, A.L., Tang, S.L., Halgamuge, S.K. (2008). Using growing self-organising maps to improve the binning process in environmental whole-genome shotgun sequencing. *Journal of Biomedicine and Biotechnology*, 2008, 513701.

Chantziaras, I., Boyen, F., Callens, B. Dewulf, J. (2014). Correlation between veterinary antimicrobial use and antimicrobial resistance in food-producing animals: a report on seven countries. *Journal Antimicrobial Chemotherapy*, 69, 827-834.

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, 265-270.

Chao, A., Lee, S.M. (1992). Estimating the number of classes via sample coverage. *Journal of the American Statistical Association*, 87(417), 210-217.

Chao, Y., Ma, L., Yang, Y., Ju, F., Zhang, X.X., Wu, W.M., Zhang, T. (2013). Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. *Scientific Reports*, 3, 3550.

Chen, H., Chen, R., Jing, L., Bai, X., Teng, Y. (2019). A metagenomic analysis framework for characterization of antibiotic resistomes in river environment: Application to an urban river in Beijing. *Environmental Pollution*, 245, 398-407.

Chen, S.Y., Feng, Z., Yi, X. (2017). A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease*, 9(6), 1725.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Müller, W.E., Wetter, T., Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Research*, 14(6), 1147-1159.

Chew, Y.V., & Holmes, A.J. (2009). Suppression subtractive hybridisation allows selective sampling of metagenomic subsets of interest. *Journal of Microbiological Methods,* 78, 136-143.

Choi, K.Y, Lee, T.K, Sul, W.J. (2015). Metagenomic Analysis of Chicken Gut Microbiota for Improving Metabolism and Health of Chickens — A Review. *Asian-Australasian Journal of Animal Sciences*, 28(9), 1217-1225.

Clavijo, V., & Flòrez, M.J.V. (2018). The gastrointestinal microbiome and its association with the control of pathogens in broiler chicken production: a review. *Poultry Science*, 97(3), 1006-1021.

Cocolin, L., & Ercolini, D. (2015). Zooming into food-associated microbial consortia: a 'cultural'evolution. *Current Opinion in Food Science*, 2, 43-50.

Cocolin, L., Alessandria, V., Dolci, P., Gorra, R., Rantsiou, K. (2013). Culture independent methods to assess the diversity and dynamics of microbiota during food fermentation. *International Journal of Food Microbiology*,167(1), 29-43.

Commission Decision 2000/96/EC of 22 December 1999 on the communicable diseases to be progressively covered by the Community network under Decision No 2119/98/EC of the European Parliament and of the Council. Official Journal of the European Communities L 28/50; 03.02.2000.

Commission Implementing Decision (EU) No 2013/652 of 12 November 2013 on the monitoring and reporting of antimicrobial resistance in zoonotic and commensal bacteria (notified under document C (2013) 7145). *Official Journal of the European Union*, L 303/26, 4.11.2013.

Commission Regulation (EC) No 1091/2005 of 12 July 2005 implementing Regulation (EC) No 2160/2003 of the European Parliament and of the Council as regards requirements for the use of specific control methods in the framework of the national programmes for the control of salmonella. *Official Journal of the European Union*, L 182, 13.7.2005, p. 3-4.

Cornejo, J., Pokrant, E., Carvallo, C., Maddaleno, A., San Martín, B. (2018). Depletion of tylosin residues in feathers, muscle and liver from broiler chickens after completion of antimicrobial therapy. *Food Additives & Contaminants: Part A*, 35(3), 448-457.

Crofts, T.S., Gasparrini, A.J., Dantas, G. (2017). Next-generation approaches to understand and combat the antibiotic resistome. *Nature Reviews Microbiology*, 15(7), 422-434.

Cuthbertson, L., Rogers, G.B., Walker, A.W., Oliver, A., Hafiz, T., Hoffman, L.R., Carroll, M.P., Bruce, K.D., Van Der Gast, C.J. (2014). Time between collection and storage significantly influences bacterial sequence composition in sputum samples from cystic fibrosis respiratory infections. *Journal of Clinical Microbiology*, 52(8), 3011-3016.

D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G.B., Poinar, H.N., Wright, G.D. (2011). Antibiotic resistance is ancient. *Nature*, 477(7365), 457-461.

Dalmasso, G., Nguyen, H.T.T., Yan, Y., Charrier-Hisamuddin, L., Sitaraman, S.V., Merlin, D. (2008). Butyrate transcriptionally enhances peptide transporter PepT1 expression and activity. *PloS One*, 3(6), e2476.

Darwich, L., Vidal, A., Seminati, C., Albamonte, A., Casado, A., López, F., Molina-Lopez, R., Migura-Garcia, L. (2019). High prevalence and diversity of extended-spectrum β-lactamase and emergence of OXA-48 producing *Enterobacterales* in wildlife in Catalonia. *PLoS One*, 14(8), e0210686.

Davies, R., & Wales, A. (2019). Antimicrobial resistance on farms: a review including biosecurity and the potential role of disinfectants in resistance selection. *Comprehensive reviews in food science and food safety*, 18(3), 753-774.

Davies, R., & Wales, A. (2019). Antimicrobial resistance on farms: a review including biosecurity and the potential role of disinfectants in resistance selection. *Comprehensive Reviews in Food Science and Food Safety*, 18(3), 753-774.

De Abreu, V.A., Perdigão, J., Almeida, S. (2021). Metagenomic Approaches to Analyze Antimicrobial Resistance: An Overview. *Frontiers in Genetics*, 1711.

De Cesare, A., Caselli, E., Lucchi, A., Sala, C., Parisi, A., Manfreda, G., & Mazzacane, S. (2019). Impact of a probiotic-based cleaning product on the microbiological profile of broiler litters and chicken caeca microbiota. *Poultry Science*, 98(9), 3602-3610.

De Cesare, A., Sirr,i F., Manfreda, G., Moniaci, P., Giardini, A., Zampiga, M., Meluzzi, A. (2017). Effect of dietary supplementation with Lactobacillus acidophilus D2/CSL (CECT 4529) on caecum microbioma and productive performance in broiler chickens. *PLoS One*, 12, e0176309.

De Filippis F, Parente E, Ercolini D. (2017b). Metagenomics insights into food fermentations. *Microbial Biotechnology*, 10(1), 91-102.

De Filippis, F., La Storia, A., Blaiotta, G. (2017a). Monitoring the mycobiota during Greco di Tufo and Aglianico wine fermentation by 18S rRNA gene sequencing. *Food Microbiology*, 63, 117-122.

De Filippis, F., La Storia, A., Villani, F., Ercolini, D. (2013). Exploring the sources of bacterial spoilers in beefsteaks by culture-independent high-throughput sequencing. *PLoS One*, 8(7), e70222.

De Filippis, F., Parente, E., Ercolini, D. (2017). Metagenomics insights into food fermentations. *Microbial Biotechnology*, 10(1), 91-102.

De Filippis, F., Parente, E., Ercolini, D. (2018). Recent Past, Present, and Future of the Food Microbiome. *Annual Review of Food Science and Technology*, 9, 589-608.

De Pasquale, I., Di Cagno, R., Buchin, S., De Angelis, M., Gobbetti, M. Microbial ecology dynamics reveal a succession in the core microbiota involved in the ripening of pasta filata Caciocavallo Pugliese cheese. (2014). *Applied and Environmental Microbiology*, 80(19), 6243-6255.

Delcher, A.L., Harmon, D., Kasif, S., White, O., Salzberg, S.L. (1999). Improved microbial gene identification with GLIMMER. *Nucleic Acids Research*, 27(23), 4636-4641.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., Von Mering, C., Vorholt, J.A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences*, 106(38), 16428-16433.

Deng, W., Zhang, A., Chen, S., He, X., Jin, L., Yu, X., Yang, S., Li, B., Fan, L., Ji, L., Pan, X., Zou, L. (2020). Heavy metals, antibiotics and nutrients affect the bacterial community and resistance

genes in chicken manure composting and fertilized soil. *Journal of Environmental Management*, 257, 109980.

DeSa (Department of Economic and Social Affairs), 2015. World population projected to reach 9.7 billion by 2050. UN DESA United Nations Department of Economic and Social Affairs, New York, NY, USA.

Deshpande, V., Fung, E.D.K., Pham, S., Bafna, V. Cerulean: A Hybrid Assembly Using High Throughput Short and Long Reads. In: Darling A., Stoye J. (eds) Algorithms in Bioinformatics. WABI 2013. Lecture Notes in Computer Science, vol 8126. Springer, Berlin, Heidelberg, pp 349-363.

Di Francesco, C.E., Smoglica, C., Profeta, F., Farooq, M., Di Giannatale, E., Toscani, T., Marsilio, F. (2021). Research Note: Detection of antibiotic-resistance genes in commercial poultry and turkey flocks from Italy. *Poultry Science*, 100(5), 101084.

Diaz, M., Kellingray, L., Akinyemi, N., Adefiranye, O.O., Olaonipekun, A.B., Bayili, G.R., Ibezim, J., Du Plessis, A.S., Houngbédji, M., Kamya, D., Mukisa, I.M., Mulaw, G., Manthi Josiah, S., Onyango Chienjo, W., Atter, A., Agbemafle, E., Annan, T., Bernice Ackah, N., Buys, E.M., Joseph Hounhouigan, D., Muyanja, C., Nakavuma, J., Odeny, D.A., Sawadogo-Lingani, H., Tesfaye Tefera, A., Amoa-Awua, W., Obodai, M., Mayer, M.J., Oguntoyinbo, F.A., Narbad, A. (2019). Comparison of the microbial composition of African fermented foods using amplicon sequencing. *Scientific Reports*, 9(1), 1-8.

Diaz-Sanchez, S., Hanning, I., Pendleton, S., D'Souza, D. (2013). Next-generation sequencing: the future of molecular genetics in poultry production and food safety. *Poultry science*, 92(2), 562-572.

Dinsdale, E.A., Edwards, R.A., Hall, D., Angly, F., Breitbart, M., Brulc, J.M., Furlan, M., Desnues, C., Haynes, M., Li, L., McDaniel, L., Moran, M.A., Nelson, K.E., Nilsson, C., Olson, R., Paul, J., Rodriguez Brito, B., Ruan, Y., Swan, B.K., Stevens, R., Valentine, D.L., Vega Thurber, R., Wegley, L., White, B.A., Rohwer, F. (2008). Functional metagenomic profiling of nine biomes. *Nature*, 452(7187), 629-632.

Directive 2003/99/EC of the European Parliament and of the Council of 17 November 2003 on the monitoring of zoonoses and zoonotic agents, amending Council Decision 90/424/EEC and repealing Council Directive 92/117/EEC. *Official Journal of European Union*, L 325, 12.12.2003, pp. 1-10.

Dixon, P., (2007). Vegan: a package of R functions for community ecology. *Journal of Vegetation Science*, 2003, 14(6), 927-930.

Dolejska, M. (2020). Antibiotic-resistant bacteria in wildlife. *Antibiotic Resistance in the Environment: A Worldwide Overview*, 19-70.

Doster, E., Lakin, S.M., Dean, C.J., Wolfe, C., Young, J.G., Boucher, C., Belk, K.E., Noyes, N.R., Morley, P.S. (2019). MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. *Nucleic Acids Research*, 48(D1), D561-D569.

Dugat-Bony, E., Straub, C., Teissandier, A., Onésime, D., Loux, V., Monnet, C., Irlinger, F., Landaud, S., Leclercq-Perlat, M.N., Bento, P., Fraud, S., Gibrat, J.F., Aubert, J., Fer, F., Guédon, E., Pons, N., Kennedy, S., Beckerich, J.M., Swennen, D., Bonnarme, P. (2015). Overview of a surface-ripened cheese community functioning by meta-omics analyses. *PloS One*,10, e0124360.

Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D., De Cesare, A. (2021). Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Scientific Reports*, 11(1), 1-10.

ECDC (European Centre for Disease Prevention and Control); EFSA (European Food Safety Authority); EMA (European Medicines Agency). Third joint inter-agency report on integrated analysis of consumption of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from humans and food-producing animals in the EU/EEA. *EFSA Journal* 2021, 19, 6712.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Steven R. Gill, S.R., Nelson, K.E., Relman, D.A. (2005). Diversity of the human intestinal microbial flora. *Science*, 308(5728), 1635-1638.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19), 2460-2461.

EFSA and ECDC (European Food Safety Authority and European Centre for Disease Prevention and Control), 2021a. The European Union One Health 2019 Zoonoses Report. *EFSA Journal*, 2021;19(2):6406, 286 pp.

EFSA and ECDC (European Food Safety Authority and European Centre for Disease Prevention and Control), 2021b. The European Union Summary Report on Antimicrobial Resistance in zoonotic and indicator bacteria from humans, animals and food in 2018/2019. *EFSA Journal* 2021;19(4):6490, 179 pp.

EFSA and ECDC (European Food Safety Authority and European Centre for Disease Prevention Control), 2021. The European Union Summary Report on Antimicrobial Resistance in zoonotic and indicator bacteria from humans, animals and food in 2018/2019. *EFSA Journal* 2021;19(4):6490, 179 pp.

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), 2014. Scientific Opinion on the evaluation of the safety and efficacy of peroxyacetic acid solutions for reduction of pathogens on poultry carcasses and meat. EFSA J. 2014, 12(3), 3599, 60 pp;

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Jenkins C, Malorny B, Ribeiro Duarte AS, Torpdahl M, da Silva Felıcio MT, Guerra B, Rossi M and Herman L, 2019. Whole genome sequencing and metagenomics for outbreak investigation, source attribution and risk assessment of food-borne microorganisms. *EFSA Journal* 2019;17(12):5898, 78 pp.

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Herman L, Hilbert F, Lindqvist R, Nauta M, Ru G, Simmons M, Skandamis P, Suffredini E, Argu€ello H, Berendonk T, Cavaco LM, Gaze W, Schmitt H, Topp E, Guerra B, Liébana E, Stella P and Peixe L, 2021. Scientific Opinion on the role played by the environment in the emergence and spread of antimicrobial resistance (AMR) through the food chain. *EFSA Journal* 2021;19(6):6651, 188 pp.

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, De Cesare A, Herman L, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Dewulf J, Hald T, Michel V, Taina N, Ricci A, Snary E, Boelaert F, Messens W and Davies R, 2019. Scientific Opinion on the Salmonella control in poultry flocks and its public health impact. *EFSA Journal*, 2019, 17(2), e5596.

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Herman L, Hilbert F, Lindqvist R, Nauta M, Ru G, Simmons M, Skandamis P, Suffredini E, Argu€ello H, Berendonk T, Cavaco LM, Gaze W, Schmitt H, Topp E, Guerra B, Liebana E, Stella P and Peixe L, 2021.

Scientific Opinion on the role played by the environment in the emergence and spread of antimicrobial resistance (AMR) through the food chain. *EFSA Journal* 2021;19(6):6651, 188 pp.

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), Koutsoumanis K, Allende A, Alvarez-Ordon~ez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Herman L, Hilbert F, Lindqvist R, Nauta M, Ru G, Simmons M, Skandamis P, Suffredini E, Argu€ello H, Berendonk T, Cavaco LM, Gaze W, Schmitt H, Topp E, Guerra B, Liebana E, Stella P and Peixe L, 2021. Scientific Opinion on the role played by the environment in the emergence and spread of anti-microbial resistance (AMR) through the food chain. *EFSA Journal*, 2021, 19(6), 6651, 188 pp

EMA (European Medicines Agency - European Surveillance of Veterinary Antimicrobial Consumption), 2020. Sales of veterinary antimicrobial agents in 31 European countries in 2018 (EMA/24309/2020).

EMA and EFSA (European Medicines Agency and European Food Safety Authority), 2017. Joint Scientific Opinion on measures to reduce the need to use antimicrobial agents in animal husbandry in the European Union, and the resulting impacts on food safety (RONAFA). *EFSA Journal* 2017;15(1):4666, 245 pp.

EMA and EFSA (European Medicines Agency and European Food Safety Authority), 2017. EMA and EFSA Joint Scientific Opinion on measures to reduce the need to use antimicrobial agents in animal husbandry in the European Union, and the resulting impacts on food safety (RONAFA). *EFSA Journal* 2017;15(1):4666, 246 pp.

EN ISO 10272-1:2006. Microbiology of Food and Animal Feeding Stuffs - Horizontal Method for Detection and Enumeration of *Campylobacter* spp. - Part 1: Detection method. *International Organization for Standardization, Geneva* (2006).

EN ISO 6579:2002/Amd.1:2007. Microbiology of Food and Animal Feeding Stuffs – Horizontal Method for the Detection of *Salmonella* spp.–Annex D: Detection of Salmonella Spp. in Animal Faeces and in Environmental Samples from the Primary Production Stage. *International Organization for Standardization, Geneva, Switzerland* (2007).

EN ISO 6888-1:1999/Amd.1:2003. Microbiology of food and animal feeding stuffs - Horizontal method for the enumeration of coagulase-positive staphylococci (*Staphylococcus aureus* and other species) - Part 1:Technique using Baird-Parker agar medium. AMENDMENT: Inclusion of precision data. *International Organization for Standardization, Geneva, Switzerland* (2003).

Ercolini, D. (2013). High-throughput sequencing and metagenomics: moving forward in the culture-independent analysis of food microbial ecology. *Applied and Environmental Microbiology*, 79(10), 3148-3155.

Ercolini, D., Moschetti, G., Blaiotta, G., Coppola, S. (2001). The potential of a polyphasic PCR-DGGE Approach in evaluating microbial diversity of natural whey cultures for water-buffalo mozzarella cheese production: bias of culture-dependent and culture-independent analyses. *Systematic and applied microbiology*, 24(4), 610-617.

Escobar-Zepeda, A., Sanchez-Flores, A., Baruch, M.Q. (2016). Metagenomic analysis of a Mexican ripened cheese reveals a unique complex microbiota. *Food Microbiology*, 57, 116-127.

European Commission. Commission Regulation (EC) No 2073/ 2005 of 15 November 2005 on microbiological criteria for foodstuffs. *Official Journal of European Union*, 2005, 1-26.

Eurostat Statistics Explained, 2121. Agricultural production- livestock and meat. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agricultural_production_-_livestock_and_meat&oldid=427096#Meat_production (accessed on 13 October 2021).

Fernández-Rubio, C., Ordonez, C., Abad-González, J., Garcia-Gallego, A., Honrubia, M.P., Mallo, J.J., Balana-Fouce, R. (2009). Butyric acid-based feed additives help protect broiler chickens from Salmonella Enteritidis infection. *Poultry Science*, 88(5), 943-948.

Feye, K.M., Baxter, M.F.A., Tellez-Isaias, G., Kogut, M.H., Ricke, S.C. (2020). Influential factors on the composition of the conventionally raised broiler gastrointestinal microbiomes. *Poultry Science*, 99(2), 653-659.

Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H., Caporaso, J.G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, 109, 21390-95.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L.L., Tate, J., Punta, M. (2014). Pfam: the protein families database. *Nucleic Acids Research*, 42(D1), D222-D230.

Fitzpatrick, W.P., Llabres, S., Neuberger, A., Blaza, J.N., Bai, X.C., Okada, U., Murakami, S., van Veen, H.W., Zachariae, U., Scheres, S.H.W., Luisi, B.F., Du, D. (2017). Structure of the MacAB–TolC ABC-type tripartite multidrug efflux pump. *Nature Microbiology*, 2(7),17070-17070.

Fleming, A. (2001). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *British Journal of Experimental Pathology,* 79(8), 780-90.

Fluckey, W.M., Sanchez, M.X., McKee, S.R., Smith, D., Pendleton, E., Brashears, M.M. (2003). Establishment of a microbiological profile for an air-chilling poultry operation in the United States. *Journal of Food Protection*, 66(2), 272-279.

Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W., DeLong, E.F. (2008). Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences*, 105, 3805-10.

Frosth, S., Karlsson-Lindsjö, O., Niazi, A., Fernström, L.L., Hansson, I. (2020). Identification of transmission routes of *Campylobacter* and on-farm measures to reduce *Campylobacter* in chicken. *Pathogens*, 9(5), 363.

Garcia-Garcerà, M., Garcia-Etxebarria, K., Coscollà, M., Latorre, A., Calafell, F. (2013). A new method for extracting skin microbes allows metagenomic analysis of whole-deep skin. *PloS One,* 8, e74914.

Gentleman, R.C., Carey, V.J., Bates, D.M. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10), 1-16.

Ghosh, A., Mehta, A., Khan, A.M. (2019). Metagenomic analysis and its applications. *Encyclopedia of Bioinformatics and Computational Biology*, 3, 184-193.

Gibson, M.K., Forsberg, K.J. Dantas, G. (2015). Improved annotation of antibiotic. Resistance determinants reveals microbial resistomes cluster by ecology. *ISME Journal*, 9(1), 207-216.

Gibson, M.K., Forsberg, K.J., Dantas, G. (2015). Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *The ISME Journal*, 9(1), 207-216.

Gill, C.O. Sources of microbial contamination a slaughtering plants. J.N. Sofos (Ed.), Improving the safety of fresh meat, CRC/Woodhead Publishing Limited, Cambridge, UK, 2005, 231-243.

Giraffa, G. & Neviani, E. (2001). DNA-based, culture-independent strategies for evaluating microbial communities in food-associated ecosystems. *International Journal of Food Microbiology*, 67(1-2), 19-34.

Glass, E.M., Wilkening, J., Wilke, A., Antonopoulos, D., Meyer, F. (2010). Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harbor Protocols*, 2010(1), dbrot5368.

Glass, E.M., Wilkening, J., Wilke, A., Antonopoulos, D., Meyer, F. (2010). Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harbor Protocols*, 2010(1), pdb-prot5368.

Godzik, A. (2011). Metagenomics and the protein universe. *Current Opinion in Structural Biology*, 21, 398-403.

Gorelick, R. Combining richness and abundance into a single diversity index using matrix analogues of Shannon's and Simpson's indices. *Ecography*, 2006, 29(4), 525-530.

Graham, J.P., Price, L.B., Evans, S.L., Graczyk, T.K., Silbergeld, E.K. (2009). Antibiotic resistant enterococci and staphylococci isolated from flies collected near confined poultry feeding operations. *Science of the Total Environment*, 407(8), 2701-2710.

Griffith, A., Neethirajan, S., Warriner, K. (2015). Development and evaluation of silver zeolite antifouling coatings on stainless steel for food contact surfaces. *Journal of Food Safety*, 35(3), 345-354.

Groat, E. F., Williams, N. J., Pinchbeck, G., Simpson, A., Schmidt, V. M. (2016, April). Canine raw meat diets and antimicrobial resistant E. coli: is there a link?. In *BSAVA Congress Proceedings 2016* (pp. 540-540). BSAVA Library.

Gupta, C.L., Tiwari, R.K., Cytryn, E. (2020). Platforms for elucidating antibiotic resistance in single genomes and complex metagenomes. *Environment International*, 138, 105667.

Gupta, S.K., Padmanabhan, B.R., Diene, S.M., Lopez-Rojas, R., Kempf, M., Landraud, L., Rolain, J.M. (2014). ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial Agents and Chemotherapy*, 58(1), 212-220.

Hanage, W.P. Microbiome science needs a healthy dose of scepticism. *Nature*, 2014, 512, 247-48.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., Goodman, R.M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology*, 5(10), R245-R249.

Hendriksen, R.S., Bortolaia, V., Tate, H., Tyson, G.H., Aarestrup, F.M., Mcdermott, P.F. (2019). Using genomics to track global antimicrobial resistance. *Frontiers in public health*, 7, 242.

Hervé M. (2021). RVAideMemoire: Testing and Plotting Procedures for Biostatistics. 2021. https://cran.r-project.org/package=RVAideMemoire. Accessed 5 Jul 2021.

Hiergeist, A., Gläsner, J., Reischl, U., Gessner, A. (2015). Analyses of Intestinal Microbiota: Culture versus Sequencing. *ILAR Journal*, 56(2), 228-40.

Hinton, M. (2000). Microbial control in the meat industry. Flair Flow Europe Technical manual 379A/2000. Ed. Gormley R. National Food Centre.

Hoff, K.J., Lingner, T., Meinicke, P., Tech, M. (2009). Orphelia: predicting genes in metagenomic sequencing reads. *Nucleic Acids Research*, 37(suppl_2), W101-105.

Hu, Y., Yang, X., Qin, J., Lu, N., Cheng, G., Wu, N., Pan, Y., Li, J., Zhu, L., Wang, X., Meng, Z., Zhao, F., Liu, D., Ma, J., Qin, N., Xiang, C., Xiao, *Y., Li, L., Yang, H., Wang, J.*, Yang, R., Gao, G.F., Wang, J., Zhu, B. (2013). Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nature Communications*, 4, 2151.

Hue, O., Allain, V., Laisney, M.J., Le Bouquin, S., Lalande, F., Petetin, I., Rouxel, S., Quesne, S., Gloaguen, P.Y., Picherot, M., Santolini, J., Bougeard, S., Salvat, G., Chemaly, M. (2011). *Campylobacter* contamination of broiler caeca and carcasses at the slaughterhouse and correlation with Salmonella contamination. *Food Microbiology*, 28(5), 862-868.

Hultman, J., Rahkila, R., Ali, J., Rousu, J., Björkroth, K. J. (2015). Meat processing plant microbiome and contamination patterns of cold-tolerant bacteria causing food safety and spoilage risks in the manufacture of vacuum-packaged cooked sausages. *Applied and Environmental Microbiology*, 81(20), 7088-7097.

Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402), 207.

Huson, D.H., & Weber, N. (2013). Microbial community analysis using MEGAN. *Methods in Enzymology*, 531, 465-85.

Illeghems, K., Weckx, S., De Vuyst, L. (2015). Applying meta-pathway analyses through metagenomics to identify the functional properties of the major bacterial communities of a single spontaneous cocoa bean fermentation process sample. *Food Microbiology*, 50, 54-63.

International Commission On Microbiological Specifications For Foods (ICMSF). *Microorganisms in Foods* 2. *Sampling for Microbiological Analysis: Principles and Specific Applications*. 2nd. Edn. University of Toronto Press, Toronto, Canada, 1986.

Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Dröge, M., Gartemann, K.H., Jünemann, S., Kaiser, O., Krause, L., Tille, F., Zakrzewski, M., Pühler, A., Schlüter, A., Goesmann, A. (2011). Comparative and joint analysis of two metagenomics datasets from a biogas fermenter obtained by 454-pyrosequencing. *PloS One*, 6(1), e14519.

Jahan, N.A., Lindsey, L.L., Larsen, P.A. (2021). The role of peridomestic rodents as reservoirs for zoonotic foodborne pathogens. *Vector-Borne and Zoonotic Diseases*, 21(3), 133-148.

Jayarao, B., Almeida, R., Oliver, S.P. (2019). Antimicrobial resistance on dairy farms. *Foodborne pathogens and disease*, 16(1), 1-4.

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., Sharma, A.N., Doshi, S., Courtot, M., Lo, R., Williams, L.E., Frye, J.G., Elsayegh, T., Sardar, D., Westman, E.L., Pawlowski, A.C., Johnson, T.A., Brinkman, F.S.L., Wright, G.D., McArthur, A.G., (2017). CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*, 45(D1), D566-D573.

Jonsson, V., Österlund, T., Nerman, O., Kristiansson, E. (2016). Statistical evaluation of methods for identification of differentially abundant genes in comparative metage- nomics. *BMC Genomics*, 17(1), 1-14.

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A.L., Madsen, K.L., Wong, G.K.S. (2016). Characterization of the gut micro- biome using 16S or shotgun metagenomics. *Frontiers in Microbiology*, 7, 459.

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Research*, 32(suppl_1), D277-280.

Keegan, K.P., Glass, E.M., Meyer, F. (2016). MG-RAST, a metagenomics service for analysis of microbial community structure and function. In *Microbial Environmental Genomics (MEG)* (pp. 207-233). Humana Press, New York, NY.

Kent, W.J. (2002). BLAT—the BLAST-like alignment tool. *Genome Research*, 12(4), 656-64.

Kim, D.W., & Cha, C.J. (2021). Antibiotic resistome from the One-Health perspective: understanding and controlling antimicrobial resistance transmission. *Experimental & Molecular Medicine*, 53(3), 301-309.

Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C., Debelius, J., Gonzalez, A., Kosciolek, T., McCall, L.I., McDonald, D., Melnik, A.V., Morton, J.T., Navas, J., Quinn, R.A., Sanders, J.G., Swafford, A.D., Thompson, L.R., Tripathi, A., Xu, Z.Z., Zaneveld, J.R., Zhu, Q., Caporaso, J.G., Dorrestein, P.C. (2018a). Best practices for analysing microbiomes. *Nature Reviews Microbiology*, 16(7), 410-422.

Knight, S.J., Klaere, S., Morrison-Whittle, P., Goddard, M. R. (2018b). Fungal diversity during fermentation correlates with thiol concentration in wine. *Australian Journal of Grape and Wine Research*, 24(1), 105-112.

Kogut, M.H. (2019). The effect of microbiome modulation on the intestinal health of poultry. *Animal Feed Science and Technology*, 250, 32-40.

Konopka, A. (2009). What is microbial community ecology? *The ISME Journal*, 3(11), 1223-1230.

Korsak, N., Taminiau, B., Hupperts, C., Delhalle, L., Nezer, C., Delcenserie, V., Daube, G. (2017). Assessment of bacterial superficial contamination in classical or ritually slaughtered cattle using metagenetics and microbiological analysis. *International Journal of Food Microbiology*, 247, 79-86.

Koutsoumanis, K.P., Kendall, P.A., Sofos, J.N. (2004). A comparative study on growth limits of *Listeria monocytogenes* as affected by temperature, pH and aw when grown in suspension or on a solid surface. *Food Microbiology*, 21(4), 415-422.

Kristiansson, E., Hugenholtz, P., Dalevi, D. (2009). Shotgun Functionalize R: An R-package for functional comparison of metagenomes. *Bioinformatics*, 25(20), 2737-2738.

Lakin, S.M., Dean, C., Noyes, N.R., Dettenwanger, A., Ross, A.S., Doster, E., Rovira, P., Abdo, Z., Jones, K.L., Ruiz, J., Belk, K.E., Morley, P.S., Boucher, C. (2017). MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Research* 45(D1), D574-D580.

Lane-Petter, W. (1962). The provision and use of pathogen-free laboratory animals. *Proceedings of the Royal Society of Medicine*, 55(4), 253-263.

Langmead, B., Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357-359.

Langmead, B., Trapnell, C., Pop, M., Sailzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3), 1-10.

Lattanzi, A., Minervini, F., Di Cagno, R., Diviccaro, A., Antonielli, L., Cardinali, G., Cappelle, S., De Angelis, M., Gobbetti, M. (2013). The lactic acid bacteria and yeast microbiota of eighteen sourdoughs used for the manufacture of traditional Italian sweet leavened baked goods. *International Journal of Food Microbiology*, 163(2-3), 71-79.

Laudadio, I., Fulci, V., Palone, F., Stronati, L., Cucchiara, S., Carissimi, C. (2018). Quantitative assessment of shotgun metagenomics and 16S rDNA amplicon sequencing in the study of human gut microbiome. *Omics: a Journal of Integrative Biology*, 22(4), 248-254.

Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A. K., Wertheim, H. F., Sumpradit, N., Vlieghe, E., Hara, G.L., Gould, I.M., Goossens, H., Greko, C., So, A.D., Bigdeli, M., Tomson, G., Woodhousen, W., Ombaka, E., Quizhpe Peralta, A., Qamar, F.N., Mir, F., Kariuki, S., Bhutta, Z.A., Coates, A., Bergstrom, R., Wright, G.D., Brown, E.D., Cars, O. (2013). Antibiotic resistance—the need for global solutions. *The Lancet Infectious Diseases*, 13(12), 1057-1098.

Lederberg, J., McCray, A.T. (2001). Ome SweetOmics—A genealogical treasury of words. *The Scientist*, 15(7), 8-8.

Legendre, P. & Gallagher, E.D. (2001). Ecologically meaningful transformations for ordination ofspecies data. *Oecologia*, 129(2), 271-280.

Leonard, S.R., Mammel, M.K., Lacher, D.W., Elkins, C.A. (2015). Application of metagenomic sequencing to food safety: detection of Shiga toxin–producing *Escherichia coli* on fresh bagged spinach. Applied Environ. Microbiol. 81(23), 8183-8191.

Leroy, S., Giammarinaro, P., Chacornac, J. P., Lebert, I., Talon, R. (2010). Biodiversity of indigenous staphylococci of naturally fermented dry sausages and manufacturing environments of small-scale processing units. *Food Microbiology*, 27(2), 294-301.

Li, D., Liu, C.M., Luo, R., Sadakane, K., Lam, T.W. (2015). MEGAHIT: An ultra-fast single- node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10), 1674-1676.

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589-595.

Li, S., Mann, D.A., Zhang, S., Qi, Y., Meinersmann, R.J., Deng, X. (2020). Microbiome-Informed Food Safety and Quality: Longitudinal Consistency and Cross-Sectional Distinctiveness of Retail Chicken Breast Microbiomes. *Msystems*, 5(5), e00589-20.

Li, S., Mann, D.A., Zhang, S., Qi, Y., Meinersmann, R.J., Deng, X. (2020). Microbiome-Informed Food Safety and Quality: Longitudinal Consistency and Cross-Sectional Distinctiveness of Retail Chicken Breast Microbiomes. *Msystems*, 5(5), e00589-20.

Liu, B. & Pop, M. (2009). ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Research*, 37(suppl_1), D443-D447.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Li, D., Lu, L., Law, M. (2012). Comparison of next-generation sequencing systems. *Journal of BioMed Research International*, 2012.

Loman, N.J., & Pallen, M.J. (2015). Twenty years of bacterial genome sequencing. *Nature Reviews Microbiology*, 13, 1-9.

Looft, T., Johnson, T.A., Allen, H.K., Bayles, D.O., Alt, D.P., Stedtfeld, R.D., Sul, W.J., Stedtfeld, T.M., Chai, B., Cole, J.R., Hashsham, S.A., Tiedje, J.M., Stanton, T.B. (2012). In-feed antibiotic effects on the swine intestinal microbiome. *Proceedings of the National Academy of Sciences, U.S.A*, 109(5), 1691-1696.

Love, M., Anders, S., Huber, W. (2014a). Differential analysis of count data–the DESeq2 package. *Genome Biology*, 15(550), 10-1186.

Love, M.I., Huber, W., Anders, S. (2014b). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1-21.

Lozupone, C., Hamady, M., Knight, R. (2006). UniFrac – an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics*, 7, 371.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I, Jansson, J.K., Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, 489, 220-30.

Luber, P. (2009). Cross-contamination versus undercooking of poultry meat or eggs—which risks need to be managed first? *International Journal of Food Microbiology*, 134(1-2), 21-28.

Luiken, R.E., Van Gompel, L., Bossers, A., Munk, P., Joosten, P., Hansen, R.B., Knudsen, B.E., Garcıa-Cobos, S., Dewulf, J., Aarestrup, F.M., Wagenaar, J.A., Smit, L.A.M., Mevius, D.J., Heederik,

D.J.J., Schmitt, H. (2020). Farm dust resistomes and bacterial microbiomes in European poultry and pig farms. *Environment International*, 143, 105971.

Ma, L., Li, B., Jiang, X.T., Wang, Y.L., Xia, Y., Li, A.D., Zhang, T. (2017). Catalogue of antibiotic resistome and host-tracking in drinking water deciphered by a large scale survey**.** *Microbiome*, 2017, 5(1), 154.

Manual, Instruction. PowerFood™ DNA Isolation Kit.

Marangoni, F., Corsello, G., Cricelli, C., Ferrara, N., Ghiselli, A., Lucchin, L., & Poli, A. (2015). Role of poultry meat in a balanced diet aimed at maintaining health and wellbeing: An Italian consensus document. *Food & Nutrition Research*, 59(1), 27606.

Marchesi, J.R., & Ravel, J. (2015). The vocabulary of microbiome research: a proposal. *Microbiome*, 1-31.

Marine, R., Polson, S.W., Ravel, J., Hatfull, G., Russell, D., Sullivan, M., Syed, F., Dumas, M., Wommack, K.E. (2011). Evaluation of a transposase protocol for rapid generation of shotgun high-throughput sequencing libraries from nanogram quantities of DNA. *Applied and Environmental Microbiology*, 77(22), 8071-8079.

Markowitz, V.M., Ivanova, N.N., Szeto, E., Palaniappan, K., Chu, K., Dalevi, D., Chen, I.M.A., Grechkin, Y., Dubchak, I., Anderson, I., Lykidis, A., Mavromatis, K., Hugenholtz, P., Kyrpides, N.C. (2008). IMG/M: A data management and analysis system for metagenomes. *Nucleic Acids Research*, 36(suppl_1), D534-D538.

Markowitz, V.M., Ivanova, N.N., Szeto, E., Palaniappan, K., Chu, K., Dalevi, D., Chen, I-M.A., Grechkin, Y., Dubchak, I., Anderson, I., Lykidis, A., Mavromatis, K., Hugenholtz, P., Kyrpides, N. C. (2008). IMG/M: a data management and analysis system for metagenomes. *Nucleic Acids Research*, 36(suppl_1), D534-D538.

Marquez, B. (2005). Bacterial efflux systems and efflux pumps inhibitors. *Biochimie*, 87(12), 1137-1147.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), 10-12.

Mavromatis, K., Ivanova, N., Barry, K., Shapiro, H., Goltsman, E., McHardy, A.C., Rigoutsos, I., Salamov, A., Korzeniewski, F., Land, M., Lapidus, A., Grigoriev, I., Richardson, P., Hugenholtz, P., Kyrpides, N.C. (2007). Use of simulated data sets to evaluate the fidelity of metagenomic processing methods. *Nature Methods*, 4(6), 495-500.

Mayo, B., TCC Rachid, C., Alegría, Á., MO Leite, A., S Peixoto, R., Delgado, S. (2014). Impact of next generation sequencing techniques in food microbiology. *Current genomics*, 15(4), 293-309.

McElhoe, J.A., Holland, M.M., Makova, K.D., Su, M.S. W., Paul, I.M., Baker, C.H., Young, B. (2014). Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. *Forensic Science International: Genetics*, 13, 20-29.

McHardy, A.C., Martìn, H.G., Tsirigos, A., Hugenholtz, P., Rigoutsos, I. (2007). Accurate phylogenetic classification of variable-length DNA fragments. *Nature Methods*, 4(1), 63-72.

McMurdie, P.J. (2018). Normalization of microbiome profiling data. In: *Microbiome Analysis*. Humana Press, New York, NY, pp. 143-168.

McMurdie, P.J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysisand graphics of microbiome census data. *PloS One*, 8(4), e61217.

Mead, G.C. (2004). Microbiological quality of poultry meat: a review. *Revista Brasileira de Ciência Avícola*, 6(3), 135-142.

Medinger, R., Nolte, V., Pandey, R. V., Jost, S. Ottenwälder, B., Schlötterer, C., Boenigk, J. (2010). Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Molecular Ecology*, 19(1), 32-40.

Mende, D.R., Waller, A.S., Sunagawa, S., Järvelin, A.I., Chan, M.M., Arumugam, M., Raes, J., Bork, P. (2012). Assessment of metagenomic assembly using simulated next generation sequencing data. *PloS One*, 7(2), e31386.

Metzker, M.L. (2010). Sequencing technologies - the next generation. *Nature Reviews Genetics*, 11(1), 31-46.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, t., Rodriguez, A., Stevens, R., Wilke, A., Wilkening J., Edwards, R.A. (2008). The metagenomics RAST server — a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9, 386 (2008).

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R.A. (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9(1), 1-8.

Minot, S.S., Krumm, N., Greenfield, N.B. (2015). One codex: a sensitive and accurate data platform for genomic microbial identification. *BioRxiv*, 027607.

Monecke, S., Ruppelt, A., Wendlandt, S., Schwarz, S., Slickers, P., Ehricht, R., de Jäckel, S.C. (2013). Genotyping of Staphylococcus aureus isolates from diseased poultry. *Veterinary Microbiology*, 162(2-4), 806-812.

Montoro-Dasí, L., Villagrá, A., Sevilla-Navarro, S., Pérez-Gracia, M. T., Vega, S., Marin, C. (2020). The dynamic of antibiotic resistance in commensal Escherichia coli throughout the growing period in broiler chickens: Fast-growing vs. slow-growing breeds. *Poultry Science*, 99(3), 1591-1597.

Moore, J.E., Rao, J.R., Moore, P.J., Millar, B.C., Goldsmith, C.E., Loughrey, A., Rooney, P.J. (2010). Determination of total antibiotic resistance in waterborne bacteria in rivers and streams in Northern Ireland: can antibiotic-resistant bacteria be an indicator of ecological change?. *Aquatic Ecology*, 44(2), 349-358.

More, S.J., (2020). European perspectives on efforts to reduce antimicrobial usage in food animal production. *Irish Veterinary Journal*, 73(2), 1-12.

Møretrø, T., Moen, B., Heir, E., Hansen, A. Å., Langsrud, S. (2016). Contamination of salmon fillets and processing plants with spoilage bacteria. *International Journal of Food Microbiology*, 237, 98-108.

Mouttotou, N., Ahmad, S., Kamran, Z., & Koutoulis, K. C. (2017). Prevalence, risks and antibiotic resistance of Salmonella in poultry production chain. *Current topics in Salmonella and Salmonellosis*, 215-234.

Mughini-Gras, L., Pasqualin, D., Tarakdjian, J., Santini, A., Cunial, G., Tonellato, F., Schiavon, E., Martino, G. D. (2021). Short-term and long-term effects of antimicrobial use on antimicrobial resistance in broiler and turkey farms. *Avian Pathology*, (just-accepted), 1-25.

Mulders, M.N., Haenen, A.P.J., Geenen, P.L., Vesseur, P.C., Poldervaart, E. S., Bosch, T., Huijsdens, X., Hengeveld, P.D., Dam-Deisz, W., Graat, E., Mevius, D., Voss, A., Van De Giessen, A. W. (2010). Prevalence of livestock-associated MRSA in broiler flocks and risk factors for slaughterhouse personnel in The Netherlands. *Epidemiology & Infection*, 138(5), 743-755.

Muller, J., Szklarczyk, D., Julien, P., Letunic, I., Roth, A., Kuhn, M., Powell, S., von Mering, C., Doerks, T., Jensen, L.J., Bork, P. (2010). eggNOG v2.0: extending the evolutionary genealogy of genes with enhanced non-supervised orthologous groups, species and functional annotations. *Nucleic Acids Research*, 38(suppl_1), D190-195.

Munk, P., Andersen, V.D., de Knegt, L., Jensen, M.S., Knudsen, B.E., Lukjancenko, O., Mordhorst, H., Clasen, J., Agersø, Y., Folkesson, A., Pamp, S.J., Vigre, H., Aarestrup, F.M. (2017). A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds. *Journal of Antimicrobial Chemotherapy*, 72(2), 385-392.

Munk, P., Knudsen, B.E., Lukjacenko, O., Duarte, A.S.R., Van Gompel, L., Luiken, R.E.C., Smit, L.A.M., Schmitt, H., Garcia, A.D., Hansen, R.B., Petersen, T.N., Bossers, A., Ruppé, E., Graveland, H., van Essen, A., Gonzalez-Zorn, B., Moyano, G., Sanders, P., Chauvin, C., David, J., Battisti, A., Caprioli, A., Dewulf, J., Blaha, T., Wadepohl, K., Brandt, M., Wasyl, D., Skarzyńska, M., Zajac, M., Daskalov, H., Saatkamp, H.W., Stärk, K.D.C., Lund, O., Hald, T., Pamp, S.J., Vigre, H., Heederik, D., Wagenaar, J.A., Mevius, D., Aarestrup, F.M., (2018). Abundance and diversity of the faecal resistome in slaughter pigs and broilers in nine European countries. *Nature Microbiology*, 3, 898-908.

Naas, T., Oueslati, S., Bonnin, R.A., Dabos, M.L., Zavala, A., Dortet, L., Retailleau, P., Iorga, B.I. (2017). Beta-lactamase database (BLDB)–structure and function. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 32(1), 917-919.

Nacke, H., Engelhaupt, M., Brady, S., Fischer, C., Tautzt, J., Daniel, R. (2012). Identification and characterization of novel cellulolytic and hemicellulolytic genes and enzymes derived from german grassland soil metagenomes. *Biotechnology. Letters*, 34, 663-675.

Naik, O.A., Shashidhar, R., Rath, D., Bandekar, J.R., Rath, A. Characterization of Multiple Antibiotic Resistance of Culturable Microorganisms and Metagenomic Analysis of Total Microbial Diversity of Marine Fish Sold in Retail Shops in Mumbai, India. *Environmental Science and Pollution Research*, 2018, 25, 6228-6239.

Namiki, T., Hachiya, T., Tanaka, H., Sakakibara, Y. (2012). MetaVelvet: an extension of velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Research*, 40, e155.

Nhung, N.T., Cuong, N.V., Thwaites, G., Carrique-Mas, J. (2016). Antimicrobial usage and antimicrobial resistance in animal production in Southeast Asia: a review. *Antibiotics*, 5(4), 37.

Ni, J., Yan, Q., Yu, Y. (2013). How much metagenomic sequencing is enough to achieve a given goal?. *Scientific Reports*, 3(1), 1-7.

Nielsen, H.B., Almeida, M., Juncker, A.S., Rasmussen, S., Li, J., Sunagawa, S., Plichta, D.R., Gautier, L., Pedersen, A.G., Le Chatelier, E., Pelletier, E., Bonde, I., Nielsen, T., Manichanh, C., Arumugam, M., Batto, J.M., dos Santos, M.B.Q., Blom, N., Borruel, N., Burgdorf, K.S., Boumezbeur, F., Casellas, F., Doré, J., Dworzynski, P., Guarner, F., Hansen, T., Hildebrand, F., Kaas, R.S., Kennedy, S., Kristiansen, K., Kultima, J.R., Léonard, P., Levenez, F., Lund, O., Moumen, B., Le Paslier, D., Pons, N., Pedersen, O., Prifti, E., Qin, J., Raes, J., Sørensen, S., Tap, J., Tims, S., Ussery, D.W., Yamada, T., MetaHIT Consortium, Renault, P., Sicheritz-Ponten, T., Bork, P., Wang, J., Brunak, S., Ehrlich, S.D. (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature Biotechnology*, 32(8), 822-828.

Noble, W.S. (2009). How does multiple testing correction work?. *Nature Biotechnology*, 27(12), 1135-1137.

Noguchi, H., Taniguchi, T., Itoh, T. (2008). MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Research*, 15(6), 387-396.

Noyes, N.R., Yang, X., Linke, L.M., Magnuson, R.J., Cook, S.R., Zaheer, R., Yang, H., Woerner, D.R., Geornaras, I., McArt, J.A., Gow, S.P., Ruiz, J., Jones, K.L., Boucher, C.A., McAllister, T.A., Belk K.E., Morley, P. S. (2016a). Characterization of the resistome in manure, soil and wastewater from dairy and beef production systems. Scientific Reports, 6(1), 1-12.

Noyes, N.R., Yang, X., Linke, L.M., Magnuson, R.J., Dettenwanger, A., Cook, S., Geornaras, I., Woerner, D.E., Gow, S.P., McAllister, T.A., Yang, H., Ruiz, J., Jones, K., Boucher, C.A., Morley, P.S., Belk, K.E. (2016b). Resistome diversity in cattle and the environment decreases during beef production. *Elife*, 5, e13195.

Nurk, S., Meleshko, D., Korobeynikov, A. Pevzner, P.A. (2017). MetaSPAdes: a new versatile metagenomic assembler. *Genome Research*, 27, 824-834.

Nychas, G.J.E., Skandamis, P.N., Tassou, C.C., Koutsoumanis, K.P. (2008). Meat spoilage during distribution. *Meat science*, 78(1-2), 77-89.

Obe, T., Nannapaneni, R., Schilling, W., Zhang, L., McDaniel, C., Kiess, A. (2020). Prevalence of Salmonella enterica on poultry processing equipment after completion of sanitization procedures. *Poultry Science*, 99(9), 4539-4548.

O'Dwyer, J., Hynds, P.D., Byrne, K.A., Ryan, M.P., Adley, C.C. (2018). Development of a hierarchical model for predicting microbiological contamination of private groundwater supplies in a geologically heterogeneous region. *Environmental Pollution*, 237, 329-338.

Oksanen, J., Blanchet, G.F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Wagner, H. (2020). Vegan: Community Ecology Package. R package version 2.5-6.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., Suggests, M.A.S.S. (2007). The vegan package. *Community Ecology Package*, 10(631-637), 719.

Oniciuc, E.A., Likotrafiti, E., Alvarez-Molina, A., Prieto, M., Santos, J.A., Alvarez-Ordóñez, A. (2018). The present and future of whole genome sequencing (WGS) and whole metagenome sequencing (WMS) for surveillance of antimicrobial resistant microorganisms and antimicrobial resistance genes across the food chain. *Genes*, 9(5), 268.

Oulas, A., Pavloudi, C., Polymenakou, P., Pavlopoulos, G.A., Papanikolaou, N., Kotoulas, G., Arvanitidis, C., Iliopoulos, L. (2015). Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinformatics and Biology Insights*, 9, BBI-S12462.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E.D., Gerdes, S., Glass, E.M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A.C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G.D., Rodionov, D.A., Rückert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., Vonstein, V. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Research*, 33(17), 5691-702.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, 42(D1), D206–D214.

Pacholewicz, E., Liakopoulos, A., Swart, A., Gortemaker, B., Dierikx, C., Havelaar, A., Schmitt, H. (2015). Reduction of extended-spectrum-β-lactamase-and AmpC-β-lactamase-producing Escherichia coli through processing in two broiler chicken slaughterhouses. *International Journal of Food Microbiology*, 215, 57-63.

Pacholewicz, E., Swart, A., Wagenaar, J.A., Lipman, L.J., Havelaar, A.H. (2016). Explanatory variables associated with Campylobacter and Escherichia coli concentrations on broiler chicken carcasses during processing in two slaughterhouses. *Journal of Food Protection*, 79(12), 2038-2047.

Pareek, C.S., Smoczynsky, R., Tretyn, A. (2011). Sequencing tech- nologies and genome sequencing. *Journal of Applied Genetics*, 52(4), 413-435.

Parks, D.H., Tyson, G.W., Hugenholtz, P., Beiko, R.G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics*, 30(21), 3123-3124.

Pasolli, E., Truong, D.T., Malik, F., Waldron, L. Segata, N. (2016). Machine learning meta-analysis of large metagenomic datasets: tools and biological insights. *PloS Computational Biology*, 12(7), e1004977.

Paszkiewicz, K., & Studholme, D.J. (2010). De novo assembly of short sequence reads. *Briefings in Bioinformatics*, 11(5), 457-72.

Pati, A., Heath, L.S., Kyrpides, N.C., Ivanova, N. (2011). ClaMS: a classifier for metagenomic sequences. *Standards in Genomic Sciences*, 5(2), 248-53.

Patil, K.R., Roune, L., McHardy, A.C. (2012). The PhyloPythiaS web server for taxonomic assignment of metagenome sequences. *PloS One*, 2012, 7(6), e38581.

Pehrsson, E.C., Tsukayama, P., Patel, S., Mejia-Bautista, M., Sosa-Soto, G., Navarrete, K.M., Calderon, M., Cabrera, L., Hoyos-Arango, W., Bertoli, M.T., Berg, D.E., Gilman, R.H., Dantas, G. (2016). Interconnected microbiomes and resistomes in low-income human habitats. *Nature*, 533, 212-216.

Peng, Y., Leung, H.C., Yiu, S.M., Chin, F.Y. (2011). Meta-IDBA: a de novo assembler for metagenomic data. *Bioinformatics*, 27(13), i94-i101.

Perrin-Guyomard, A., Jouy, E., Urban, D., Chauvin, C., Granier, S.A., Mourand, G., Chevance, A., Adam, C., Moulin, G. Kempf, I. (2020). Decrease in fluoroquinolone use in French poultry and pig production and changes in resistance among *E. coli* and *Campylobacter*. *Veterinary Microbiology*, 243, 108637.

Pesciaroli, M., Orsini, S., Marchi, L., Massacci, F.R., Tofani, S., Maresca, C., Scoccia, E., Filippini, G., Magistrali, C.F., Pezzotti, G. (2018). The impact of the farming system on the antibiotic resistance of commensal bacteria in the intestine of broiler. In *XVIII Congresso Nazionale SI Di. LV, Perugia (PG), Italia, 7-9 Novembre 2018* (pp. 45-46). Società Italiana di Diagnostica di Laboratorio Veterinaria (SIDiLV).

Petersen, T.N., Lukjancenko, O., Thomsen, M.C.F., Maddalena Sperotto, M., Lund, O., Møller Aarestrup, F., Sicheritz-Pontén, T. (2017). MGmapper: reference-based mapping and taxonomy annotation of metagenomics sequence reads. *PloS One*, 12(5), e0176469.

Petersen, T.N., Lukjancenko, O., Thomsen, M.C.F., Sperotto, M.M., Lund, O., Aarestrup, F.M., Sicheritz-Pontén, T. (2017). Mgmapper: Reference based mapping and taxonomy annotation of metagenomics sequence reads. *PloS One*, 12(5), e0176469.

Phan, G., Picard, M., Broutin, I. (2015). Focus on the outer membrane factor OprM, the forgotten player from efflux pumps assemblies. *Antibiotics*, 4(4), 544-566.

Pisacane, V., Callegari, M. L., Puglisi, E., Dallolio, G., Rebecchi, A. (2015). Microbial analyses of traditional Italian salami reveal microorganisms transfer from the natural casing to the meat matrix. *International Journal of Food Microbiology*, 207, 57-65.

Pitta, D.W., Dou, Z., Kumar, S., Indugu, N., Toth, J.D., Vecchiarelli, B., Bhukya, B. (2016). Metagenomic Evidence of the Prevalence and Distribution Patterns of Antimicrobial Resistance Genes in Dairy Agroecosystems. *Foodborne Pathogens and Disease*, 13, 296-302.

Podolsky, S.H. (2017). Historical perspective on the rise and fall and rise of antibiotics and human weight gain. *Annals of Internal Medicine*, 166(2), 133-138.

Poudel, A., Hathcock, T., Butaye, P., Kang, Y., Price, S., Macklin, K., Walz, P., Cattley, R., Kalalah, A., Adekanmbi, F., Wang, C. (2019). Multidrug-resistant *Escherichia coli*, *Klebsiella pneumoniae* and *Staphylococcus* spp. in houseflies and blowflies from farms and their environmental settings. *International Journal of Environmental Research and Public Health*, 16(19), 3583.

Powell, S., Forslund, K., Szklarczyk, D., Trachana, K., Roth, A., Huerta-Cepas, J., Gabaldón, T., Rattei, T., Creevey, C., Kuhn, M., Jensen, L.J., von Mering, C., Bork, P. (2014). eggNOG v4. 0: nested orthology inference across 3686 organisms. *Nucleic Acids Research*, 42(D1), D231-D239.

Prescott, S.L. (2017). History of medicine: Origin of the term microbiome and why it matters. *Human Microbiome Journal*, 4, 24-25.

Pruitt, K.D., Tatusova, T., Maglott, D.R. (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, 33, 501-504.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.M., Hansen, T., Paslier, D.L., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Doré, J., Guarner, F., kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, 464, 59-65.

Quigley, L., O'Sullivan, D. J., Daly, D., O'Sullivan, O., Burdikova, Z., Vana, R., Beresford, T.P., Ross, R.P., Fitzgerald, G.F., McSweeney, P.L.H., Giblin, L., Sheehan, J.J., Cotter, P.D. (2016). Thermus and the pink discoloration defect in cheese. *MSystems*, 1(3), e00023-16.

Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J., Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature biotechnology*, 35(9), 833-844.

Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J., Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, 35(9), 833-844.

Rabello, R.F., Bonelli, R.R., Penna, B.A., Albuquerque, J.P., Souza, R.M., Cerqueira, A.M. (2020). Antimicrobial resistance in farm animals in Brazil: An update overview. *Animals*, 10(4), 552.

Raffatellu, M. (2018). Learning from bacterial competition in the host to develop antimicrobials. *Nature Medicine*, 24(8), 1097-1103.

Ranjitkar, S., Lawley, B., Tannock, G., Engberg, R. M. (2016). Bacterial succession in the broiler gastrointestinal tract. *Applied and Environmental Microbiology*, 2016, *82*(8), 2399-2410.

Rasschaert, G., De Zutter, L., Herman, L., Heyndrickx, M. (2020). *Campylobacter* contamination of broilers: The role of transport and slaughterhouse. *International journal of Food Microbiology*, 322, 108564.

Regulation (EU) 2019/4 of the European Parliament and of the council of 11 December 2018 on the manufacture, placing on the market and use of medicated feed, amending Regulation (EC) No 183/2005 of the European Parliament and of the Council and repealing Council Directive 90/167/EEC. *Official Journal of the European Union*, L 4/1, 7.01.2019.

Reuter, J.A., Spacek, D.V., Snyder, M.P. (2015). High-throughput sequencing technologies. *Molecular Cell*, 58,586-97.

Rho, M., Tang, H., Ye, Y. (2010). FragGeneScan: Predicting genes in short and error-prone reads. *Nucleic Acids Research*, 2010, 38(20), e191-e191.

Robinson, M.D., McCarthy, D.J., Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 139- 40.

Roca, I., Akova, M., Baquero, F., Carlet, J., Cavaleri, M., Coenen, S., Cohen, J., Findlay, D., Gyssens, I., Heuer, O.E., Kahlmeter, G., Kruse, H., Laxminarayan, R., Liébana, E., López-Cerero, L., MacGowan, A., Martins, M., Rodríguez-Baño, J., Rolain, J.M., Segovia, C., Sigauque, B., Tacconelli, E., Wellington, E., Vila, J. (2015). The global threat of antimicrobial resistance: science for intervention. New Microbes and New Infections, 6, 22-29.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlén, M., and Nyrén, P.L. (1996). Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry*, 242(1), 84-89.

Roto, S.M., Rubinelli, P.M., Ricke, S.C. (2015). An introduction to the avian gut microbiota and the effects of yeast-based prebiotic-type compounds as potential feed additives. *Frontiers in Veterinary Science*, 2, 28.

Rouger, A., Remenant, B., Prévost, H., Zagorec, M. (2017a). A method to isolate bacterial communities and characterize ecosystems from food products: Validation and utilization in as a reproducible chicken meat model. *International Journal of Food Microbiology*, 247, 38-47.

Rouger, A., Tresse, O., Zagorec, M. (2017b). Bacterial contaminants of poultry meat: sources, species, and dynamics. *Microorganisms*, 5(3), 50.

Royden, A., Wedley, A., Merga, J.Y., Rushton, S., Hald, B., Humphrey, T., Williams, N.J. (2016). A role for flies (Diptera) in the transmission of *Campylobacter* to broilers?. *Epidemiology & Infection*, 144(15), 3326-3334.

Rozman, V., Matijašić, B.B., Možina, S.S. (2019). Antimicrobial resistance of common zoonotic bacteria in the food chain: An emerging threat. *Antimicrobial Resistance a Global Threat*.

Ruppé, E., Ghozlane, A., Tap, J., Pons, N., Alvarez, A.S., Maziers, N., Cuesta, T., Hernando-Amado, S., Clares, I., Martínez, J.L., Coque, T.M., Baquero, F., Lanza, V.F., Máiz, L., Goulenok, T., de Lastours, V., Amor, N., Fantin, B., Wieder, I., Andremont, A., van Schaik, W., Rogers, M., Zhang, X., Willems, R.J.L., de Brevern, A.G., Batto, J.M., Blottière, H.M., Léonard, P., Léjard, V., Letur, A., Levenez, F., Weiszer, K., Haimet, F., Doré, J., Kennedy, S.P., Ehrlich, S.D. (2019). Prediction of the intestinal resistome by a three-dimensional structure-based method. *Nature Microbiology*, 4, 112-123.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J.E., Li, K., Kravitz, S., Heidelberg, J.F., Utterback, T., Rogers, Y.H., Falcón, L.I., Souza, V., Bonilla-Rosso, G., Eguiarte, L.E., Karl, D.M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M.R., Strausberg, R.L., Nealson, K., Friedman, R., Frazier, M., Venter, J.C. (2007). The Sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. *PloS Biology*, 5(3), e77.

Sabino, Y.N.V., Santana, M.F., Oyama, L.B., Santos, F.G., Moreira, A.J.S., Huws, S.A., Mantovani, H.C. (2019). Characterization of antibiotic resistance genes in the species of the rumen microbiota. *Nature Communications*, 10(1), 1-11.

Sala, C., Mordhorst, H., Grützke, J., Brinkmann, A., Petersen, T.N., Poulsen, C., Cotter, P.D., Crispie, F., Ellis, R.J., Castellani, G., Amid, C., Hakhverdyan, M., Le Guyader, S., Manfreda, G., Mossong, J., Nitsche, A., Ragimbeau, C., Schaeffer, J., Schlundt, J., Tay, M.Y.F., Aarestrup, M.F., Hendriksen, R.S., Pamp, S.J., De Cesare, A. (2020). Metagenomics-based proficiency test of smoked salmon spiked with a mock community. *Microorganisms*, 8(12), 1861.

Salipante, S.J., Kawashima, T., Rosenthal, C., Hoogestraat, D.R., Cummings, L.A., Sengupta, D.J., Harkins, T.T., Cookson, B.T., Hoffman, N.G. Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. (2014). *Applied and Environmental Microbiology*, 80(24), 7583-91.

Samapundo, S., de Baenst, I., Aerts, M., Cnockaert, M., Devlieghere, F., Van Damme, P. (2019). Tracking the sources of psychrotrophic bacteria contaminating chicken cuts during processing. *Food Microbiology*, 81, 40-50.

Sanger, F., Nicklen, S., Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12), 5463-5467.

Sapkota, A.R., Lefferts, L.Y., McKenzie, S., Walker, P. (2007). What do we feed to food-production animals? A review of animal feed ingredients and their potential impacts on human health. *Environmental Health Perspectives*, 115(5), 663-670.

Savin, M., Bierbaum, G., Hammerl, J. A., Heinemann, C., Parcina, M., Sib, E., Voigt, A., Kreyenschmidt, J. (2020). Isolation and characterization of ESKAPE-bacteria and ESBL-producing *E. coli* from waste-and process water of German poultry slaughterhouses. *Applied and Environmental Microbiology*, 86.

Savin, M., Bierbaum, G., Hammerl, J.A., Heinemann, C., Parcina, M., Sib, E., Voigt, A., Kreyenschmidt, J. (2020). Isolation and characterization of ESKAPE-bacteria and ESBL-producing *E. coli* from waste-and process water of German poultry slaughterhouses. *Applied and Environmental Microbiology*.

Schadt, E.E., Turner, S., Kasarskis, A. (2010). A window in third-generation sequencing. *Human Molecular Genetics*, 19(R2), R227-R240.

Schloss, P.D., & Handelsman, J. (2008). A statistical toolbox for metagenomics: assessing functional diversity in microbial communities. *BMC Bioinformatics*, 23, 9(1), 1-15.

Schmieder, R., & Edwards, R. (2011a) Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27, 863-864.

Schmieder, R., & Edwards, R. (2011b). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PloS One*, 6, e17288.

Schmieder, R., & Edwards, R. (2012). Insights into antibiotic resistance through metagenomic approaches. *Future Microbiology*, 7(1), 73-89.

Seeger, M.A., Schiefner, A., Eicher, T., Verrey, F., Diederichs, K., Pos, K.M. (2006). Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science*, 313(5791), 1295-1298.

Sender, R., Fuchs, S., Milo, R. (2016). Revised estimates for the number of human and bacteria cells in the body. *PLoS Biology*, 14(8), e1002533.

Sharpton, T.J. (2014). An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science*, 5, 209.

Sharpton, T.J., Riesenfeld, S.J., Kembel, S.W., Ladau, J., O'Dwyer, J.P., Green, J.L., Eisen, J.A., Pollard, K.S. (2011). PhylOTU: a high-throughput procedure quantifies microbial community diversity and resolves novel taxa from metagenomic data. *PloS Computational Biology*, 20, 7(1), e1001061.

Shaufi, M.A.M., Sieo, C.C., Chong, C.W., Gan, H.M., Ho, Y.W. (2015). Deciphering chicken gut microbial dynamics based on high-throughput 16S rRNA metagenomics analyses. *Gut Pathogens*, 7(1), 1-12.

Simonsen, B. (1995). Microbiological criteria for poultry products. In: *Processing of Poultry*. Springer, Boston, MA, p. 221-250.

Singh, K.M., Jakhesara, S.J., Koringa, P.G., Rank, D.N., Joshi, C.G. (2012). Metagenomic analysis of virulence-associated and antibiotic resistance genes of microbes in rumen of Indian buffalo (Bubalus bubalis). *Gene*, 507(2), 146-151.

Smith, D.L., Harris, A.D., Johnson, J.A., Silbergeld, E.K., Morris, J.G. (2002). Animal antibiotic use has an early but important impact on the emergence of antibiotic resistance in human commensal bacteria. *Proceedings of the National Academy of Sciences U. S. A.*, 99(9), 6434-6439.

Smith, E., Lichten, C.A., Taylor, J., MacLure, C., Lepetit, L., Harte, E., Martin, A., Ghiga, I., Pitchforth, E., Sussex, J., Dujso E., Littmann, J. (2016). Evaluation of the EC action plan against the rising threats from antimicrobial resistance. *European Commission, Brussels*.

Souza, A.I., Saraiva, M.M., Casas, M.R., Oliveira, G.M., Cardozo, M.V., Benevides, V.P., Barbosa, F.O., Freitas Neto, O.C., Almeida, A.M., Berchieri, A. (2020). High occurrence of β-lactamase-producing Salmonella Heidelberg from poultry origin. *PloS One*, 15(3), e0230676.

Spielmeyer, A., Ahlborn, J., Hamscher, G. (2014). Simultaneous determination of 14 sulfonamides and tetracyclines in biogas plants by liquid-liquid-extraction and liquid chromatography tandem mass spectrometry. *Analytical and Bioanalytical Chemistry*, 406(11), 2513-2524.

Srivastava, A., Singhal, N., Goel, M., Virdi, J.S., Kumar, M. (2014). CBMAR: A comprehensive β-lactamase molecular annotation resource. *Database*, 2014.

Stellato, G., La Storia, A., De Filippis, F., Borriello, G., Villani, F., Ercolini, D. (2016). Overlap of spoilage-associated microbiota between meat and the meat processing environment in small-scale and large-scale retail distributions. *Applied and Environmental Microbiology*, 82(13), 4045-4054.

Sun, J., Deng, Z., Yan, A. (2014). Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. *Biochemical and Biophysical Research Communications*, 453(2), 254-267.

Sun, S., Chen, J., Li, W., Altintas, I., Lin, A., Peltier, S., Stocks, K., Allen, E.E., Ellisman, M., Grethe, J., Wooley, J. (2010). Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. *Nucleic Acids Research*, 39(suppl_1), D546-551.

Talon, R., & Leroy, S. (2011). Diversity and safety hazards of bacteria involved in meat fermentations. *Meat Science*, 89(3), 303-309.

Tang, J., Y. Bu, X., Zhang, X.X., Huang, K., He, X., Ye, L., Shan, Z., Ren, H. (2016). Metagenomic analysis of bacterial community composition and antibiotic resistance genes in a wastewater treatment plant and its receiving surface water**.** *Ecotoxicology and Environmental Safety*, 132, 260-269.

Tang, W., Bischof, J., Desai, N., Mahadik, K., Gerlach, W., Harrison, T., Wilke, A., Meyer, F. (2014). Workload characterization for MG-RAST metagenomic data analytics service in the cloud. In 2014 *IEEE International Conference on Big Data,* 56-63.

Tanner, M.A., Goebel, B.M., Dojka, M.A., Pace, N.R. (1998). Specific ribosomal DNA sequences from diverse environmental settings correlate with experimental contaminants. *Applied and Environmental Microbiology,* 64, 3110-3113.

Tanner, W.D., VanDerslice, J.A., Goel, R.K., Leecaster, M.K., Fisher, M.A., Olstadt, J., Gurley, C.M., Morris, A.G., Seely, K.A., Chapman, L., Korando, M., Shabazz, K.A., Stadsholt, A., VanDeVelde, J., Braun-Howland, E., Minihane, C., Higgins, P.J., Deras, M., Jaber, O., Jette, D., Gundlapalli, A.V. (2019). Multi-state study of Enterobacteriaceae harboring extended-spectrum beta-lactamase and carbapenemase genes in US drinking water. *Scientific Reports*, 9(1), 1-8.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D, Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics,* 4, 41.

Team, R.D.C. (2008). R: A Language and Environment for Statistical Computing. *Vienna, Austria: R Foundation for Statistical Computing*.

Teeling, H., Waldmann, J., Lombardot, T., Bauer, M., Glockner, F.O. (2004). TETRA: a web-service and a stand-alone program for the analysis and comparison of tetra-nucleotide usage patterns in DNA sequences. *BMC Bioinformatics*, 5(1), 1-7.

Thai, Q.K., Bös, F., Pleiss, J. (2009). The lactamase engineering database: A critical survey of TEM sequences in public databases. *BMC Genomics* ,10(1), 1-9.

Thomas, C J., & McMeekin, T.A. (1980). Contamination of broiler carcass skin during commercial processing procedures: an electron microscopic study. *Applied and Environmental Microbiology*, 40(1), 133-144.

Thomas, T., Gilbert, J., Meyer, F. (2012). Metagenomics-a guide from sampling to data analysis. *Microbial Informatics and Experimentation*, 2(1), 1-12.

Tong, P., Ji, X., Chen, L., Liu, J., Xu, L., Zhu, L., Zhou, W., Liu, G., Wang, S., Guo, X., Feng, S., Sun, Y. (2017). Metagenome analysis of antibiotic resistance genes in fecal microbiota of chickens. *Agri Gene*, 5, 1-6.

Torres, G.J., Piquer, F.J., Algarra, L., de Frutos, C., Sobrino, O.J. (2011). The prevalence of Salmonella enterica in Spanish feed mills and potential feed-related risk factors for contamination. *Preventive Veterinary Medicine*, 98(2-3), 81-87.

Treangen, T.J., Koren, S., Sommer, D.D., Liu, B., Astrovskaya, I., Ondov, B., Darling, A.E., Phillippy, A.M., Pop, M. (2013). MetAMOS: a modular and open source metagenomic assembly and analysis pipeline. *Genome Biology*, 14(1), 1-20.

Truong, D.T., Tett, A., Pasolli, E., Huttenhower, C., Segata, N. (2017). Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Research,* 27(4), 626-38.

Van Boeckel, T.P., Brower, C., Gilbert, M., Grenfell, B.T., Levin, S.A., Robinson, T.P., Teillant, A., Laxminarayan, R. (2015). Global trends in antimicrobial use in food animals. *Proceedings of the National Academy of Sciences of the United States of America*, 112(18), 5649-5654.

Van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends in Genetics,* 30(9), 418-426.

Van Gompel, L., Dohmen, W., Luiken, R. E., Bouwknegt, M., Heres, L., Van Heijnsbergen, E., Jongerius-Gortemaker, B.G.M., Scherpenisse, P., Greve, G.D., Tersteeg-Zijderveld, M.H.G., Wadepohl, K., Ribeiro Duarte, A.S., Munoz-Gomez, V., Fischer, J., Skarzynska, M., Wasyl, D., Wagenaar, J.A., Urlings, B.A.P., Dorado-Garcıa, A., Wouters, I.M., Heederik, D.J.J., Schmitt, H., Smit, L. A. (2020). Occupational exposure and carriage of antimicrobial resistance genes (tetW, ermB) in pig slaughterhouse workers. *Annals of work exposures and health*, 64(2), 125-137.

Van Schothorst, M., Zwietering, M.H., Ross, T., Buchanan, R.L., Cole, M.B. (2009). Relating microbiological criteria to food safety objectives and performance objectives. *Food Control*, 20(11), 967-979.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-tillson, H., Pfannkoch, C., Rogers, Y.U., Smith, H.O. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304, 66-74.

Waldron, L. (2018). Data and statistical methods to analyze the human microbiome. *Msystems*, 3(2), e00194-17.

Wallace, J.C., Port, J.A., Smith, M.N., Faustman, E.M. (2017). FARME DB: a functional antibiotic resistance element database. *Database*, 2017, baw165.

Walsh, A.M., Crispie, F., Claesson, M.J. Cotter, P.D. (2017). Translating Omics to Food Microbiology. *Annual Review of Food Science and Technology*, 8, 113-134.

Walsh, A.M., Crispie, F., O'Sullivan, O., Finnegan, L., Claesson, M.J., Cotter, P.D. (2018). Species classifier choice is a key consideration when analysing low-complexity food microbiome data. *Microbiome*, 6(1), 1-15.

Walsh, A.M., Macori, G., Kilcawley, K.N., Cotter, P.D. (2020). Meta-analysis of cheese microbiomes highlights contributions to multiple aspects of quality. *Nature Food*, 1(8), 500-510.

Wang, Y., Hu, Y., Cao, J., Bi, Y., Lv, N., Liu, F., Liang, S., Shi, Y., Jiao, X., Fu, G.G., Zhu, B. (2019). Antibiotic resistance gene reservoir in live poultry markets. *Journal of Infection*, 78(6), 445-453.

Wang, Y., Hu, Y., Liu, F., Cao, J., Lv, N., Zhu, B., Zhang, G., Gao, G.F. (2020). Integrated metagenomic and metatranscriptomic profiling reveals differentially expressed resistomes in human, chicken, and pig gut microbiomes. *Environment International*, 138, 105649.

Wang, Y., Na, L., Liu, F., Liu, W.J., Gao, G.F. (2021). More diversified antibiotic resistance genes in chickens and workers of the live poultry markets. *Environment International*, 153, 106534.

West, A.M., Teska, P.J., Lineback, C.B., Oliver, H.F. (2018). Strain, disinfectant, concentration, and contact time quantitatively impact disinfectant efficacy. *Antimicrobial Resistance & Infection Control*, 7(1), 1-8.

Whipps, J., Lewis, K., Cooke, R. (1988). Mycoparasitism and plant disease control. In: Burge M, editor. Fungi in Biological Control Systems. Manchester University Press; p. 161-187.

White, J.R., Nagarajan, N., Pop, M. (2009). Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Computational Biology*, 5, e1000352.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Computational Biology*, 13(6), e1005595.

Wilke, A., Bischof, J., Gerlach, W., Glass, E., Harrison, T., Keegan, K. P., Paczian, T., Trimble, W.L., Bagchi, S., Grama, A., Chaterji, S., Meyer, F. (2016). The MG-RAST metagenomics database and portal in 2015. *Nucleic Acids Research*, 44(D1), D590-D594.

Wilke, A., Gerlach, W., Harrison, T., Paczian, T., Trimble, W. L., Meyer, F. (2017). MG-RAST manual for version 4, revision 3; 2017.

Wolfe, B.E., Button, J.E., Santarelli, M., Dutton, R.J. (2014). Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. *Cell*, 158,422-33.

Wommack, K.E., Bhavsar, J., Ravel, J. (2008). Metagenomics: read length matters. *Applied and Environmental Microbiology*, 2008, 74 (5).

Wood, D.E., & Salzberg, S.L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, 15(3), 1-12.

World Health Organization Antibiotic resistance. URL. https://www.who.int/news-room/fact-sheets/detail/antibiotic-resistance, 2018.

Woyke, T., Teeling, H., Ivanova, N.N., Huntemann, M., Richter, M., Gloeckner, F.O., Boffetti, D., Anderson, I.J., Barry, K.W., Shapiro, H.J., Szeto, E., Kyrpides, N.C., Mussmann, M., Amann, R., Bergin, C., Ruehland, C., Rubin, E.M., Dubilier, N. (2006). Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature*, 443, 950-955.

Wright, G.D. (2007). The antibiotic resistome: the nexus of chemical and genetic diversity. *Nature Reviews Microbiology*, 5(3), 175-186.

Wrighton, K.C., Thomas, B.C., Sharon, I., Miller, C.S., Castelle, C.J., VerBerkmoes, N.C., Wilkins, M.J., Hettich, R.L., Lipton, M.S., Williams, K.H., Long, P.E., Banfield, J.F. (2012). Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science*, 337, 1661-1665.

Wu, L.H., Lu, Z.M., Zhang, X.J., Wang, Z.M., Yu, Y.J., Shi, J.S., Xu, Z.H. (2017). Metagenomics reveals flavour metabolic network of cereal vinegar microbiota. *Food Microbiology*, 62, 23-31.

Wylezich, C., Papa, A., Beer, M., Höper, D.A. (2018). Versatile sample processing workflow for metagenomic pathogen detection. *Scientific Reports*, 8, 13108.

Xiong, W., Wang, Y., Sun, Y., Ma, L., Zeng, Q., Jiang, X., Li, A., Zeng, Z., Zhang, T. (2018). Antibiotic-mediated changes in the fecal microbiome of broiler chickens define the incidence of antibiotic resistance genes. *Microbiome*, 6(1), 1-11.

Xu, J., Sangthong, R., McNeil, E., Tang, R., Chongsuvivatwong, V. (2020). Antibiotic use in chicken farms in northwestern China. *Antimicrobial Resistance & Infection Control*, 9(1), 1-9.

Yan, Q., Wi, Y.M., Thoendel, M.J., Raval, Y.S., Greenwood-Quaintance, K.E., Abdel, M.P., Jeraldo, P.R., Chia, N., Patel, R. (2019). Evaluation of the CosmosID bioinformatics platform for prosthetic joint-associated sonicate fluid shotgun metagenomic data analysis. *Journal of Clinical Microbiology*, 57(2), e01182-18.

Yan, Q., Wi, Y.M., Thoendel, M.J., Raval, Y.S., Greenwood-Quaintance, K.E., Abdel, M.P., Jeraldo, P.R., Chia, N., Patel, R. (2019). Evaluation of the CosmosID bioinformatics platform for prosthetic joint-associated sonicate fluid shotgun metagenomic data analysis. *Journal of Clinical Microbiology*, 57(2), e01182-18.

Yang, H., Dettman, B., Beam, J., Mix, C., Jiang, X. (2006). Occurrence of ceftriaxone-resistant commensal bacteria on a dairy farm and a poultry farm. *Canadian Journal of Microbiology*, 52(10), 942-950.

Yang, J., Tong, C., Xiao, D., Xie, L., Zhao, R., Huo, Z., Tang, Z., Hao, J., Zeng, Z., Xiong, W. (2021). A reference gene catalogue of chicken gut antibiotic resistomes, 2021.

Yang, X., Noyes, N.R., Doster, E., Martin, J.N., Linke, L.M., Magnuson, R.J., Yang, H., Geornaras, I., Woerner, D.R., Jones, K.L., Ruiz, J. Boucher, C., Morley, P.S., Belk, K. (2016). Use of metagenomic shotgun sequencing technology to detect foodborne pathogens within the microbiome of the beef production chain. *Applied and Environmental Microbiology*, 82, 2433-43.

Yang, Y., Jiang, X., Chai, B., Ma, L., Li, B., Zhang, A., Cole, J.R., Tiedje, J.M., Zhang, T. (2016b). ARGs-OAP: Online analysis pipeline for antibiotic resistance genes detection from metagenomic data using an integrated structured ARG-database. *Bioinformatics*, 32(15), 2346-2351.

Yang, Y., Zhou, R., Chen, B., Zhang, T., Hu, L., Zou, S. (2018). Characterization of airborne antibiotic resistance genes from typical bioaerosol emission sources in the urban environment using metagenomic approach. *Chemosphere*, 213, 463-471.

Yewale, P.P., Lokhande, K.B., Sridhar, A., Vaishnav, M., Khan, F.A., Mandal, A., Swamy, K. V., Jass, J., Nawani, N. (2020). Molecular profiling of multidrug-resistant river water isolates: insights into resistance mechanism and potential inhibitors. *Environmental Science and Pollution Research, 27*(22), 27279-27292.

Yin, X., Jiang, X.T., Chai, B., Li, L., Yang, Y., Cole, J.R., Tiedje, J.M., Zhang, T. (2018). ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. *Bioinformatics*, 34, 2263-2270.

Yozwiak, N.L., Skewes-Cox, P., Stenglein, M.D., Balmaseda, A., Harris, E., and DeRisi, J.L. (2012). Virus identification in unknown tropical febrile illness cases using deep sequencing. *PloS Neglected Tropical Diseases*, 6, e1485.

Yuan, S., Cohen, D.B., Ravel, J., Abdo, Z., Forney, L.J. (2012). Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PloS One*, 7, e33865.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V. (2012). Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, 67(11), 2640-2644.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V. (2012). Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, 67(11), 2640-2644.

Zerbino, D.R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, 18(5), 821-829.

Zhang, X.X., Zhang, T. (2011). Occurrence, abundance, and diversity of tetracycline resistance genes in 15 sewage treatment plants across China and other global locations. *Environmental Science & Technology*, 45(7), 2598-2604.

Zhao, Y., Yang, Q.E., Zhou, X., Wang, F. H., Muurinen, J., Virta, M.P., Brandt, K.K., Zhu, Y.G. (2021). Antibiotic resistome in the livestock and aquaculture industries: Status and solutions. *Critical Reviews in Environmental Science and Technology*, 51(19), 2159-2196.

Zhu, A., Ibrahim, J.G., Love, M.I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics*, 35(12), 2084-2092.

Zurek, L., & Ghosh, A. (2014). Insects represent a link between food animal farms and the urban environment for antibiotic resistance traits. *Applied and Environmental Microbiology*, 80(12), 3562-3567.

8. Acknowledgments

The research performed in the study 1 was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 643476, named “Collaborative management platform for detection and analyses of (Re-)emerging and foodborne outbreaks in Europe - COMPARE (http://www.compare-europe.eu).

In the research performed in the study 2 the authors acknowledge the help received from the workers of the poultry company, their support in the identification of the farms to investigate, in the planning of the slaughtering of the tested groups as first of the day and in the sample collection during evisceration and at the end of the refrigeration tunnel.

The research performed in the study 3 has been supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 818290, named “Controlling microbiomes circulations for better food systems” - CIRCLES (https://circlesproject.eu/).

The research performed in the study 4 was supported by partnership for research and innovation programme in the mediterranean area (PRIMA) under grant agreement No 181- ARTOSANEFOOD, named “Innovative Bio-interventions and Risk Modelling Approaches for Ensuring Microbial Safety and Quality of Mediterranean Artisanal Fermented Foods” - ArtiSaneFood (http://www.ipb.pt/artisanefood/)

The work was performed at the Food Safety unit of the Department of Agricultural and Food Sciences (DISTAL), Alma Mater Studiorum - University of Bologna (Italy). Nevertheless, a huge part of the work was also performed at the Department of Physics and Astronomy, University of Bologna (Italy) for providing access to cloud computing resources and for the statistical analyses and bioinformatics.

I gratefully acknowledge Prof. Richard Ducatelle and Dr. Prof. Dr. Hauke Smidt for the contribution and time they offered as referees and for the constructive comments and suggestions.

The author expresses sincere appreciation to: During these of doctoral studies, a lot of people have guided and trained me with multiple tasks, and I would like to express my appreciation to all of you. I would like to thank all people at DISTAL for their great contribution in different ways in performing this work. I would like to thank my supervisor Prof. Gerardo Manfreda for giving me the opportunity to perform these interesting research activities, for giving me the chance to further my education, and for his support and guidance during my Ph.D. Prof. De Cesare for accompanying me during my Ph.D. and for sharing her huge knowledge with me. Her immense experience and wise advice as well as her patience, kindness and caring attitude towards me and my work helped me in all the time of my Ph.D. and dissertation writing. I would to special thank Alex (Dr. Lucchi) for having taught me, advised me and above all supported in these years of PhD in laboratory activities and beyond. In particular, I would like to thank my friends/ colleagues Frédérique (Dr. Pasquali), Pietro (Dr. Olivi), Cecilia (Dr.

Crippa), Lucia (Dr. Gambi) my colleagues from upstairs, you all helped me and taught me so much and also for all good laughs and coffee breaks. I would like to Dr. Sala, of the Department of Physics and Astronomy, University of Bologna, for assistance in sequence analysis, statistical and bioinformatics analysis. I would to special thank my family and my big family of friends for supporting me whatever things I do and without whom none of this would be possible. I learned and enjoy with all of you and it has been an endless pleasure to work and share special moments with you. Thank you!