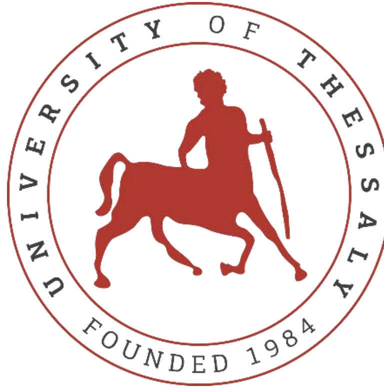


UNIVERSITY OF THESSALY

MSc Program: Host-Microbe interaction (HOSMIC)

MSc Thesis: α/β diversity changes and differential abundance of generic functional, gene mobility, resistance and virulence DNA markers on the soil plasmidome, induced by application of manure and the veterinary antibiotic tiamulin.

Name: Domna Nikolaidou

Supervisor: Dr. Vasileiadis Sotirios

Larisa 2024

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Tripartite committee:

Dr. Vasileiadis Sotirios: (Supervisor), Assistant Professor of Microbial Molecular-Ecology/ Genomics, Department of Biochemistry and Biotechnology, University of Thessaly.

Dr. Dimitrios G. Karpouzas: Professor of Environmental Microbiology and Biotechnology, Department of Biochemistry and Biotechnology, University of Thessaly.

Dr. Evangelia S. Papadopoulou: Assistant Professor of Environmental Microbiology, Department of Environmental Sciences University of Thessaly.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

ACKNOWLEDGMENTS

This thesis took place in the Plant and Environmental Biotechnology Laboratory in the Department of Biochemistry and Biotechnology, University of Thessaly, between March 2024 and September 2024.

First of all, I owe huge thanks to Assistant Professor Sotirios Vasileiadis for the opportunity to carry out my thesis with him and for his help throughout its execution and writing, as well as for its final proofreading.

I would also like to thank Professor Karpouzas Dimitrios for accepting me in his laboratory and being a member of my three-member committee, as well as Assistant Professor Evangelia S. Papadopoulou for his support during the preparation of my thesis.

Finally, a big thank you to all the members of the lab for their advice, help, smooth cooperation and excellent atmosphere during the preparation of my thesis.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Abstract

Fertilization of agricultural soils with manure derived from animals treated with antibiotics may pose ecotoxicological threat for the environment and cause dispersal of pathogens and antibiotic resistance (AMR), threatening livestock animals and the public health, due to produce consumption derived from of these soils. Antibiotics such as tiamulin (TIA) have demonstrated ecotoxicological effects on the soil microbiota and the effect of the increase of genes associated with antibiotic resistance and mobile genetic elements, while putative pathogenic strains were increased as well. Antibiotic resistance in very frequently resides on the microbial mobilome. We aimed at assessing the effects of TIA on the dispersal of resistance and microbial mobilome marker genes. For achieving our aim, in a microcosm approach, we have treated an agricultural soil with TIA and manure fortified with TIA, extracted cells form the resulting microcosms using a Nycodenz gradient approach, extracted the plasmidial DNA using an optimized protocol, and performed shotgun sequencing analysis of the DNA extracts. The data were subjected to quality control, mapping of the sequencing reads on known relevant databases, analysis of α and β diversity with suitable statistical tools. The reduction of chromosomal DNA along the plasmidome extraction was measured on the basis of the reduction of the rRNA coding gene by qPCR and was 75-90%. Our results showed that the ratio of the index values between samples differed for different alpha diversity indices. According to the nMDS plot, we saw that the effect of manure was stronger than the effect of antibiotic on β-diversity, as the samples were grouped according to whether or not they contained manure. We also saw the relative abundance of several genes increase in antibiotic-treated samples compared to those that were not treated with antibiotics. Among these genes were efflux pumps, membrane transporters, integrase genes and antibiotic resistance genes that have the same target as tiamulin. In addition to these we saw the increase of some heavy metal resistance genes, aromatic hydrocarbon degradation genes and virulence factors. This was an initial exploratory analysis on the aforementioned samples, extra analysis including more controls and more detailed analysis using horizontal gene transfer prediction software will further elucidate the effect of TIA concerning the dispersal of AMR and mobile genetic elements.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Table of Contents

1. Introduction 7
1.1 Agricultural practices and environmental impacts 7
1.2 Mobilome 10
1.3 Antibiotics 15
1.3.1 History of antibiotics 15
1.3.2 Antibiotic mechanisms......16
1.3.3 Antibiotic resistance 19
1.4 Metagenome study strategies 22
1.4.1 Sampling and sample processing......25
1.4.2 Nucleic acid extraction......27
1.4.3 Sequencing with high throughput methods 27
1.4.4 Bioinformatics......29
1.4.4.1 Quality control......29
1.4.4.2 Assembly......31
1.4.4.3 Binning 33
1.4.4.4 Annotation 34
1.4.5 Biostatistical analysis of the data 36
1.4.5.1 Alpha diversity 36
1.4.5.2 Beta diversity 37
1.5 Plasmidome analysis strategies 38
1.6 Aim of the study 41
2. Materials/methods......42
2.1 Experimental design 42
2.2 Cell and plasmidome extraction 43
2.3 Exonuclease digestion and quantitative PCR 44
2.4 Multiple displacement amplification (MDA) of intact circular genetic elements 44
2.5 DNA sequencing 44
2.6 Bioinformatics......44
2.7 Biostatistical analysis 45
2.7.1 Coverage estimation and alpha diversity indices......45
2.7.2 Beta diversity......46
3. Results 46
3.1 Dataset features 46
3.2 Elimination of 16S rDNA......47

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

3.3 Alpha diversity indices 47

3.4 Beta diversity 50

3.4.1 General patterns 50

3.4.2 Differential abundance of ARGs, MGEs and pathogenicity related marker genes 51

3.4.3 Other differences between the samples...56

4. Discussion 64

5. References 70

6. Supplementary 90

6.1 Supplementary I 90

6.2 Supplementary II 103

6.3 Supplementary III..111

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

1. Introduction

1.1 Agricultural practices and environmental impacts

One of the biggest global contemporary challenges is the exponential population growth, which implies the need to produce more food (Lee & Tuljapurkar, 2008). This food relies mainly on primary food production (agriculture, fisheries and livestock farming), while increasing the quantity of food relies on improving the techniques involved in these processes (O'Connor, Kleemann, & Attard, 2022). In the livestock sector, a multitude of practices have been implemented over the years to increase production, including the use of antibiotics, which has been very popular in recent years (Aslam et al. 2021). Another common practice is to increase organic matter and nutrients in arable soils through the application of animal manure from conventional livestock farms, though there is a high probability that antibiotics exist in the manure in notable levels, due to prior administration to the animals (Jechalke et al. 2014). Therefore, the combination of these two with the lack of moderation in the use of antibiotics has led to serious problems, such as a chain of infection with antibiotic-resistant bacteria, as discussed in length below.

More specifically, in recent years, antibiotic-resistant bacterial species have been increasing in soil samples (Forsberg et al. 2012, Forsberg et al. 2014). This is, as expected, partially due to the inherent resistance of soil microorganisms, through the selection and retention of random mutations in DNA when microbes come into contact with a limiting agent. Indeed, the more intense the contact with a toxic substance is, the more likely it is for microbes to select for mutations enabling them to develop resistance (D'Costa et al. 2011). In addition to that, there is also the possibility of transferring resistance genes between different bacteria, as observed following the use of manure as organic fertilizer without the necessary detoxification treatment (Sun et al. 2019, Fang et al. 2015). Antibiotics are administered at therapeutic levels at diseased animals, while they may also be administered at sub-therapeutic levels in cases of preventive use (Van Epps et al. 2016). Sub-therapeutic levels refer to the use of antibiotics to prevent disease, because in the event of illness, economic damage will be caused to the producer, but also to promote animal growth in order to increase profit. In any case, there is a possibility that antibiotics and their metabolites may pass into the soil through

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

the manure if it is not properly treated. In addition, the recommended duration of administration (two weeks for humans) is often violated (Aslam et al. 2021), while recommended administration levels have been found to vary by compound, animal and country (Van Epps et al. 2016). Moreover, even in the case of pre-detoxification of manure before its use as fertiliser or reclaimed water, the lack of knowledge and techniques to enable wastewater treatment plants to effectively remove antibiotics can lead to increased antibiotic concentrations in soil (Marti et al. 2013; Zhu et al. 2017). Subsequently, resistant soil bacteria can migrate into the food chain through food cultivation in manure-enriched soils and end up in humans through the consumption of these foods [Figure 1].

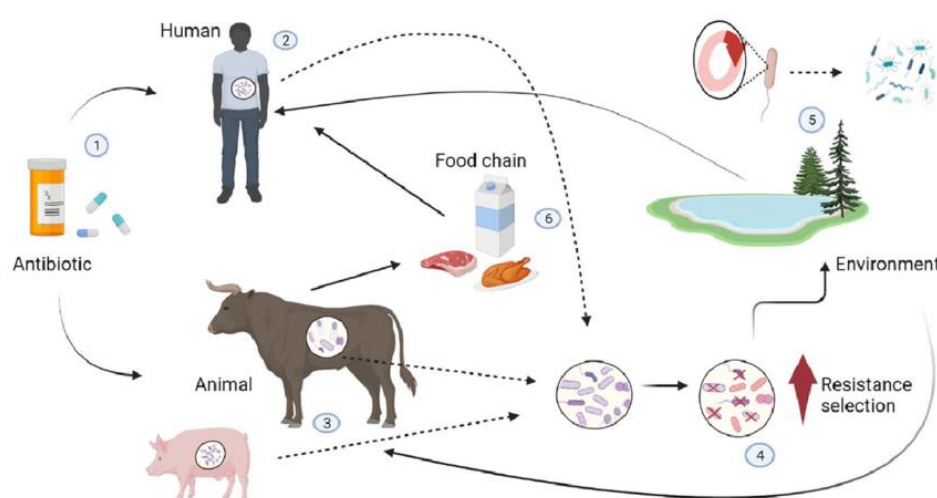


Figure 1 Chain of infection with antibiotic resistant bacteria

Source: Mancuso, G. et al. 2023

Another problem arising from the lack of moderation in the antibiotic use is the ecotoxicological effects on soil ecosystems (Zheng et al. 2022). Maintaining soil microbial diversity is of utmost importance considering the range of different functions these different microorganisms perform (e.g. nutrient cycling, nutritional support for plant growth, pathogen biocontrol, climate regulation, biodegradation of xenobiotics, production of therapeutic bioactive compounds) (Bender, Wagg & van der Heijden, 2016). Little is known about the mechanisms underlying the potential ecotoxicity of veterinary antibiotics on soil microbial communities, due to limited research on the effects of antibiotics on microorganisms essential for essential ecosystem functions (e.g., fungi, protists, and ammonia-oxidizing microbes) (Nguyen et al. 2022).

Initially the majority of veterinary doctors, producers and animal health stakeholders argued that the use of antibiotics is not only beneficial but also essential for the normal development

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

of animals and ensures the health, welfare and productive efficiency of animals (Kasimanickam et al. 2021). This view has completely changed in recent years, now unanimously highlighting the dangers of this perception and officially committed to addressing antibiotic resistance and promoting responsible and prudent use of antibiotics. The human, animal and plant health sectors have a collective responsibility to prevent or minimise use as much as possible to reduce the selection pressures that lead to the development and spread of antimicrobial resistance to various pathogens, human and non-human. To achieve this, responsible and prudent use of antibiotics in animals in the animal husbandry and veterinary sectors, is therefore required. Reducing the use of antibiotics in farm animals is mandatory to protect human health by preventing the spread of antibiotic resistance (Poupaud et al. 2021). For this reason, a framework that takes into account human and animal health, and the impact to the environment was developed, employing an integrated approach, the "One Health Initiative" initiative (Kasimanickam et al. 2021). In the context of the "One Health Initiative" initiative, animals are protected and treated responsibly, while maintaining the safety and quality of the food put on the market. Many governments, international organisations, public health agencies, researchers and private companies, starting in Europe and then expanding worldwide, have acknowledged the threat of antimicrobial resistance and have taken steps to reduce the risk (Kasimanickam et al. 2021, Swann MM. 1969). In order to promote global welfare and protect health and the environment, companies and animal health organisations are committed to operating according to the following strategies (Kasimanickam et al. 2021):

i. protection of animal health and welfare,
ii. limiting the use of antibiotics,
iii. strengthening disease prevention and increasing vaccination,
iv. diagnostic techniques and veterinary expertise, paying more attention to research and development for the prevention and treatment of diseases, and
v. enriching knowledge, transparency and communication through scientific and evidence-based solutions to solve these interconnected challenges.

At the legislative level, the first attempt to ban the sub-therapeutic use of antibiotics in livestock farming to avoid the possible development of antibiotic resistance in human pathogens was made as early as 1969 by the Swann Committee of the United Kingdom (Swann MM. 1969). Within the European Union (EU), this decision on the ban of sub-therapeutic use of antibiotics to increase animal biomass has been adopted since 2006 (More 2020), yet preventive and curative use of antibiotics remained legitimate for a few years, and in several

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

cases resulted in increased administered levels of antibiotics. However, 2011 saw also the beginning of a ban on the preventive use of antibiotics in animals in Denmark (More 2020), while on 28 January 2022, as part of the European One Health Action Plan, the EU restricted preventive use to individual animals, instead of mass administration to groups of animals.

However, it is necessary to allow some time for validating the outcome of these regulatory amendments. An important factor contributing to the decisions leading to this regulatory framework is the fact that some antibiotics are difficult to metabolise and this results to the massive release of the parent compound (up to 75% or more) through animal excreta (faeces, urine) into the manure, which is then used for organic fertilization (Perruchon et al. 2022, Xu et al. 2022). Even when undergoing transformations by the animal's catabolic enzymes or present at sub-inhibitory concentrations, antibiotics are still able to cause the selection/evolution of resistance in soil microorganisms (Chen et al. 2019, Stanton et al. 2020).

In order to reduce the problems caused by the over-/mis-use of antibiotics in livestock farming, alternatives have been tested for increasing the production and addressing the ecological problem without reducing the associated revenues (Gadde et al. 2017). These solutions include, probiotics, prebiotics, synbiotics, organic acids, enzymes, phytogens, antimicrobial peptides, egg hyperimmune antibodies, bacteriophages, clay and minerals.

The spread of resistance induced in soil bacteria by the use of manure containing antibiotics has often been studied under laboratory conditions by simulating field conditions during manure application (Katsivelou et al. 2023).

1.2 Mobilome

One approach in order to study plasmids, viruses, circular DNA elements and non-chromosomal genetic elements in general, is the mobilome studies (Browne et al. 2020). Mobilome studies focus on the screening of the transfer of genetic material between microorganisms, a notable driver of their adaptation and, by extension, their evolution (Siefert J. L. 2009). The mobilome, or in other words the total mobile genetic elements (MGEs) in the cell, is defined as any type of DNA that can move within or between genomes (de Nies et al. 2022). So, under the broader concept of mobilome, the terms plasmids, transposons, insertion sequences, mobilisable genetic elements, integrons and bacteriophages are included (Carr et al., 2021).

Plasmids are highly heterogeneous, non – chromosomal DNA replicons, present in bacteria and archaea (Helinski D. R., 2022). The total number of plasmids in a sample called plasmidome (Dib et al. 2015). They are also called episomes when they are incorporated into

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

the bacterial genome (Dib et al. 2015). Among mobilome genes, signature genes are those responsible for MGE self - replication, but also genes enabling conjugation between microbial cells facilitating intercellular mobilization of genetic material like e.g. antibiotic/metal-resistance genes and virulence factors for tolerating selective stresses and occupying novel niches. Such a transfer can be achieved in prokaryotes through plasmids. Plasmids are considered to be transferred horizontally between cells in the same generation utilising conjugation and transformation [Figure 2]. Briefly, during the conjugation, DNA moves from a bacterial cell to another through a tube (Pilus) (Virolle et al., 2020). In transformation, a bacterium uptakes extracellular DNA from the environment, probably naturally originating from a dead cell after lysis (Gingold E. B., 1985).

Very important members of the MGEs are the insertion sequences (Vandecraen et al., 2017). They are short transposable elements which encapsulate transposition genes, more specifically a transposase gene that has inverted repeats at both ends and this entire element is usually surrounded by two direct repeats. At their simplest form, transposable elements code only for those genes (simple insertion sequences), but usually they can be found to have additional survival genes, like antibiotic resistance genes (Vandecraen et al., 2017). In the case of these composite insertion sequences, the aforementioned system consisting of a transposase gene between inverted repeats is located at both ends of an antibiotic resistance gene and all of this together is surrounded by direct repeats. Insertion sequences spread among cells of the same generation with three ways, by transposing to conjugative elements followed by conjugation, via transformation and transduction [Figure 2]. Regarding the transduction procedure, bacteria-specific viruses (bacteriophages) are being involved, transferring mobile genetic elements between infected bacteria (Thierauf, Perez & Maloy, 2009), though it should be noted that transduction is not related to plasmids in nature.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

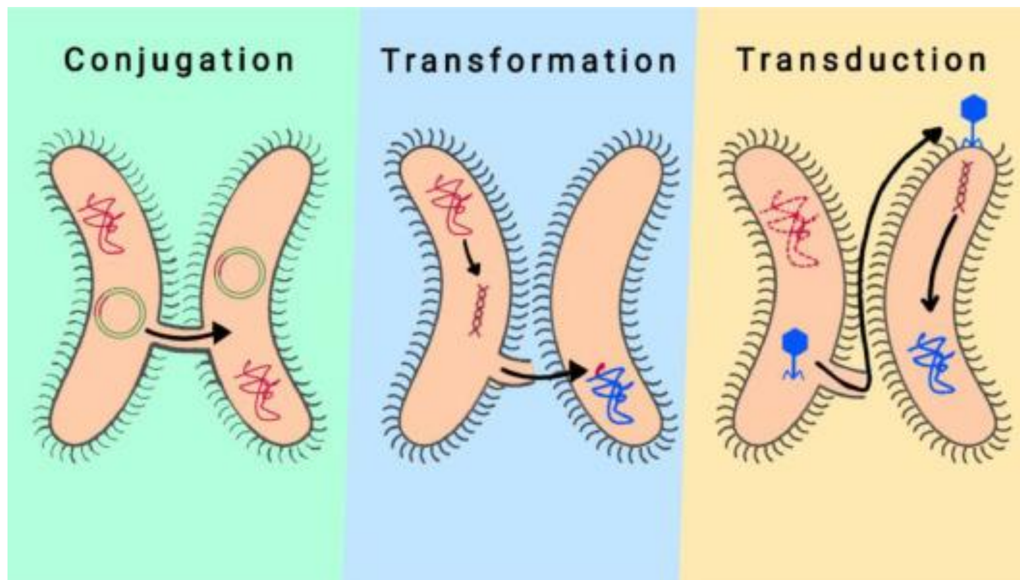


Figure 2: horizontal gene transfer (HGT) mechanisms of bacteria.

Source: Alawi, Velasco-Torrijos & Walsh, 2022

Additionally, MGEs include transposons, or in the alternative term, Integrative Conjugative Elements (ICEs) (Ladd & Bordoni, 2023). These are highly heterogeneous elements, transferred between bacteria through conjugation (Ladd & Bordoni, 2023), while they possess the capacity of insertion into bacterial genomes, through site-specific recombination. After incorporation into the bacterial genome, they passively proliferate during chromosomal replication and cell division (Johnson & Grossman, 2015). Similarly, to the previous MGEs, they usually contain genes that enhance cell survival in the presence of a stress factor . Essentially, when the ICE gene is expressed it results in the induction of a conjugation machinery and the transfer of DNA to appropriate recipient cells, while within this transferred DNA there are likely to be antibiotic resistance genes (Johnson & Grossman, 2015). This conjugation machinery is conserved between ICEs and is homologous to that encoded by conjugating plasmids, i.e. the classic case where two cells come into contact and DNA is transferred from one to the other via a tube extending from the donor cell to the recipient cell. The above process is not usually activated when the ICE gene is integrated into the chromosome, as it needs certain conditions, but it is possible that it can also occur or spontaneously.

The mobilisable genetic elements are highly heterogeneous elements that exist as plasmids, integrated or not in the chromosome (Carr et al., 2021). Like the previous categories, they contain genes promoting survival under stress conditions and they can be moved to another bacterial cell by using the transfer mechanisms of conjugative plasmids or ICEs.

Integrons have the capacity to integrate gene cassettes (Carr et al., 2021). Gene cassettes are small mobile elements consisting of a single gene and a recombination site (attC), with the

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

absence of a promoter. These cassettes can possess a variety of different functions, antibiotic resistance included (Partridge et al, 2018). They can jump to another cell by the site-specific recombination mechanism (Carr et al., 2021). Gene cassettes create an intermediate form of circular DNA molecule and that helps them to travel in another integron and form a larger array of genes. Moreover, the integrons are able to move to another cell by using a combination of composite transposons, conjugative elements, plasmids, or by transformation. In other words, they jump from one genetic element to another within the same cell. Integrons consist of three main elements: the gene that encodes a tyrosine recombinase, the adjacent recombination site (attI) that is recognized by the recombinase and the common promoter (Pc) (Domingues, da Silva & Nielsen, 2012). The tyrosine recombinase is essentially an integrase encoded by the intI gene that performs site-specific recombination within the integron. Based on the genetic relatedness of the intI gene sequence, integrons were initially classified into a few classes (class 1, class 2, class 3, etc.). The promoter (Pc) is located upstream of the integration site and serves to efficiently transcribe and express the gene cassettes present in the integron. LexA binding sites near the promoter regions in many integrases and this enables the host to control transcription of integrons through the LexA protein, which is a transcriptional repressor and results in a reduction in integrase activity (Domingues, da Silva & Nielsen, 2012). Integrons can incorporate one or more gene cassettes; studies have found integrons with up to nine antibiotic resistance genes. The gene cassettes are circular in their free form and take a linear shape when they are integrated into the integron. The attC position is recognized by the integrase and recombination occurs between this position and the attI position, resulting in the insertion of the cassette into the integron (Domingues, da Silva & Nielsen, 2012). Due to the nature of the site-specific recombination, a portion of the attC site is located at the beginning of the gene cassette and a portion at the end. It is understood that this mode of insertion results in the promoter being at the beginning of the integron followed by the attl position, then the gene cassettes separated by the attC positions. For example, the most common resistance gene, sul1 (Sulfonamide resistance gene) is carried due to anthropogenic prossure in class I integrons (Poey, Azpiroz & Laviña, 2019).

Last but not least, the Bacteriophages (phages) are viruses that replicate only in bacterial and archaeal cells, and transfer of genetic elements by phages was previously mentioned briefly under the name "transduction" (Kasman & Porter, 2022). Phages are the most abundant biological agents at a global level. Although they vary enormously in size, morphology and genomic organisation, there is a common structure shared by all of them. This structure consists of a nucleic acid genome, surrounded by a shell of capsid proteins, encoded by the

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

phage, that protect its genetic material and mediate its distribution to the next host cell (Kasman & Porter, 2022). Like other viruses, bacteriophages have a limited range of hosts, this may be a single bacterial species or even specific strains within a species. After attaching to the host, they can replicate and destroy their hosts (lytic phages), or just integrate DNA into the host DNA (lysogenic phages). The lytic cycle involves attachment of the virus to the bacterium, insertion of the viral genome into the host cytoplasm and use of the bacterial ribosomes to produce viral proteins (Kasman & Porter, 2022). The newly synthesized viral genomes and proteins are then assembled to produce many copies of the original phage, which are released into the environment after the death (lysis) of the bacterium and are ready to infect other cells. In the lysogenic cycle there is also the attachment of the phage to a bacterium and the introduction to its genome, but in this case the viral genome is incorporated into the host genome or retained as an episomal element (Kasman & Porter, 2022). In either case, during cell division the viral nucleic acid is replicated and passed on to the daughter bacterial cells without requiring their death. It is possible for one cycle to switch to the other in response to environmental signals.

MGEs can cause problems in human health as they can facilitate the dissemination of genes which give to the microbes the ability to resist antimicrobials (de Nies et al. 2022). This can impact the speed and the efficiency by which a community of microbes adapt and evolve to occupy new niches. In the case of pathogens, such horizontal transfer of resistance genes may ultimately lead to non-treatable human infections.

Eukaryotes have a strong defence mechanism that prevents exogenous DNA from being incorporated into their genome, so the likelihood of transferring MGEs between their cells is low (de Nies et al. 2022). In the other hand prokaryotic genomes are characterized by the presence of multiple MGEs, with the capacity to move between cells in the same generation, known as horizontal gene transfer (HGT).

The effort to understand the mechanism of HGT has been underway for a long time and now sequencing methods and bioinformatics analyses are enhancing our understanding of the ecology of this mechanism. (Siefert J. L. 2009). If the aim is to sequence a specific type of MGEs then targeted genomics are used, but it is also possible to sequence the whole metagenome and then use bioinformatic analysis to focus on one or more MGEs (Carr et al., 2021). Notably, the biggest issue with the de novo MGE discovery is the limitations inherent to the current sequencing technologies, such as the fact that sequencing can help identify sequences similar to those that exist in databases but cannot discover entirely new sequences. Usually, predictions of non-host DNA in the whole metagenome are mediated by the use of

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

bioinformatic methods like CRISPR spacer recognition. Concerning sequencing, the combination of short and long read technologies could be a mean to advance our capabilities to discover novel MGEs, as, instead of getting confused in a sea of sequences present in a sample, it would be more convenient to have sequences that are both long and of high quality, to reduce the possibility of confusing different bacterial genomes with each other (Carr et al., 2021). Short reads are more accurate and are prone to less errors, while long reads are helpful for the scaffolding of short read produced contigs. Therefore, the combined use of short and long reads will allow obtaining larger fragments, which will simultaneously be accurate, as larger fragments correspond to a more general view of the genome of existing microorganisms and thus increase the probability of identifying a pattern related to MGE. Finally, the sequencing is followed by the bioinformatic analysis and the annotation of the resulting contiguous sequences, with the ultimate goal of creating as complete a representation of the mobilome as possible, and this will lead to a more representative description of the microbes of the community and their genes.

1.3 Antibiotics

1.3.1 History of antibiotics

One of the biggest advancements in medicine in the 20th century was undoubtedly the use of antibiotics to treat infectious diseases (Katz & Baltz 2016). Generally, the term antibiotics includes substances, produced by micro-organisms or synthesised in laboratories, that kill or inhibit the growth of bacteria by interfering with their biochemical processes (detailed information is provided further in the section: “1.3.2 Antibiotic mechanisms of action”) (Hutchings et al. 2019). Initially their use was limited to the treatment of infectious diseases, but now they are also used as a mean to prevent infection, for example after an organ transplant, as there is a high risk of infection due to the complexity of surgical procedures combined with the impact of immunosuppression (Prescott 2014, Chan et al. 2020). Their overuse and misuse, meaning use in large quantities and without being necessary, led to the rapid emergence of resistance of the targeted microorganisms, and alternative forms of treatment against pathogens are now being explored.

Before the discovery of antibiotics, the treatment of infectious diseases was inadequate, their treatment and prevention was ineffective, and most patients ended up dying resulting in an overall reduced life expectation (Jayachandran 2018). The first steps in the study of microorganisms were taken in 1676, when

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Antonie van Leeuwenhoek saw and described bacteria under the microscope for the first time in history (Gould 2016). Then, in the 19th century, Robert Koch, while studying *Mycobacterium tuberculosis*, observed and described a relationship between certain bacteria and disease (Nicolaou & Rigol 2018). This gradually created lucrative ground for the discovery of antibiotics in the following years by Alexander Fleming (Uddin et al. 2021), who accidentally discovered that the fungus *Penicillium notatum* inhibited the growth of *Staphylococcus aureus* and attributed this capacity to the secretion of an antibacterial compound. In 1929 this compound was isolated and named "penicillin", which went down in history as the first antibiotic.

Unfortunately, the lack of control in the use of antibiotics both in clinical practice and in animal husbandry has led to the emergence of resistance mechanisms in target microorganisms due to evolutionary pressure (Uddin et al. 2021). In 1961 the first superbacterium in history was isolated and found to possess a resistance towards methicillin (a semi-synthetic derivative of penicillin). This bacterium is known as methicillin resistant *Staphylococcus aureus* (MRSA) (Gaynes 2017). Fewer new antibiotics are nowadays being released in the market compared to previous years (WHO 2017). The lack of regulation of the antibiotic use and the uncontrolled prescription have led to a lack of moderation in their use (Ribeiro et al. 2019). In addition, the fact that antibiotics are administered for a short period of time (1-2 weeks), compared to drugs aimed at treating chronic diseases, leads to a lack of economic interest and a decrease in funding for the discovery of new antibiotics, resulting in few new antibiotics ending up in clinical trials. Most antibiotics were developed during their golden age of research and after this period mostly variant forms of existing antibiotics were marketed.

Even when new antibiotics are licensed, antimicrobial resistance is developed relatively quickly (Michael et al. 2014). In order to reduce the possibility of developing resistance, newer antibiotics are used less frequently, only in severe cases, while older antibiotics with similar efficacy are used for mild infections. Another measure to reduce resistance is to use certain groups of antibiotics only for humans, while other categories are utilized in animal therapy.

1.3.2 Antibiotic mechanisms of action

Depending on their mechanism of action, antibiotics are classified into one of

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

five categories:

1. interfering with bacterial cell wall synthesis,
2. inhibition of bacterial protein biosynthesis,
3. inhibition of bacterial nucleic acid synthesis,
4. inhibition of metabolic pathways; and
5. inhibition of bacterial membrane function (Kapoor et al. 2017).

These mechanisms and some examples are discussed below.

Antibiotics that target the cell wall

The cell wall is a large structure that lies around the bacterial membrane like a shield, designed to counteract intracellular osmotic pressure, determine the shape of bacterial cells and protect against environmental stresses (Garde et al. 2021). Linear glycan strands are covalently joined by short peptide chains to form a single polymer, peptidoglycan. Glycans consist of two monosaccharides linked by β-1,4 glycosidic bonds (Garde et al. 2021).

For the creation of the cross-links, either in the direct connection between the peptides, or in the indirect interpeptide bridges, the main enzymes responsible are d,d-transpeptidases of penicillin-binding proteins (PBPs) and l,d-transpeptidases (Ld) (Garde et al. 2021). The purpose of these enzymes is to reinforce the stiffness, strength, flexibility and structural integrity of the cell wall under stress conditions. Therefore, antibiotics that target these enzymes render bacterial cells more vulnerable and can lead to their death. For example, this is the mode of action for the most notorious and most studied antibiotic, penicillin (Bush, & Bradford, 2016). β-lactams, the structural category in which penicillin belongs, resemble the part of the peptide chain that is bound by PBP enzymes, thus their irreversible binding and inhibition of PBP activity means that peptidoglycan synthesis is no longer possible.

Protein biosynthesis inhibitors

Some antibiotics also target the 30S and 50S ribonucleoprotein subunits that make up the bacterial ribosome 70S, with the ultimate goal of disrupting protein biosynthesis (Yoneyama & Katsumata, 2006). For example, tetracyclines bind the 16S rRNA of the 30S subunit and

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

block access to t-RNA. This class of antibiotics also includes tiamulin, a semisynthetic derivative of the diterpene antibiotic pleuromutilin, widely used in veterinary medicine [Figure 3] (Islam, Klein & Burch, 2009). In general, TIA is one of the most widely used, broad-spectrum veterinary antibiotic (Cycoń, Mrozik & Piotrowska-Seget, 2019).

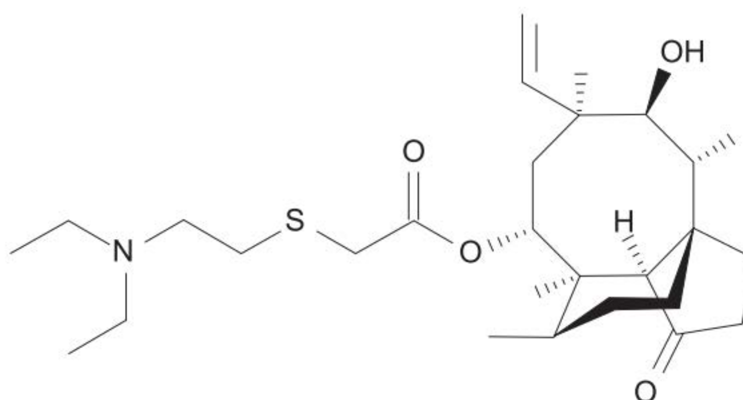


Figure 3 Chemical structure of tiamulin

Source: Krasucka et al. 2010

Its mode of action is similar to other pleuromutilins (Krasucka et al. 2010) and is based on the asite of the 50S ribosomal subunit [Figure 4] (Killeavy, Jogl & Gregory, 2020), due to its strong affinity with the 23 rRNA present in the 50S ribosomal subunit (Krasucka et al. 2010). All of the above result in the cessation of protein synthesis, as thiamin binds to the peptidyl transferase centre and prevents peptide bond formation. However, it does not affect the function of the cell if elongation of the peptide chain (at peptidyl site) has already begun.

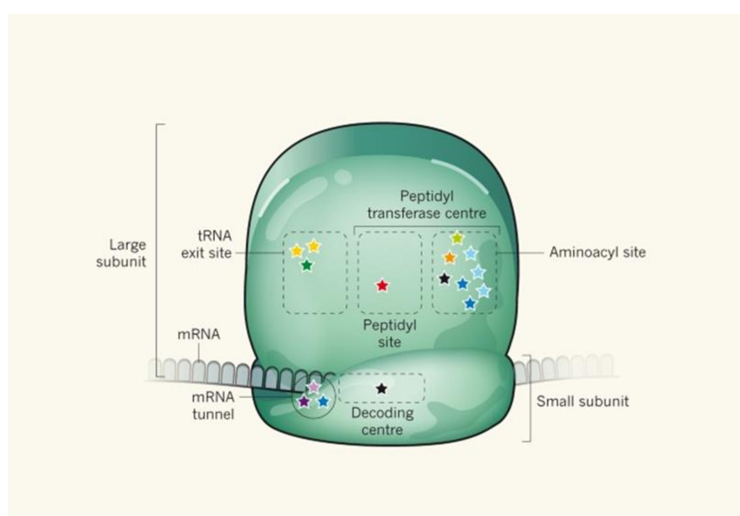


Figure 4 Ribosomal structure

Source: Olivier, 2014

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

DNA replication inhibitors

To stop replication, enzymes that carry it out are usually targeted (Yoneyama & Katsumata 2006). For example, one of the most important enzymes for normal cell function is DNA gyrase (Pozdeev, Mogre, & Dorman, 2021). This enzyme acts during replication or transcription, and aims to cut off overhangs in double-stranded DNA and, subsequently, rejoin the ends. Antibiotics such as fluoroquinolones target DNA gyrase and inhibit the processes it performs (Yoneyama & Katsumata 2006). This antibiotic is safe for humans in low amounts, as in mammalian cells the similar function of the DNA gyrase is performed by topoisomerase II, which has a low affinity for fluoroquinolones (Liu et al., 1983 ,Mukherjee, Sen, & Agarwal, 1993).

Inhibitors of folic acid metabolism

Folic acid (or folate) is an essential molecule for the synthesis of DNA bases (Yoneyama& Katsumata 2006). Inhibitors of folic acid biosynthesis enzymes can act as antibiotics. For greater effectiveness, combinations of two or more different antibiotics that act at different stages in the same biosynthetic pathway, can be considered.

1.3.3 Antibiotic resistance

The development of resistance is an inevitable outcome, as it is a natural phenomenon intensified by human intervention, and not something that is exclusively due to the use of antibiotics in medicine in recent years (von Wintersdorff et al. 2016). Afterall, most types of antibiotics are produced by microorganisms in nature. Apart from the expected mutations due to DNA polymerase errors, resistance spreads also due to the possibility of transferring genetic material between bacteria (von Wintersdorff et al. 2016). Usually resistance genes are located in mobile genetic elements and are transferred to neighbouring bacteria, even if they belong to a different genus or species. There is also the selective pressure factor, which maintains mutations that promote the survival of the best adapted organism in a given environment (Holmes et al. 2016). In a population exposed to an antibiotic, the dominant form of bacteria are ones

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

that have resistance genes against that specific antibiotic (Zhao R. et al. 2019). Resistance genes are located on the chromosome or in mobile genetic elements (Uddin et al. 2021), while, depending on the biochemical pathway conferring resistance, resistance mechanisms fall into the following categories:

- modifications of the antibiotic molecule itself,
- inhibition of the antibiotic's access to the target cell (reduced introduction of the antibiotic into the target cell or its active extrusion),
- implementation of target modifications and
- Modulation of regulatory networks, in order to implement general procedures for the adaptation of target cells to major metabolic pathways.

Modifications to the antibiotic molecule

There are enzymes that inactivate the chemical compound, either by adding chemical groups or by degrading it (Uddin et al. 2021). For example, the β-lactams mentioned earlier are degraded by the action of β-lactamases, due to the cleavage of the amide bond of the β-lactam ring (Abraham & Chain 1988). Changing the stereo-configuration of the antibiotic, either through bond cleavage or group addition, results in a decrease in affinity to its target (Wilson 2014).

Reduced penetration or active extrusion of the antibiotic

Preventing the antibiotic from reaching the target can be achieved by reducing the influx of antibiotic into the cell or by utilizing efflux pumps (Uddin et al. 2021).

Antibiotics have to cross the cytoplasmic membrane in case of intracellular targets or target the inner cytoplasmic membrane in Gram-negative bacteria [Figure 5] (Pagès et al. 2008). In this case, bacteria have to limit drug entry to survive. If the antibiotic is hydrophilic in nature, it is affected by changes in the permeability of the outer membrane, as its introduction usually depends on pores, which are water-filled diffusion channels.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

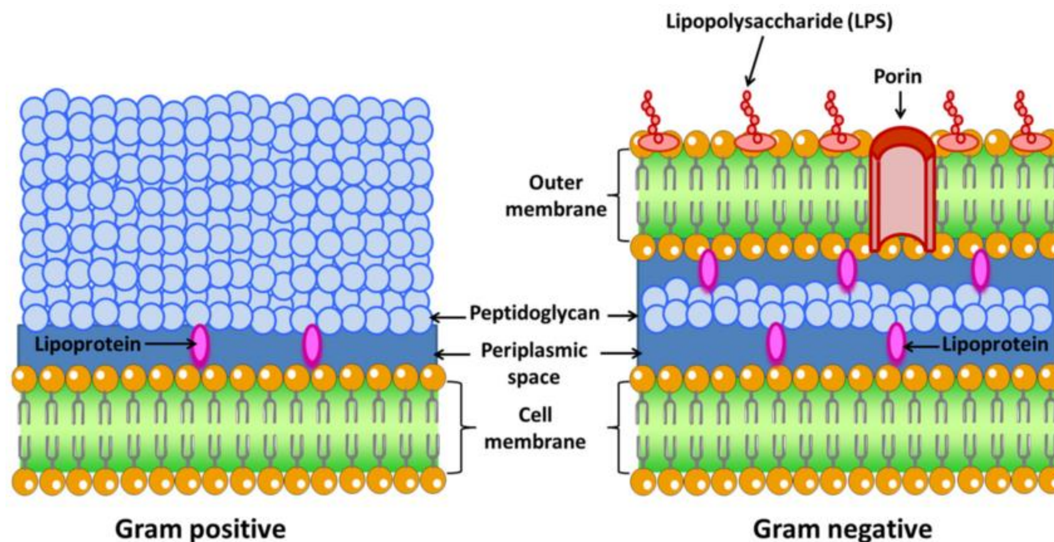


Figure 5 differences between Gram positive and gram negative bacteria in membrane structure

Source: Jiménez-Jiménez, Moreno & Vallet-Regí, 2022

Changes to the antibiotic target

Genetic, epigenetic and/or post-translational modifications in the genetic material of bacteria can lead to a reduction in the affinity of the drug with its target (Munita & Arias 2016). Among others, these modifications include point mutations in the target genes and enzymatic, post-translational modifications at the binding site of the antibiotic (e.g. addition of methyl groups). Complete substitution of the target in the biochemical pathway has even been observed, resulting in similar biochemical functions taking place without being inhibited by the antibiotic.

For example, although resistance to tiamulin is a rare event (Krasucka et al. 2010), it is not impossible to develop (Killeavy, Jogl & Gregory, 2020). This resistance is mainly based on mutations in the genes of the functions targeted by the specific antibiotic, that encode 23S rRNA or ribosomal proteins. These are the main modes of resistance to pleuromotilins, although the contribution of the various mutations in either the ribosome itself, or associated ribosomal proteins has not yet been clearly described. In addition to these, efflux pumps or ribosome protection transporters (not attached to membranes but floating around ribosomes) have also been found (Nagy et al. 2023, Bøsling et al. 2003).

General cell changes

Examples of more general defence mechanisms are: cell wall composition and membrane homeostasis, through changing metabolic pathways and functions in

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

order to adapt to the specific environment (Munita & Arias 2016).

The role of epigenetics to the antibiotic resistance

Genetics is the focus of attention, but it is unable to explain the speed of resistance development and reversibility in a susceptible phenotype (Ghosh et al. 2020). When bacteria are exposed to increasing sub-inhibitory concentrations of an antibiotic, they temporarily develop adaptive resistance. But this resistance is lost when the bacterium stops being exposed to the antibiotic. This kind of survival is not due to genetic factors, as the expected mutation rates are lower than the observed survival rates (Ghosh et al. 2020). Neither is the reversal in the susceptible phenotype explained genetically, as reverse mutations are needed, which are not as frequent as needed for this phenomenon. A fairly reasonable explanation is rooted in epigenetics and stochastic heterogeneity in gene expression patterns. As the changes that occur in this case are not permanent, they are therefore relatively easily reversed (Ghosh et al. 2020). Also, no division needs to be mediated to show its effects in the next generation, justifying the rapid development of resistance.

Practices to avoid

In addition to the inherent ability of bacteria to escape the mechanisms of antibiotics, inconsistent use by humans usually makes the situation worse (Uddin et al. 2021). Patients often use antibiotics without a doctor's prescription or do not complete the antibiotic cycle, resulting in residual bacteria having an increased chance of developing resistance. Even doctors themselves sometimes prescribe antibiotics to patients when they should not, such as in the case of a viral infection, where antibiotics have no effect against viruses, or give broad-spectrum antibiotics when a narrow-spectrum one is appropriate for the situation (Uddin et al. 2021). The situation is worsened by the spread of microbes due to the inability to maintain clean areas in hospital environments, a result of inadequate procedures and protocols. Finally, as discussed at length in section 1.1, one of the most important factors in the development of resistance is the use of antibiotics in animal husbandry and the subsequent use of animal manure as fertiliser.

1.4 Metagenome study strategies

The term metagenomics refers to the direct analysis of genetic material from an environmental sample (Thomas, Gilbert, & Meyer, 2012). The whole concept was coined in

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

1998 by Handelsman and her colleagues (Muhamad Rizal et al. 2020). Through metagenomics, it is possible to analyse functional genes, such as biocatalysts or enzymes, in microorganisms that cannot be cultured in the laboratory (Thomas, Gilbert, & Meyer, 2012). Beyond that, metagenomics delineates the evolutionary relationships of microorganisms and provide information for phylogeny in an economically viable approach, in this context by the sequencing of specific DNA regions (e.g. the 16S r gene) and/or directly, by random shotgun sequencing [Figure 6].

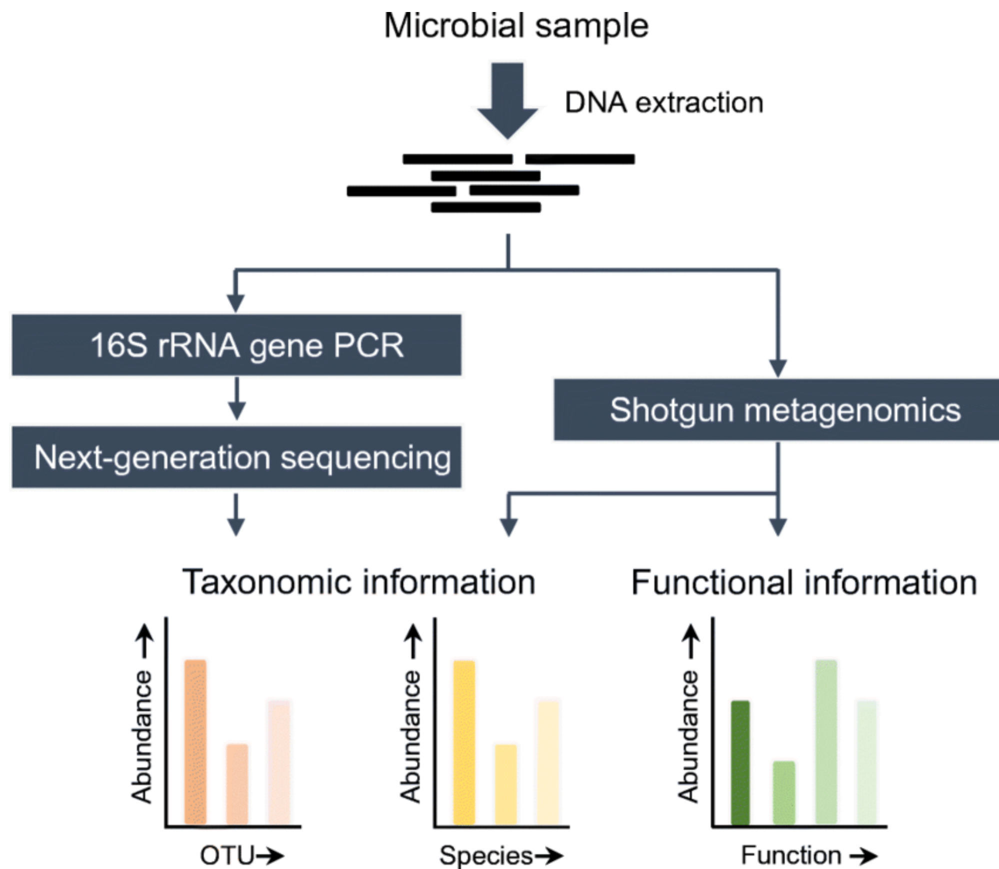


Figure 6 Two different metagenome approaches

Source: Boers, Jansen, & Hays, 2019

In modern microbiome studies, the ultimate goal is to gather as much genetic information as possible to create more accurate metagenome assembled genomes (MAGs), which represent the genome of the microorganisms that make up a sample without prior isolation (Taş et al., 2021). In order to attribute the DNA fragments to the different microorganisms and generate the respective MAGs, various genome properties such as DNA sequence, GC percentage present, tetranucleotide frequency combined with sequence coverage depth and sequence abundance in the sample are taken into consideration.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

The generation of MAGs is based on the shotgun sequencing method (Taş et al., 2021). In this approach, the DNA from the sample is broken into smaller pieces, these pieces are sequenced and then are computationally mapped into position, resulting in MAGs that correspond to the different microorganisms present in the sample (Quince et al., 2017). Cases exist where the MAG generation attempt is not suggested due to lack of the necessary coverage of the genetic content of the samples, owing to their large genetic diversity (Rodriguez-R and Konstantinidis, 2014b). In cases for example where the chieved coverage by the sequencing effort is less than 60%, mapping of the sequencing reads directly on database genes or genome sequences is a good alternative for obtaining important functional and taxonomy information (Karaolia et al, 2021).

This method is untargeted and can be used both to analyse the taxonomic composition of the community and its functional potential (Quince et al., 2017). More specifically, the five key steps of a shotgun metagenome study are [Figure 7](Quince et al., 2017):

i. preliminary steps including the collection, necessary processing and sequencing of samples,
ii. processing of the reads retrieved during sequencing (e.g. quality control, error correction, grouping into tables that are easy to process later),
iii. processing of the data for the purpose of characterising the microbiome taxonomically, functionally and genomically,
iv. statistical analysis (e.g. indices of α- and β- diversity) and biological explanation,
v. Validation of the results

The potential limitations that may arise lie in potential biases during the experiments and the very nature of the approach, which involves complex concepts and interpretations of results (Quince et al., 2017).

A different approach, in order to reduce costs, is to initially process the data by sequencing the 16S rRNA gene, which is a less costly method, and subsequently select samples to which shotgun sequencing analysis will be applied (Tickle et al., 2013).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

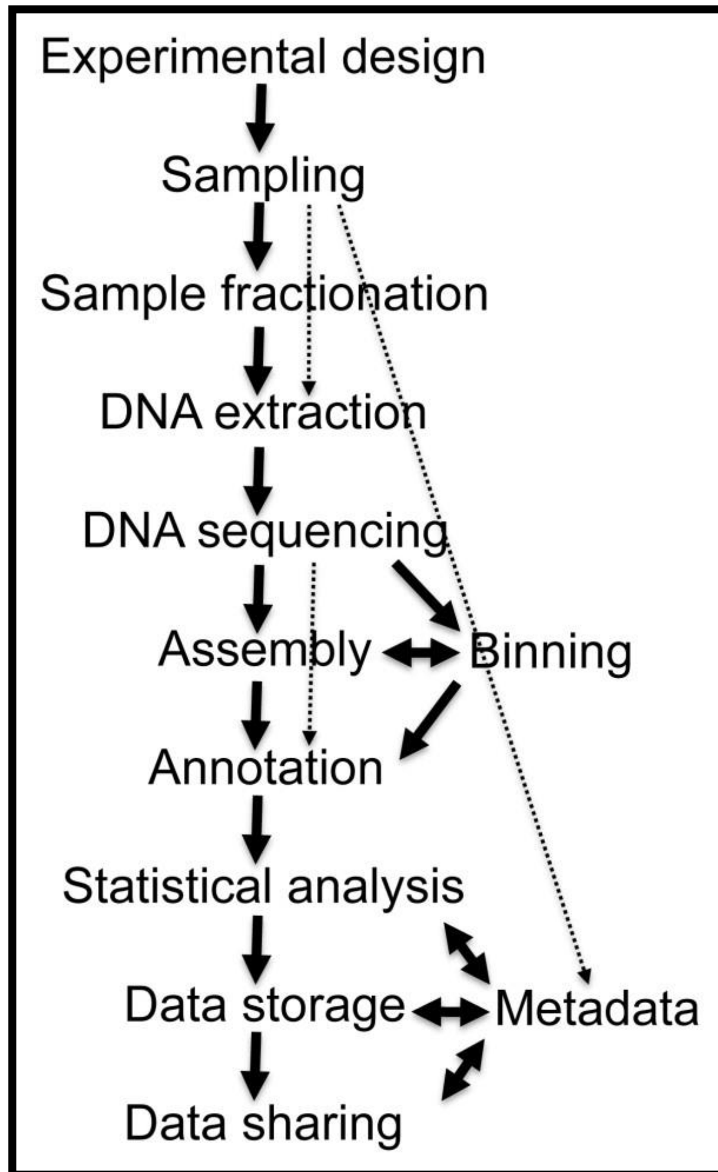


Figure 7 Example of a typical metagenomic analysis protocol. These with dashed arrows can be skipped

Source: Thomas, Gilbert, & Meyer, 2012

1.4.1 Sampling and sample processing

First, based on the subject to be studied, the experiment is designed and outlined, including its conditions and the questions to be answered (Thomas, Gilbert, & Meyer, 2012). In most studies the ideal situation is to compare samples from the same habitat over different time frames (Knight et al., 2012). These studies do not rely on a single sample as it may not be

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

representative. Thus, a key component for the reliability of the results is the existence of a representative sample and its comparison with positive and negative controls in a framework of appropriate parameters (Andrade, 2020). The sampling process is repeated in order to have the appropriate number of samples required for each analysis. To determine the ideal sample size, a statistical software based on some assumptions (power tests) is used, for example the free software G*Power (Andrade, 2020). G*Power requires that the number of groups, number of observations, effect size and significance level are given as data. In case it is unable to reach a hypothesis, then there is the possibility off setting an arbitrary sample size based on previous studies.

The subsequent sampling depends on the questions to be answered (Anderson-Sprecher, Flatman, & Borgman, 1994). The DNA in the samples must represent the type and quantity of different microorganisms and be qualitatively and quantitatively sufficient (Thomas, Gilbert, & Meyer, 2012). It is recommended to remove samples suspected to have been affected by an undesirable factor as to avoid the production of unreliable results (Quince et al., 2017).

If the microbial community under investigation colonises a host, then fractionation or selective lysis is used to get rid of the host DNA (Thomas et al., 2010). If this is not done, in all likelihood the host genome will be at a higher concentration than that of the microbes and will cause a huge amount of noise in the data analysis. If the aim is to enrich a specific part of the community (e.g. viruses), then physical fractionation, induced after selective filtration or centrifugation steps, or flow cytometry is used (Thomas, Gilbert, & Meyer, 2012).

It is also very important how the physical separation and isolation of cells from the samples will be done, so that in the end the maximum amount of DNA can be obtained without the presence of enzyme inhibitors (Thomas, Gilbert, & Meyer, 2012). When processing samples for DNA extraction, enzymatic inhibitors (such as humic acids) that will cause problems in the subsequent analysis must be removed from the samples (Delmont et al., 2011). To avoid bias regarding microbial diversity, DNA quantity and sequence length, cells are usually separated from the soil before lysis is performed. If very small amounts of DNA have been recovered by DNA extraction it is possible to amplify this amount to be sufficient to produce libraries (Thomas, Gilbert, & Meyer, 2012). For example, multiple displacement amplification (MDA) can be performed using random hexamers and phi29 phage polymerase, but it is understood that biases can occur during amplification, so it is advisable to limit the use of the method (Lasken, 2009).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

1.4.2 Nucleic acid extraction

The next step is the cell lysis and the separation of a sufficient quality (e.g. on the size of the fragments) and quantity of nucleic acids from the other components of the cell and the environment (Gupta 2019). DNA extraction should be performed with caution, as the protocol followed is likely to lead to alterations in the microbiome due to the presence of microbes that are more easily lysed (Wesolowska-Andersen et al., 2014). The two most common methods for cell lysis are mechanical methods (bead beating) and the usage of chemical agents (Yuan et al., 2012).

For DNA extraction, there are various laboratory kits that mimic laboratory techniques (Gupta 2019). The process starts with the rupture of the cell wall and membrane, which involves the dissolution of DNA. In this state the DNA is vulnerable against the action of nucleases, so their inactivation takes place. The process is completed by the chemical or enzymatic removal of the remaining cell components (e.g. macromolecules, lipids, RNA, proteins or PCR inhibitors), so that at the end the sample is as pure as possible (Gupta 2019).

DNA extraction can be:

- organic extraction using organic solvents, e.g. by the phenol-chloroform method
- non-organic extraction, e.g. using salts and proteinase K
- extraction using silica gel membrane adsorption.

If only plasmids need to be selected from the total DNA in the sample, then the Nycodenz method can be used (Rickwood, Ford & Graham, 1982). Essentially, it takes advantage of an organic compound, which after centrifugation causes a graded concentration of organic elements depending on the density of particles, at which the cells sit at a particular point (Diebold et al., 2021). Plasmid extraction is then performed using exonuclease, as large chromosomes break more easily due to their size, so they are destroyed by exonuclease and in the end only the plasmid remains which makes up about 1% of the total DNA in the sample (Dib et al., 2015).

1.4.3 Sequencing with high throughput methods

After DNA extraction, transposase is usually preferred for DNA fragmentation due to its low cost (Baym et al., 2015). There is the option of direct sequencing or performance of a preliminary cloning in order to create libraries for sequencing (Knapik, 2013).

The vectors used to create libraries are: plasmids, cosmids, fosmids, phages, but also artificial chromosomes (usually from yeasts or bacteria) (Rondon et al. 2000). The sample is introduced into the vector and transformation of cells that are easy to grow in laboratory

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

conditions (e.g. Escherichia coli) is performed. All cells are screened and those that have accepted the vector with the sample are selected, as their phenotype differs from cells that have not accepted the plasmid at all or cells that have accepted the plasmid that does not have the sample DNA. Non-transformed cells do not develop colonies due to their inability to survive on a substrate containing an antibiotic, as the plasmid contains a respective resistance gene (Green & Sambrook 2019). On the other hand, selection of recombinant plasmids is possible when exogenous DNA is inserted into a plasmid gene, then the plasmid marker gene is no longer expressed and this is usually distinguishable to the naked eye (Chaffin & Rubens, 1998). The β-galactosidase gene is the most common case of a marker gene, being present in many plasmid vectors (e.g. the pUC, Bluescript, pGem series) and when inserted into cells these plasmids express the carboxyl-terminal portion of β-galactosidase, whereas host cells express the amino-terminal portion of the protein (Green & Sambrook 2019). When both parts of the protein are present within a cell, then and active form of β-galactosidase is produced, catabolising the X-Gal substrate and staining the cells with a blue colour. Insertion of the DNA sample in the vector cloning site interferes with the amino-terminal gene, thus an inactive protein is produced that cannot metabolize X-Gal and the cells appear white.

The next step is sequencing either the library or directly the sample DNA (Kimura 2018). There are three generations of sequencing methods (Heather & Chain 2016). The first sequencing technology used is the Sanger method (1st generation sequencing). Essentially, in this method *in vitro* replication of the sample DNA fragments is initialized, with fluorescent dideoxyribonucleotides being randomly incorporated into the DNA chain and resulting in the interruption of replication. The results are either visualized by separation in polyacrylamide gel electrophoresis, with the initial sequences reconstructed based on the size of the DNA replication results, or by using laser beam stimulation in combination with a sequencing machine.

The protagonists of 2nd generation sequencing are methods that depend on the DNA ligase enzyme (one such example is SOLiD technology) and DNA polymerisation methods (base-base or homopolymer-homopolymer) (Heather & Chain 2016). The last category includes technologies like Pyrosequencing, Illumina and Ion-torrent.

The latest technologies so far comprise the 3rd generation of sequencing methods, including methods that are able to record DNA polymerisation in real time, as PacBio does (Heather &

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Chain 2016),or technologies such as Nanopore, which relies on disrupting ion flow due to the passage of a DNA strand through a membrane pore. Common to both aforementioned methods are the huge reads that can be obtained as output.

Of the 2nd generation methods mentioned, the most requently used among researchers is Illumina (Heather & Chain 2016; Metzker 2010). In this technique, the DNA sample is first processed to add sequencing adaptors and multiplexing markers. Multiplexing is carried out based on multiplexing protocols that have the ultimate goal of simultaneously analysing multiple samples with a single sequencing run and thus reducing costs. This is done by extending the 5′ ends of the sequences with oligonucleotides indicating each sample (sample barcode). These barcodes will allow in subsequent bioinformatics analysis to classify each sequence according to the sample it comes from, while the sequencing adapters are designed to bind the sequences to the plate that has the complementary oligonucleotides attached (Heather & Chain 2016, Metzker 2010). This fusion leads to the creation of clusters containing clonal copies, a process called "bridge amplification". Sequence extension is carried out using a polymerase, sequencing primers and fluorescent nucleotides that are differently coloured according to the base they represent, as well as a reversible terminator that prevents the binding of the next nucleotide after the incorporation of each nucleotide. This inhibition is based on the fluorophore at the 3' position of the hydroxyl group, and, after the enzymatic cleavage of the terminator, the nucleotide chain continues to extend (Heather & Chain 2016, Metzker 2010). The fluorophores are excited with high photons and lower energy photos are then emitted by the fluorophores which are is recorded, so the complementary sequence of the strand is sequenced rather than the clone itself.

The shotgun method usually uses illumina technology, as it exhibits very high fidelity levels fidelity (error rate of 0.1-1%) combined with many reads (up to 1.5 Tb per run) (Quince et al., 2017).

1.4.4 Bioinformatics

1.4.4.1 Quality control

Using algorithms, the fluorescence colours during sequencing are converted into text files in FASTQ format, which includes the DNA sequence and the corresponding sequencing quality at each position (Escalona et al. 2016). The first processing that such a file will undergo is demultiplexing, i.e. based on the sample barcode sequences added to the DNA fragments

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

during the library preparation step, the reads will be separated into the samples from which they were derived.

The drawback with most tools used in shotgun sequencing data analysis is that they perform different processing steps and they are built with high-level programming languages, such as Python and Java, that provide limited multi-threading support and are generally time-consuming (Chen et al., 2018). In this context, the fastp program, an ultra-fast, all-in-one sequence pre-processor, utilizing FASTQ files as an input, was developed as an alternative for sequence trimming. With that method, quality control, adapter trimming, quality filtering and per-read quality pruning can be completed fast with one tool (Chen et al., 2018). Similarly to others, this tool uses C++, and provides multi-threading support, resulting in 2–5 times faster processing.

The final step is to classify the sequences according to databases and remove sequences of taxonomic groups irrelevant to the experimental questions (Spohn & Young 2018).

Part of the quality control is the coverage estimation of existing diversity, as this process estimates whether the amount of sequences in the sample analyzed covers the total diversity of the community (Rodriguez-R & Konstantinidis, 2014b). One algorithm that performs this task is Nonpareil, which provided estimates of the relative complexity of the communities. More specifically, this method is based on the fact that datasets with higher coverage are more redundant because the sequence reads are almost random (Dohm et al., 2008). Essentially, the term "redundancy" refers to the fraction of reads in a dataset that matches at least one other read. Its calculation is computationally expensive and prohibitive for real-size sequence datasets consisting of millions of sequence reads (Rodriguez-R & Konstantinidis, 2014b). A solution to this problem is provided by Nonpareil, which estimates the redundancy value by generating a subset of query reads from the entire dataset, based on which the number of matches per query read in the entire dataset is calculated. For each query reading, the total number of matches in the data set (match-vector) is calculated (Rodriguez-R & Konstantinidis, 2014b). The next step is to create a redundancy saturation function by iteratively sampling the identity vector. To perform this, a subset of query reads is selected with a Bernoulli trial per read (with a parameter equal to the sampling fraction) and for each selected query read, the probability of matching another read in the sample is calculated based on a binomial distribution. As a result of the above, the number of expected matches of the reading in the sample decreases proportionally with the sample size (Rodriguez-R & Konstantinidis, 2014b). The result of this technique is the avoidance of redundant comparisons between readings, because all comparisons are calculated once in

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

advance, allowing the calculation of a Nonpareil curve with high resolution (i.e. with sampling segments close to each other), thus saving time and computational power. Another advantage of the algorithm is the possibility of multiple replications in each sampling portion (1024 times by default), reducing the effect of randomness in the sampling. The mean, median and standard deviation at each attempt at sequencing are then calculated in order to estimate the average coverage (Rodriguez-R & Konstantinidis, 2014b). At the end of the whole procedure, a log-gamma regression is fit to the calculated redundancy values using the weighted NL2SOL algorithm (Dennis, Gay & Welsch, 1981). The sequencing effort required to reach a fixed average coverage is calculated from the projected regression line (Rodriguez-R & Konstantinidis, 2014b). Depending on the coverage, a different suggestion concerning the follow-up analysis strategy is followed (Rodriguez-R & Konstantinidis, 2014b). If a large part of the diversity is covered then the whole metagenome is assembled, while if the coverage is not sufficient then the microbial community is annotated based on sequences in the databases.

1.4.4.2 Assembly

If the purpose of the analysis is to retrieve the genome of non-cultured organisms or to obtain protein coding sequences, rather than providing functional community description, short read fragments will be assembled to generate the largest possible DNA fragments (Thomas, Gilbert, & Meyer, 2012). The assembly can be based on reference databases or created from scratch (*de novo*).

Reference-based assembly can be done using software packages that include fast and memory-efficient algorithms (Thomas, Gilbert, & Meyer, 2012). Thus, they are not demanding in computational resources compared with *de novo* approaches described further on (Chevreux, Wetter & Suhai , 1999). This type of assembly requires the processed dataset to be mapped on (meta)genomes present in the databases. Differences between them, such as a large insertion, deletion or polymorphisms, may mean result in a fragmented assembly, or divergent regions not being covered.

In the case of *de novo* assembly, larger computing resources are needed, often requiring hundreds of gigabytes of memory on a single machine and runtimes that are commonly several days long (Miller, Koren & Sutton, 2010). Such approaches provide the possibility of novel discoveries since they do not rely on databas knowledge.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

The diversity of microbial communities at the strain level, which implies high similarity of genetic material, is a problem in the use of assembly algorithms that assume clonal genomes (Thomas, Gilbert, & Meyer, 2012), as clonal assumptions make it difficult to form contigs for certain heterogeneous taxa at specific parameter settings (Peng et al., 2011). This results in assemblies breaking due to non-conserved regions between strains of the same species, with the exception of a few strain-level oriented assemblers, usually this is a concern of the contig binning part (as described it further below).

However, the area of metagenome assembly is still in an early stage and it is difficult to compare different tools (Thomas, Gilbert, & Meyer, 2012). One way to make an unbiased comparison is to assemble a community with known reference sequences in order to evaluate the assemblers.

Limiting factors in assembly are the length of sequencing reads and the size of the total dataset (Thomas, Gilbert, & Meyer, 2012). The results concerning sequence identification are more valid when the reads are more accurate (e.g. illumina reads are highly accurate compared with other chemistries) (Wommack, Bhavsar & Ravel, 2008). In contrast, the large reads (PacBio, Oxford Nanopore) provide less accuracy regarding the actual sequence of the samples screened, but longer contigs provide more confidence regarding sequence functional and phylogenetic annotation, as well as more information and comparison with known genetic data (e.g. through homology searches) would be easier to give interpretable results. It is also more valid to classify DNA fragments based on long contiguous sequences (McHardy et al., 2007). Tools such as Phylopythia, only work reliably above a certain cut-off point, for example 1 Kb. On the other hand, the smaller the size of the data set, the smaller the data processing requirements, meaning a reduction in required computing power (Thomas, Gilbert, & Meyer, 2012). Alternatively, similar reads can be grouped together to reduce the dataset. Merging reads also increases the quality of information, as individual readings generally have lower confidence in accuracy.

However, when the sequencing depth is small, the probability of two reads covering the same DNA fragment is reduced, the assembly will be of lower or even unlikely fidelity, in which case the assembly has no reason to be done and usually the results are directly annotated with the help of databases (Thomas, Gilbert, & Meyer, 2012).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

1.4.4.3 Binning

During binning, the DNA sequences are sorted into groups representing a gene or group of genes originating from closely related microorganisms (Thomas, Gilbert, & Meyer, 2012). This can reduce computational requirements and execution speed. The algorithms used in this process use three types of information: (a) common DNA structural traits (eg a certain GC percentage, or tetranucleotide frequencies) (McHardy et al., 2007); (b) the similarity of the unknown DNA fragment to known phylogentic marker genes (frequently single copy genes – SCG – are employed for this purpose) in databases (Monzoorul-Haque et al.,2009); (c) the variability in relative abundances of contigs of several different samples where differences in conditions/treatments cause a similar variability in the abundances of the DNA owners. Some algorithms combine all aforementioned modes (Brady & Salzberg, 2009).

DNA structural trait binning is not reliable for short reads due to missing information. In such a case, compositional assignment can be improved with the help of extra data, such as a long read generatin sequencing method data or database data (McHardy et al., 2007).

If the unknown sequence is only distantly related to known reference genomes, only a taxonomic assignment at a very high level (eg, phylum) is possible (Thomas, Gilbert, & Meyer, 2012). This assignment approach requires that the fragments exceed a certain length. The MIGA algorithm could contribute to this process (Rodriguez-R et al. 2018). The Microbial Genomes Atlas (MiGA) project offers robust taxonomic classification of a query genome or assembled contig sequences based on Average Nucleotide Identity (ANI) or, for more divergent (deep branching) sequences, Average Amino Acid Identity (AAI) values against a reference genome database. ANI/AAI terms refer to the average nucleotide/amino acid identity of all orthologous genes shared between two genomes and provides robust analysis between strains of the same or closely related species (i.e. showing 80-100% ANI) (Rodriguez-R et al. 2018). After determining the best matching reference gene for the query sequence based on the ANI/AAI values, MiGA evaluates whether the query sequence is open to the same taxonomic rank (e.g. species, genus, etc.) or represents a novel taxon at that rank. Genomes annotated in the same species usually share more than 95% ANI, so a query gene with a best match ANI value <95% most likely represents a new species. The same logic applies to other taxonomic levels. In addition, the query genomic sequence can be searched against unclassified genomes of isolates as well as selected metagenome-assembled genomes (MAGs) and single-cell amplified genome (SAGs) collections to identify the closest

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

relatives in the total available genome sequence space (Rodriguez-R et al. 2018). This approach however is unable to name the unknown sequence may determine if it has been found in another study and facilitate future new taxon descriptions.

An algorithm such as CheckM could help identify problematic genomes before they are deposited in public databases, to prevent errors that could lead to future studies (Parks et al. 2015). CheckM is an automated method designed to estimate the completeness and contamination of a genome based on marker genes that are specific to a genome's inferred lineage within a reference genome tree. Using simulated genomes of varying degrees of quality, the utility of lineage-specific marker genes in estimating the completeness and contamination of a genome has been investigated in comparison to the universal marker genes or domain-level marker genes commonly used (Parks et al. 2015). Marker genes that are consistently collocated within a lineage do not provide independent evidence of genome quality, so they are usually grouped into marker sets to further improve estimates of genome quality. Completeness and contamination estimates tend to have generally low absolute error even when genomes are relatively incomplete (70%) with moderate contamination (10%).

Overall, increasing the number and phylogenetic range of reference genomes would lead to more refined future binning approaches (Thomas, Gilbert, & Meyer, 2012). Post-assembly the binning of contigs can lead to the discovery of genomes of unknown or uncultivable organisms. However, new genomes must be validated, otherwise false information can be spread and cause problems in future analyses.

1.4.4.4 Annotation

There are two possible approaches to annotating metagenomes, one involving genome reconstruction and the other performing community-wide annotation (Thomas, Gilbert, & Meyer, 2012). More specifically, if the data are sufficient and the contigs large enough, the individual genomes in a sample can be reconstructed and the assembly process can result in large assemblies, as shown in previous studies (Aziz et al., 2008). However, contigs in this approach must be at least 30,000 bp in length. Alternatively, annotation can be performed on the entire community, according to unassembled reads or short contigs (Thomas, Gilbert, & Meyer, 2012). In this case, the tools for genome annotation are less flexible than those developed specifically for metagenomic analyses. This process consists of two steps, starting

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

with identification (feature prediction) for the genes of interest, followed by putative gene functions and assignment of taxonomic neighbours for evaluating the interactions between them (functional annotation).

Sequences can be identified as genes or genomic elements based on an algorithm that must be accurate (> 95% accuracy) and have a low false negative ratio (Lukashin & Borodovsky, 1998). This classification is based on data such as codons, a case where the algorithm can locate stop and start codons or species-specific patterns of ribosomal binding site, and identify genes in a sea of sequences. An example of such a tool is the MetaGeneAnnotator (MGA)/ Metagene (Noguchi, Taniguchi & Itoh, 2008).

It is also possible to identify non-protein-coding genes, such as tRNAs (Lowe & Eddy, 1997), signal peptides (Bendtsen et al., 2004) or CRISPRs (Bland et al., 2007). However, this requires large computational resources or large consecutive sequences.

Unfortunately, a percentage of metagenomic sequences cannot be annotated (depends on the cutoffs used), this mainly lies on the fact that annotation is based on databases and there is currently low capability of *de novo* function prediction and annotation (Gilbert et al., 2010). Although software exists where *de novo* prediction of protein structures can be used in protein-ligand interactions which can provide functional information hints, such approaches are nevertheless still at their infant stages (Zanghellini, 2014). The unassigned sequences are called ORFans and can occur due to:

1. erroneous protein coding sequences based on imperfect detection algorithms,
2. so far unknown gene functions or
3. structural homology with known genes but not homologous sequences

Techniques to reduce ORFan can include the structural analysis of proteins (e.g. by NMR and X-ray crystallography) and subsequent biochemical characterisation (Godzik A., 2011).

A variety of databases are available for functional gene annotation, such as KEGG (Kanehisa et al., 2004), eggNOG (Muller et al., 2010) and PFAM (Finn et al., 2010). They are used in combination, as none covers all biological functions. Tools such as MG-RAST (Glass et al., 2010) merge and visualize the results of the databases. MG-RAST functions as a data repository, a fully automated analysis pipeline and a comparative genomics environment. It

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

finds application in quality control, feature prediction and functional annotation. The results it gives are in the form of abundance profiles for specific taxa or functional annotations. There is an option for users to download all data products generated by MG-RAST, share and publish them within the portal. Statistical analysis of the results is also possible.

1.4.5 Biostatistical analysis of the data

1.4.5.1 Alpha diversity

Alpha diversity is defined as the diversity at a local scale, i.e. it outlines the species (or genes) diversity (richness) within a functional community (Andermann T. et al. 2022). In the case of microorganisms or their genes a community is an environmental sample.

In other words, alpha diversity indices are a function of the structure of a community of (micro)organisms (or their genes), its richness (number of taxa) and evenness (distribution of group abundance) (Willis A. D. 2019). Comparison of community structure through alpha diversity is recommended because it is susceptible to environmental changes. Usually, one of the first steps in data analysis in the field of microbial ecology is alpha diversity analysis, so that some conclusions can be drawn about the differences between environments. Microbial ecologists apply indices adopted by macroecology in microecology communities with a view to address questions relating to taxonomy (conserved marker) or/and function (metagenome-based data) (Finn D. R. 2024). The most common alpha diversity indices are:

- The observed richness: the set of different (micro)organisms observed (Li et al., 2022),
- Estimated richness: the actual set of microorganisms expected to be present in the sample (Li et al., 2022),
- the ACE index: another method to find the estimated richness in a community (Chao, 1987),
- Shannon index: provides the entropy and is more representative of the broad community (includes also the lower in abundance members) (Jost L, 2006),
- Simpson index: provides the probability that two individuals randomly selected from a sample will belong to the same species and shows the high abundant microbes (Simpson, 1949),
- Inverse Simpson index: like the Simpson, is more representative of the more dominant members of the community compared with the Shannon (Jost L, 2006) and
- Fisher's alpha index (Fisher et al., 1943): indicative of the highly dominant microbes

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Instead of microbes it is possible to have every other event, for example you can study the gene ration in a sample by using this alpha diversity indexes.

1.4.5.2 Beta diversity

Beta diversity is described as the differentiation of microorganisms or genes between different samples (Andermann T. et al. 2022). In other words, β-diversity is defined as the variability in species composition (and diversity) between sampling units for a given area (Anderson, Ellingsen, & McArdle, 2006). For these studies usually two or more samples are compared and their differences are identified. This, in conjunction with the different treatments of the samples, leads to an interpretation of the results and the association of a specific treatment with a specific species richness. During the study, differences between samples, such as those relating to false positive results, statistical power and depth of sequencing, should be taken into account. (Zhu & Yu,2009).

Many studies aim to capture beta-diversity, the methods used are aimed at testing population structure in mapping genomic correlations of complex traits (Zhu & Yu,2009). One approach to describing differences between communities is the Non-metric Multi Dimensional Scaling (NMDS) method. In this method, the differences of complex multidimensional data are represented in a small number of dimensions (2 or 3) in the form of diagrams (Zhu & Yu,2009), thereby reducing the data, visualising their differences and highlighting the gradient structure within the data. This approach highlights which samples are similar to each other and which are the most different, while it also captures the pattern of major structure in correlation mapping samples. The samples that are more similar to each other are closer on the diagram compared to the residues that are more distant
The clustering heatmap provides an interactive visualization for classifying the relationship between the genes present in the samples (Yu et al., 2020). If it is known which genes are differentially abundant in which samples, then there is the possibility of developing predictive models (Yao & Liu, 2018). Thus, heatmaps are often used in differential abundance analysis studies, as they are very useful for data visualization and quality control (Zhao S. et al., 2014). NMDS and heatmaps can be used in combination, so that one can show the relationships between the samples and the other can show the relationships between the different genes in the different samples respectively (Silva et al., 2017)

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Also, to compare proportions of a categorical outcome according to different independent groups, such as for example gene copies between samples, a statistical test such as Fisher's exact test can be used (Kim H. Y. 2017), which presents genes differentially abundant between samples. Assessing the independence between two variables is possible under the assumption that the compared groups are independent and uncorrelated. The outcome is more accurate for small-sized samples, but, in all actuality, it is valid for all sample sizes (Kim H. Y. 2017). Replicates of the samples are not necessary for the use of the Fisher's exact test. Fisher's exact test works by testing the null hypothesis of independence by applying the hypergeometric distribution of gene numbers between samples

1.5 Plasmidome analysis strategies

For the analysis of horizontal gene transfer events, nucleotide sequence data analyses have become predominant. These analyses are divided into two subgroups, one comprised of parametric and the other of phylogenetic methods [Figure 8](Ravenhall et al. 2015). More specifically, parametric ones analyse parts of a genome that differ significantly from the rest of it. For example, they may be based on the GC ratio in the genome or different codons (Ravenhall et al. 2015). Phylogenetic methods, on the other hand, study the evolutionary histories of the respective genes, based on their sequence, comparing them and identifying conflicting phylogenies. Phylogenetic methods are further divided into those that use phylogenetic trees (explicit) and those that use other measures (implicit) (Ravenhall et al. 2015). An example of an implicit method is comparing sequences between species, where if the similarity between specific genes differs from the rest of the species genome it is likely the result of horizontal transfer. The common feature of all implicit phylogenetic methods is the bypass of gene tree reconstruction, in contrast to explicit methods based on it. Even though both methods are based on distances (Ravenhall et al. 2015) the distances are used in a different way. In one case there is a regression that combines the gene distance with the whole genome distance based on the genes that definitely did not come from HGT and if the comparison between the query gene and the microorganism genome is significantly different from regression line, it means it is a result of horizontal transfer (implicit). In the other case a phylogenetic tree is created based on some reference gene (e.g. 16S rRNA), the similarity is compared with the phylogenetic tree that creates the gene we want to test and if the two trees differ then the gene being tested is a product of horizontal transfer (explicit).

Phylogenetic methods are most often used because of the large amount of sequences derived by different genomes, so it is possible to determine even the donor species and

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

when the gene was transferred (Ravenhall et al. 2015). However, the possibility of being led to wrong conclusions is not zero. It is possible that the model does not take into account the paralogous genes, which exist in the same organism, originate from the same ancestral gene, and are created due to duplication of genetic loci (Koonin E. V., 2005). In addition, there may not be enough data to lead to a reliable phylogenetic tree (Ravenhall et al. 2015), while computational cost may also be a limiting factor. Usually the combination of different methods (parametric and phylogenetic) can lead to a prediction that is more likely to be correct. This analysis has been used to study the spread of antibiotic resistance.

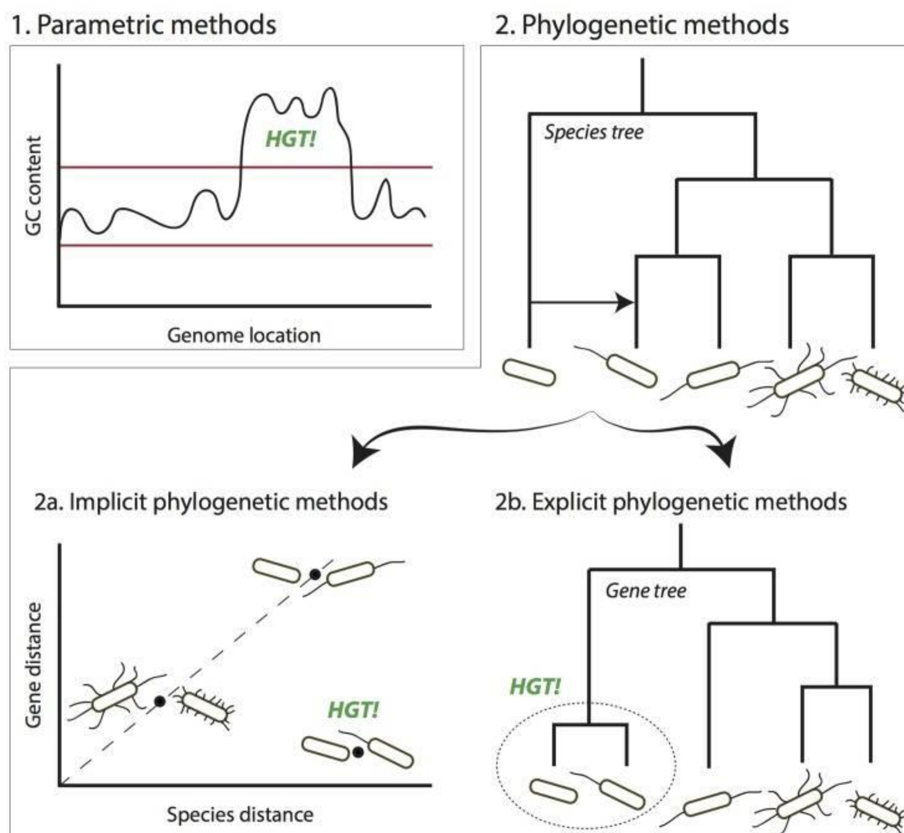


Figure 8 Different strategies for analyzing horizontal gene transfer

Source: Ravenhall 2015

As mentioned above, plasmid analysis belongs in a family of methods implemented to for the identification of MGEs [Figure 9] (Carr et al, 2021). One set of such strategies involves the plasmidome enrichment (targeted metagenomic method), while another approach is to sequence the whole metagenome and, during the bioinformatics analysis, identify plasmid sequences (whole metagenomic method). These methods are used to identify known and unknown MGEs and to predict host MGEs.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

For example, targeted metagenomic methods may rely on the purification of MGEs before sequencing (Carr et al, 2021). Circular plasmids can be isolated by the exonuclease method (Rickwood, Ford & Graham, 1982), which is described in section "1.4.2 Nucleic acid extraction". However, these targeted approaches may incorrectly estimate the potential MGE content, i.e. during the use of the exonuclease method, as, due to the fact that the nucleic acid purification is done randomly (based on DNA breaks) plasmids of interest can be destroyed by the exonuclease if they possess any breaks, thus the procedure would yield MGEs that do not reflect the original microbial community (Carr et al, 2021). Another alternative for plasmid isolation is the Transposon-aided capture (TRACA) method, which uses a DNase that selectively removes the linear DNA, leaving the circular plasmids, which are then selected with the help of a transposon (in an *in vitro* transfer reaction) with an origin of replication and selection marker (Jones & Marchesi, 2007). It is then cloned, sequenced and additional PCR is used to close the gaps in the sequences (Smalla, Jechalke & Top, 2015). There are several limitations, as TRACA does not capture smaller plasmids (between 3 and 10 kb), excludes linear plasmids, and transposon insertion may disrupt gene expression (Dib et al., 2015). Alternatively, to identify small circular plasmids in metagenomic samples, reverse PCR techniques and multiple displacement amplification could be used, described in section "1.4.1 Sampling and sample processing" (Jørgensen et al., 2014). Moreover, the repetitive regions characteristically flanking transposable elements can be employed to identify them in a targeted way via PCR amplification (Tansirichaiya, Mullany & Roberts, 2016). Amplified repetitive regions can be purified and ligated into plasmids, and then transformed into cells. Finally, after clonal expansion the plasmids are isolated, sequenced and followed by bioinformatic analysis.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

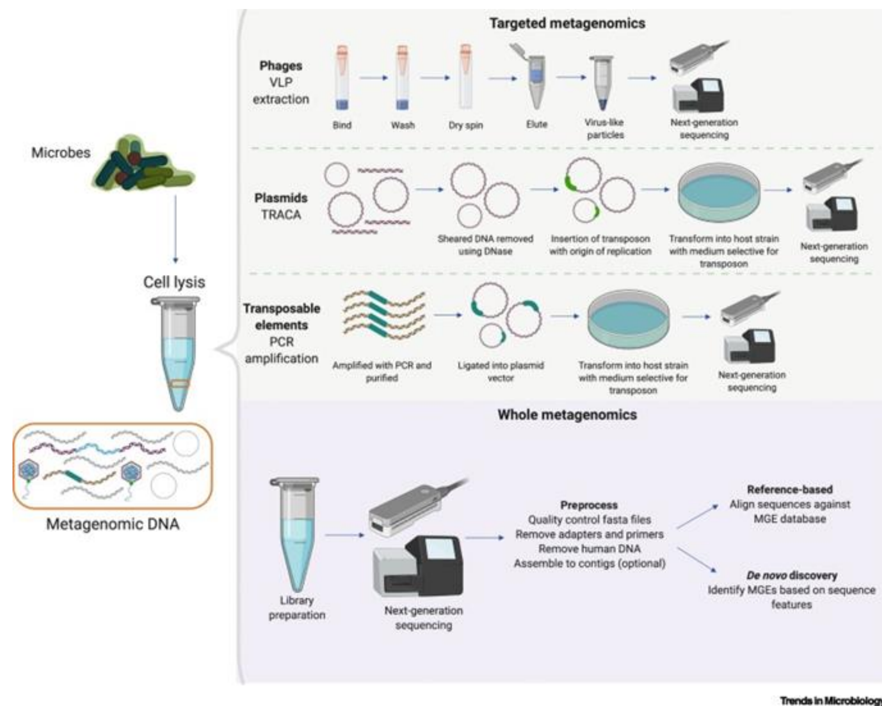


Figure 9: Targeted and whole metagenomic methods

Source: Carr et al, 2021

1.6 Aim of the study

In this study, we wanted to investigate the AMR and HGT associated functional changes imposed to the soil mobilome by the selective pressure of the antibiotic tiamulin, after being introduced to the soil via manure application. For achieving our goals we have fortified our samples with tiamulin, either directly or through manure, we have performed shotgun sequencing with the Illumina technology, and analysed the resulting reads by contrasting them against suitablereference databases. We have analysed the α-/β-diversity of the soil mobilome. The beta diversity was assessed by utilizing NMDS to reveal the different patterns of relationships between the samples, and heatmaps were generated to observe changes in statistically significant genes between samples. The ultimate goal was to determine whether the antibiotic tiamulin, if present in the samples, leads to an increase in the abundance of resistance genes or genes present in the mobilome in general. The generation of the data used in this paper was based on DNA extracts of samples collected at a previous work of the team (Katsivelou et al 2023) and the procedure is briefly presented in the materials/methods section.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

2. Materials/methods

2.1 Experimental design

The data we used in this analysis came from a previous experiment (Katsivelou et al. 2023), and the process used for their generation is described below.

Material used

Antiobiotics

The veterinary antibiotic used in the experimental procedure was tiamulin fumarate in form of a technical standard (TIA- Biosynth® Carbosynth®, Staad, Switzerland, purity > 97 %).

Soils

Soils from two different origins were also used, one from Rodia and one from Livadi. These soils differed in their tiamulin biodegradation capacity and pH, as found in a previous study (Perruchon et al., 2022). Regarding pH, the soil from Rodia had a value of 7.9 while from Livadi 5.9. Due to its alkaline pH, the soil from Rodia had a higher biodegradation capacity of the antibiotic. The two different soils were freshly sieved with a 2 mm mesh and stored at 4°C until the experiment was applied.

Manure

The manure used did not have TIA, as the antibiotic was never administered to the pigs that produced it, but for verification purposes it was also tested by HPLC analysis. At the start of the study pigs were 30 to 45-day old and were not treated with any antibiotics at that time, while any antibiotics received prior to treatment were for therapeutic purposes only. If for any reason there was a need for a continued treatment with an antibiotic, the pigs receiving it were removed from the test group 10 days before the manure sample was taken. In addition, the pigs were weaned (< 3 weeks post-weaning) and their diet contained 72 mg kg^{-1} of zinc in order to promote growth and prevent diarrhea (Castillo et al., 2008). Manure samples were stored for 3 weeks at 4 °C in the dark until used for the experiment.

Process

A total of 5 samples were obtained: 1) pure manure, 2) pure soil, 3) manure with tiamulin, 4)soil with tiamulin and 5) soil with manure and tiamulin. Manure application was performed at a rate of 2 % (w/w), simulating a crop area corresponding to 25 tons of manure per hectare, which according to studies is the maximum amount of pig manure used in agriculture without any environmental risk, for example exposure to heavy metals (Köninger

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

et al., 2021). In the simulation experiment it was assumed that the manure was applied 10 cm from the soil surface, while the bulk density of the soil was 1.5 g cm^{-3}.

The treatment involving application of the antibiotic directly to the soil was established to determine the inherent ability of tiamulin to affect soil microbial communities, while, the treatment where the antibiotic was mixed with manure before soil application represents the more realistic application conditions. In each case, the concentration of TIA in soil was 2.5 mg kg^{-1} soil (dry weight), based on a previous laboratory study reporting the concentration of TIA in the faeces of animals treated with TIA (Perruchon et al., 2022). Control samples included soils with no TIA added to manure and soils with water only. After treatment, the samples were incubated in ventilated plastic bags, in the dark, at 25 °C and the moisture content of the soils was maintained at 40 % of their maximum water holding capacity throughout the experiment. Samples were collected, with three replicates each, on days 0, 7, 15, 25 and 50 post application, with the sampling strategy relying on a previous study of the group (Perruchon et al., 2022).

2.2 Cell and plasmidome extraction

For cell and plasmidome extraction, the Ouyang protocol (Ouyang Y. et al 2021) was employed with some amendments. First, cell separation from the other components of the environment was carried out, with 40g of soil and 80ml of Tween20 (0.5% in PBS) being mixed at high speed for 3 min by using a mixer tucked in ice, in order to avoid the overheating. After that, 18mL Nycodenz (90%) were added to a sterilized falcon (50ml upper limit) and 20mL of soil suspension was slowly added at the top, resulting in a biphasic mixture. Then, the samples were centrifuged at 10000g for 40min at 4°C. Two extractions were carried out. At the first extraction, the cell band was collected in a sterilized falcon and the soil pallet was resuspended in 20mL Tween20 and vortexed for 3 min. The sample was transferred in 18mL Nycodenz and a second centrifugation was performed at 10000g for 40 min at 4°C. At the second extraction, a new cell band was selected and added to the first one. The cells were washed twice with a 5mL PBS, centrifugation was performed at 10000g for 5min at 4oC and the final cell pellet was retained. The above procedure was undertaken entirely in sterilized solutions and utilized sterile buffers.

For further processing, cells were lysed using the bead-beating technique and DNA was recovered by isopropanol precipitation. In more detail, the cells were dissolved in 1mL PBS and transferred to the PowerBead Pro Tube containing the beads. Vortexing was performed for 20 s to remove the DNA from the cells, the solution was allowed to incubate for 2

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

minutes at room temperature and was then transferred to a sterile falcon tube. Finaly, the NucleoSpin Plasmid kit was used for the completion of the DNA extraction.

2.3 Exonuclease digestion and quantitative PCR

The Plasmid-Safe™ ATP-Dependent DNAse (EPICENTRE® Biotechnologies, USA) was used for digestion, according to the manufacturer's instructions, in a 50 μL reaction volume at 37°C. The primer set Eub338/518, which amplifies the 16S rRNA gene, was used for the qPCR to verify whether there was complete removal of chromosomal DNA. The remaining DNA was washed twice by Sigma water and then concentrated to 20 μL by using Amino Ultra 0.5 mL Centrifugal Filter Units (Merck Millipore, Germany).

2.4 Multiple displacement amplification (MDA) of intact circular genetic elements

2 μL of buffer A (400 mM KOH, 10 mM EDTA) and 2 μL of DNA (0.5-5 ng) were added in a precooled PCR tube. 3 min of denaturation was allowed at room temperature and then 2 μL of buffer N [200 mM HCl, 300 mM Tris-HCl (pH 7.5)] (Merck KGaA, Germany) and 10 μL of 1.2 M trehalose were added and mixed with 16 μL of DNA. The mixture was kept in ice, while a mixture of 1.0 μL phi29 DNA polymerase and 29 μL REPLI-g Reaction buffer was also prepared in ice. After stirring, 24 μL of this mixture was added to 16 μL of DNA and the samples were left at 30°C for 10 h. After that, in order to stop the reaction, the temperature was raised to 65°C for 5 min and finally the samples were cooled on ice. The PCR products were quantified by Qubit™ fluorometer (Invitrogen, USA) and stored in the freezer at -20°C.

2.5 DNA sequencing

The sample sequencing was carried out utilising the Illumina HiSeq 2X250bp platform at Admera Health (New Jersey, USA).

2.6 bioinformatics

After sequencing, quality control of the reads was performed, trimming sequences with minimum Phred Q cutoff equal to 20 at a sliding window size of 4 bases and maintaining sequences longer than 70 bp mil length. After that, matching of the read-pairs with their inserts of origin was made by using the fastp v0.21.0 software (Chen et al., 2018). Then, anything that survived from the previous step was compared with a range of databases with the help of the diamond v2.0.15.153 software (Buchfink et al., 2015), with cutoff e value and identities of 1e-10 and 70% respectively, over a minimum length of 30 amino acids of translated sequence. These databases were:

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Aromadeg for aromatic hydrocarbon degradation genes (Duarte et al. 2014);
- BacMet v2.0 for biocide (metals and disinfectants) resistance genes (Pal et al., 2014)
- deepARG for antibiotic resistance genes (Arango-Argoty et al. 2018),
- NanoARG for mobile genetic element (mge) marker genes (Arango et al., 2019)
- Resfinder v4.0.0 for clinically relevant antibiotic resistance genes (Bortolaia et al. 2020);
- SEED for generic functional gene annotation (Overbeek et al., 2014);
- VFDB for virulence factor coding genes (Liu et al., 2022) and
- The Parallel BLAT v36x2 algorithm (Wang and Kong, 2019) was used for searching against the SILVA v138 database (Prüsse et al., 2011) for counting 16S rRNA genes, for hit normalization purposes. According to the DNA extraction method the ideal would be to not have 16s rRNA genes in the sequencing results, but if any were obtained, they would be useful for normalization.

Also, the Kaiju v1.7.3 software (Menzel et al., 2016) was used for the taxonomic annotation of prokaryotes, eukaryotes and viruses from metagenomic whole genome sequencing experiment.

2.7 Biostatistical analysis

2.7.1 Coverage estimation and alpha diversity indices

To evaluate the coverage of the sequences actually present in the sample compared to the sequences that were eventually sequenced, the Nonpareil v3.304 algorithm (Rodriguez-R and Konstantinidis, 2014b; Rodriguez-R and Konstantinidis, 2014a) was used, in combination with the Nonpareil data analysis package v3.3.1 (Rodriguez-R, 2018) in the R v4.4.0 software (R Core Team, 2024). After that, the post annotation gene matrices were used for alpha and beta diversity analysis.

The Good's coverage estimate was further used for assessing the achieved gene coverage (Good, 1953), was used, through the "entropart" v1.6-11 package (Marcon, E., & Hérault, B. 2015). The alpha diversity indices were calculated and visualized by using the "vegan" v2.6-6 (Oksanen et al., 2024) and "graphics" (R Core Team, 2024) R software packages respectively. These alpha indices were:

- the Shannon index (Shannon, 1948),
- the Simpson index (Simpson, 1949),
- the inverse Simpson (Gregorius, H.R. & Gillet, E.M., 2008),

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- the Fisher's α index (Fisher et al., 1943),
- the ACE index (Chao, 1987),
- the observed richness (Li et al., 2022) and
- the estimated richness (Li et al., 2022)

2.7.2 Beta diversity

The Non-metric multidimensional scaling (NMDS) plots were generated by using the "vegan" v 2.6-6 package (Oksanen J et al. 2024) to the relative abundance data. The visualization was based on functions from the "base" package in r studio (R Core Team, 2024).

Only the top 100 most abundant genes were retained, in order to lower the time requirements. Then, the differentially abundant genes were discovered by the observation of the differential abundance without replicates with fisher's exact test between the samples, by using the package "edgeR" (Chen Y, Lun ATL & Smyth GK, 2016). The biological coefficient of variation (bcv) was set to 0.2. The adjustment of the p-value was performed by using the Benjamini-Hochberg procedure (Benjamini, Y & Hochberg, Y, 1995). After that, the values were converted to relative abundance values and, in order to reduce the effect of the dominant taxa, each value was transformed to square root and ranked by using the "vegan" v2.6-6 package (Oksanen J et al. 2024). By the end, the "pheatmap" package (Kolde R, 2019) was used to visualize the statistically significant changes in the relative abundance of the top 100 most abundant genes at the five different treatments.

3 Results

3.1 Dataset features

The above table [Table 1] gives the portion of raw data, the data after the quality control and the 16S rDNA markers.

Table 1: Portion of raw data, the data after the quality control and the 16S rDNA markers.

	Sample 1 (manure)	Sample 2 (soil)	Sample 3 (soil and manure)	Sample 4 (soil and tiamulin)	Sample 5 (soil, manure and tiamulin)
Raw reads	81,186,214	29,798,588	38,364,982	40,414,636	40,464,585
Total reads in analysis	78,722,636	28,703,406	37,042,909	39,027,850	39,252,588
16S markers	483,328	10,612	90,739	312,984	318,092

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

3.2 Elimination of 16S rDNA

As it is shown in the table 1 we still had 16S rDNA marker in or samples, even after the processing which keeps only the non-genomic DNA. But we didn't have sequenced a control to see if the 16S rDNA marker decreasing after the processing (testing was performed via qPCR in the DNA extracts that showed a significant 16S rRNA gene reduction, yet, not its elimination). The results of qPCR performed on the samples before and after the use of endonuclease showed that the reduction of rDNA was 75-90%.

3.3 Alpha diversity analysis

All samples were thoroughly investigated as far as the analysed database markers are concerned according to the results of the Good's coverage estimate. According to results of the Good's coverage estimate, all the samples were representative of the community (more information about Good's coverage estimate is available in the section "6.1 Supplementary I").

Bellow, a barplot is provided (Figure 10) showing the differences between the analysed samples in terms of the Shannon index of ARGs according to the deepARG database screening (the corresponding plots for other databases showed similar results and are provided in Supplementary I). The indices in sample 1 (manure) consistently provided the lowest values in comparison with the other samples. The other four samples differ in their deviations in each index and will be described in detail below, although in general the manure amended soil samples were always characterized by lower diversity compared to the manure-free soils, and the tiamulin receiving samples always show reduced diversity compared to those of the same substrate (e.g. soil+ tiamulin versus soil and soil+ manure+ tiamulin versus soil+ manure) that did not receive antibiotics.
According to the Shannon index for the ARG analysis using the DeepARG approach, sample 5 (soil, manure and tiamulin) has the second lowest gene diversity after sample 1 (manure), suggesting that manure and tiamulin act in combination and reduce gene diversity more than either of them alone [Figure 10]. Sample 2 (soil) has the highest diversity, while, as mentioned above, the addition of Tiamulin (sample 4) reduces the diversity of genes and the addition of manure (sample 3) reduces it even more. Similar results were obtained for the Simpson and inverse Simpson indices, these graphs are available in the section: "Supplementary I".

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

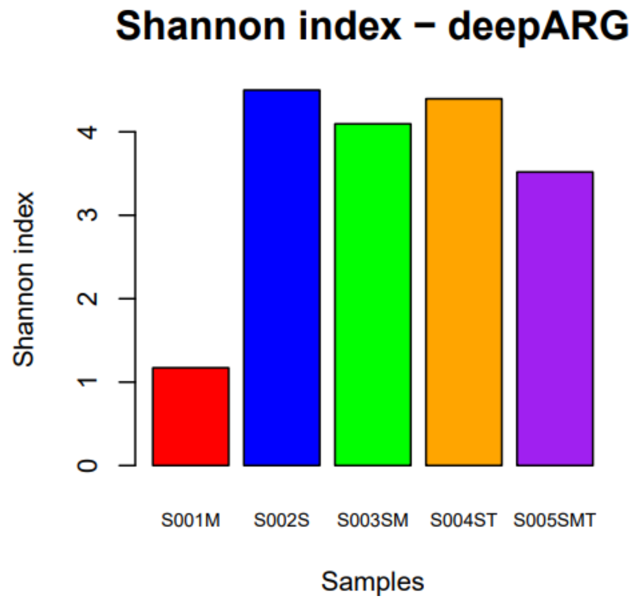


Figure 10 Barplot for the Shannon index according to deepARG database genes. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

However according to Fisher's α index for the same tested database and search algorithm, tiamulin increases diversity when added to soil (sample 4) and decreases it when added to soil mixed with manure (sample 5) [figure 11], while diversity reaches maximum levels when manure is added to soil without any antibiotic addition (sample 3). Here the reverse of the previous plot occurs, tiamulin increases the diversity of the soil, manure increases it even more, but both together synergistically reduce it.

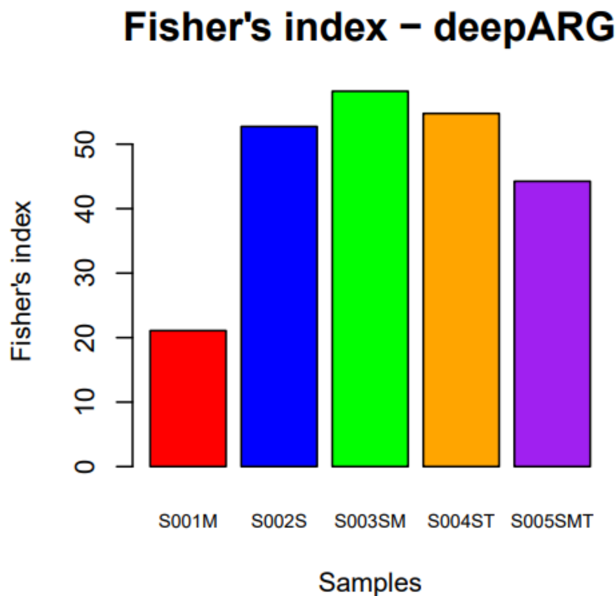


Figure 11 Barplot for the Fisher's α index according to deepARG database genes. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

According to the ACE estimate, the addition of tiamulin to the soil (sample 4) increases gene diversity and manure increases it even more (sample 3), while when they act cooperatively (sample 5) there is an intermediate increase in diversity (more than tiamulin alone but less than manure alone) [figure 12].

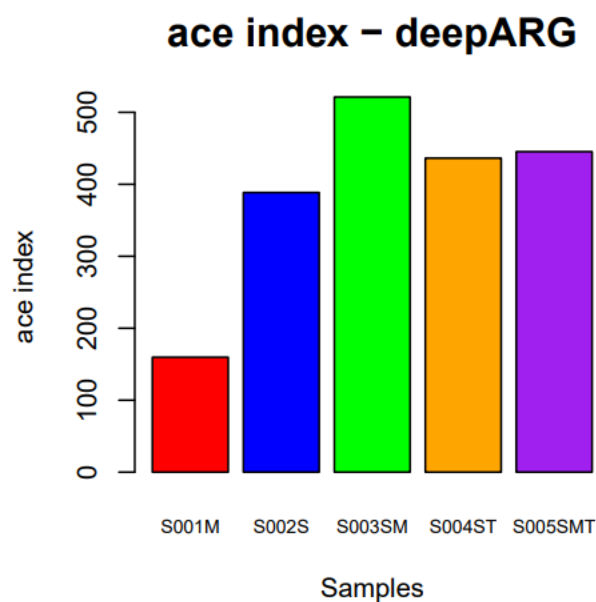


Figure 12 Barplot for the ACE index according to deepARG database genes. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

The observed richness provided quite similar patterns (but not identical) as the two previous indices. The addition of tiamulin to the soil (sample 4) increases the diversity, the addition of manure to the soil increases it even more, while their combination increases it less than each alone [Figure 13]. Similar results were obtained for the estimated richness index, this graph is available in the section: "Supplementary I".

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

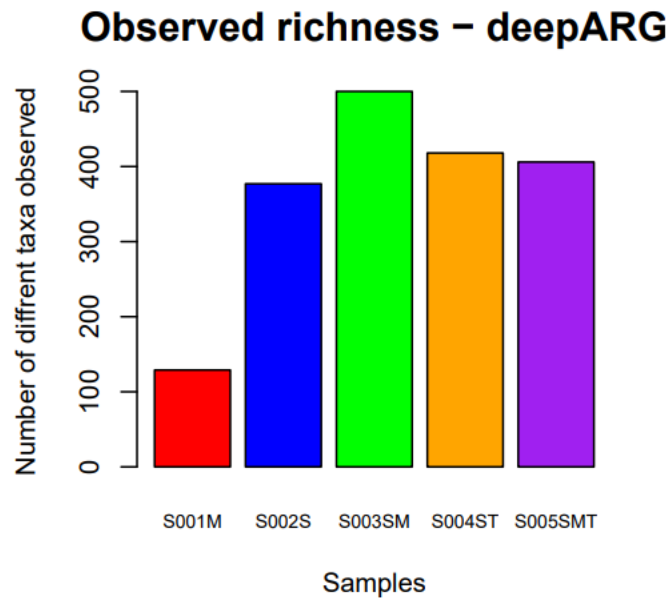


Figure 13 Barplot for observed richness index according to deepARG database genes. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

3.4 Beta diversity

3.4.1 General patterns

General patterns regarding clustering between samples are shown below in the NMDS plots. The NMDS diagram using genes from the deepARG database is given as an example [Figure 14] since the rest databases responded similarly (the corresponding plots for other databases are provided in Supplementary II) It is obvious that samples number 2 (soil) and 4 (soil and tiamulin) are grouped together, sample number 3 (soil and manure) is grouped with sample number 5 (soil, manure and tiamulin) and sample number 1 is separated from the rest. NMDS plots from the other databases are accessible in the section: "6.2 Supplementary II", but the trend is the same as in the deepARG database.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

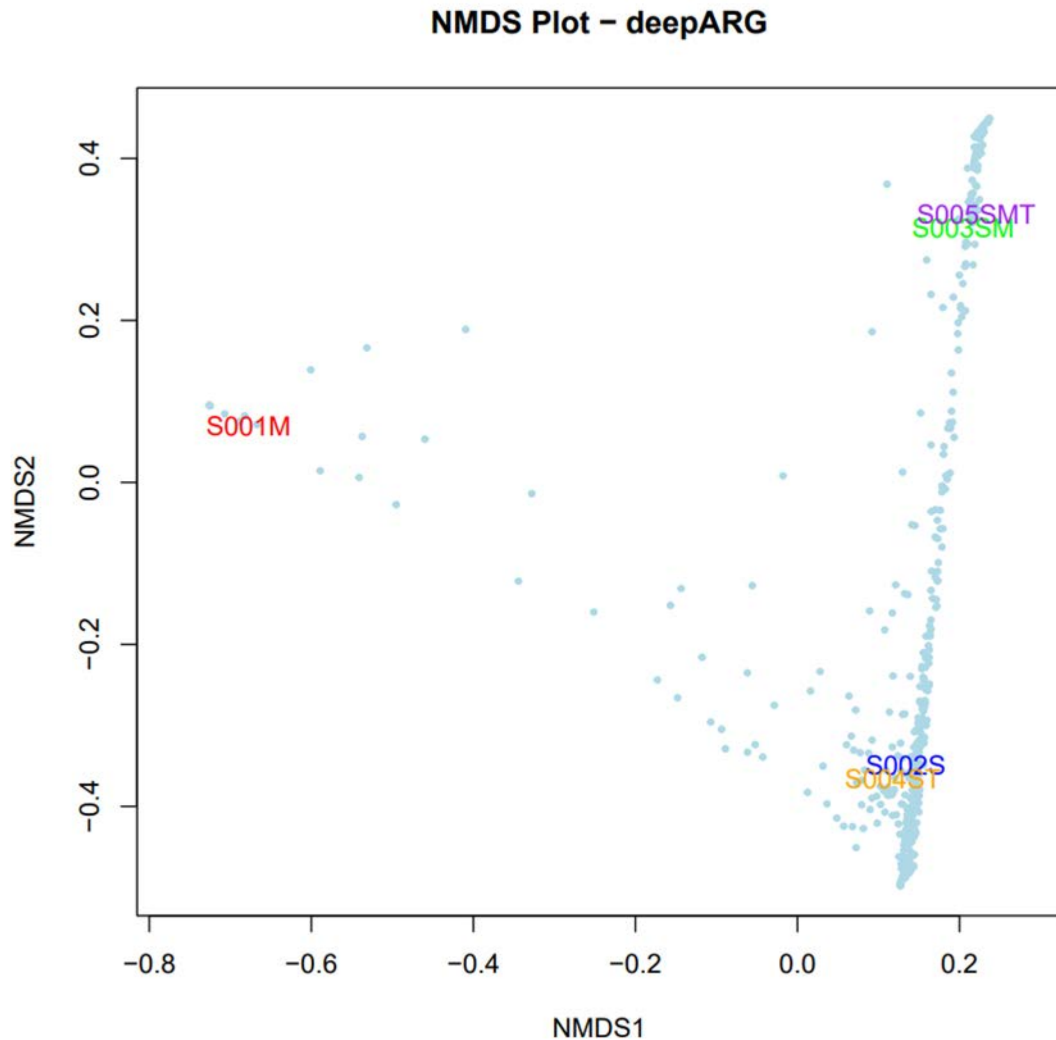


Figure 14 NMDS plot for deepARG database. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

3.4.2 Differential abundance of ARGs, MGEs and pathogenicity related marker genes

In order to support the effect of tiamulin on the microbial community, the prerequisite exists that (i) the abundance of reads must appear increased in sample 4 (soil and tiamulin) compared to sample 2 (soil), while, simultaneously, (ii) the reads must differ notably between high abundance in sample 5 (soil, manure and tiamulin) and lower abundance in samples 1 (manure), 2 (soil) and 3 (soil and manure). The main purpose is to clarify whether the addition of tiamulin results in an increase in specific genes associated with antibiotic resistance. Also, in order to examine the different effects caused by the addition of tiamulin in the case of soil as opposed to manured soil, genes with increased abundance in sample 4

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

(soil + tiamulin) compared to sample 2 (soil) and genes with increased abundance in sample 5 (soil + manure + tiamulin) compared to sample 3 (soil + manure) were observed. Out of the tested databases the deepARG and resfinder are the two concerning antibiotic resistance.

In the deepARG heatmap, the genes that are upregulated simultaneously in samples 4 and 5 are as follows: "*multidrug___major_facilitator_superfamily_transporter*", "*peptide___ugd*", "*unclassified___transcriptional_regulatory_protein_cpxr_cpxr*", "*mls___tlcc*", "*diaminopyrimidine___dfra14*", "*tetracycline___teth*" and "*multidrug___adej*" [Figure 15]. Also, the genes that are up-regulated in sample 4 compared to sample 2 are: "*nitroimidazole___MSBA*", "*multidrug___MAJOR_FACILITATOR_SUPERFAMILY_TRANSPORTER*", "*peptide___UGD*" , " *multidrug___SMEE*", " *bacitracin___BCRA*", "*multidrug___RPOB2*", "*MLS___TLCC*", "*unclassified___TRANSCRIPTIONAL_REGULATORY_PROTEIN_CPXR_CPXR*", "*bacitracin___BACA*", "*rifamycin___RPHB*", "*glycopeptide___VANR*", "*unclassified___TRUNCATED_PUTATIVE_RESPONSE_REGULATOR_ARLR*", "*rifamycin___RPHA*", "*beta−lactam___TEM*", "*fluoroquinolone___PATA*", "*tetracycline___TETA(48)*", " *glycopeptide___VANS*", "*diaminopyrimidine___DFRA14*", "*tetracycline___TETH*" and "*multidrug___ADEJ*".

While, the genes that are up-regulated in sample 5 compared to sample 3 are: "*multidrug___MAJOR_FACILITATOR_SUPERFAMILY_TRANSPORTER*", "*peptide___UGD*", "*multidrug___ACRB*", "*fosmidomycin___ROSA*", "*MLS___TLCC*", "*unclassified___TRANSCRIPTIONAL_REGULATORY_PROTEIN_CPXR_CPXR*", "*MLS___MACA*", "*multidrug___MEXW*", "*multidrug___MEXA*", "*aminoglycoside___APH(6)−I*", "*aminoglycoside___AADA*", "*diaminopyrimidine___DFRA14*", "*sulfonamide___SUL2*", "*aminoglycoside___APH(3'')−I*", "*multidrug___ACRA*", "*multidrug___MDTE*", "*tetracycline___TETH*", "*MLS___LNUG*", "*tetracycline___TETR*", "*multidrug___ADEJ*", "*MLS___LPEB*", "*aminoglycoside___KASUGAMYCIN_RESISTANCE_PROTEIN_KSGA*",

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

"*multidrug___MDTF*", "*multidrug___ACRF*" and "*bicyclomycin___BCR−1*".

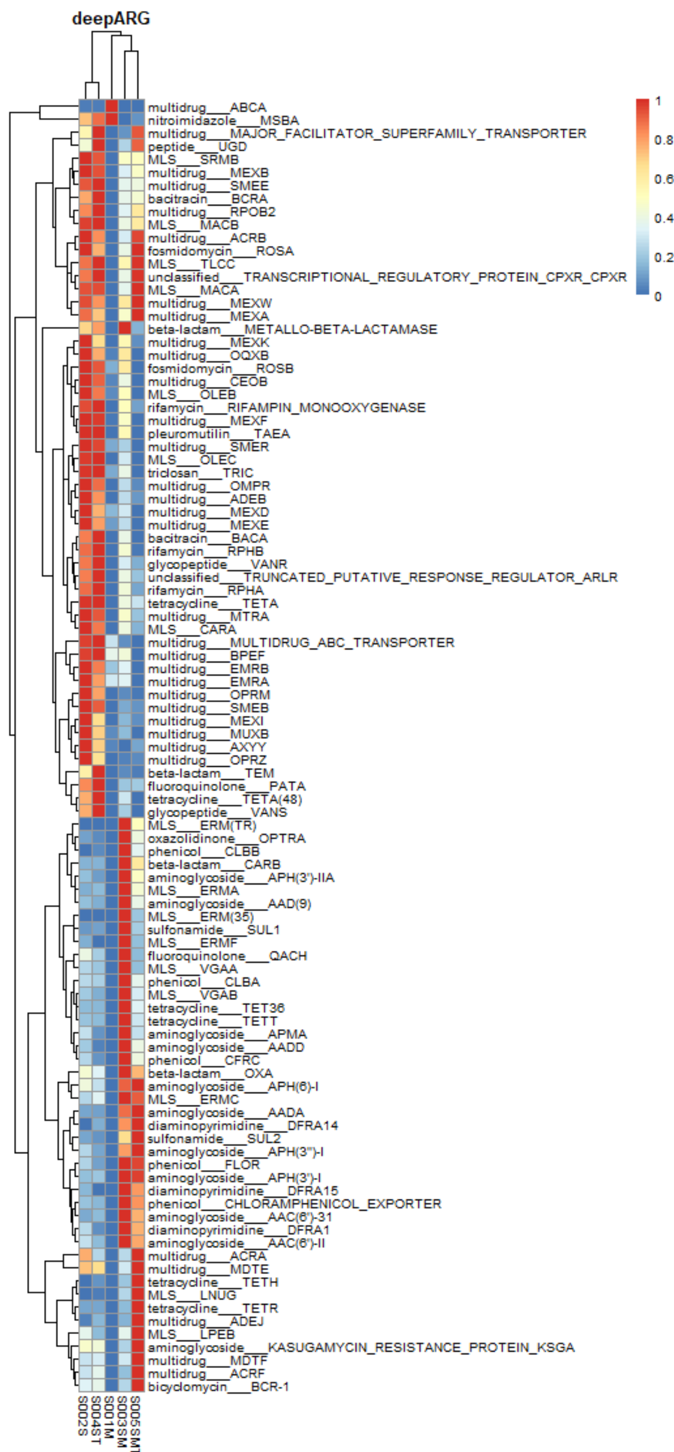


Figure 15 heatmap for deepARG. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

From the resfinder database, only the genes "*folate pathway antagonist_____dfra14*" and "*tetracycline___Increased efflux__tet(h)*" are co-upregulated simultaneously in samples 4 and 5 [Figure 16].

Also, the genes that are up-regulated in sample 4 compared to sample 2 are:

"Tetracycline___Increased efflux__tet(L)", "Beta−lactam___Enzymatic inactivation__blaTEM−178", "Beta−lactam___Enzymatic inactivation__blaTEM−207", "Lincosamide, Streptogramin A___Target protection__lsa(B)", "Aminoglycoside___Enzymatic modification__str", "Beta−lactam___Enzymatic inactivation__blaTEM−205", "Beta−lactam_____blaL1", "Quinolone___Increased efflux__qepA4", "Beta−lactam___Enzymatic inactivation__blaTEM−102", "Beta−lactam___Enzymatic inactivation__blaTEM−187", "Beta−lactam___Enzymatic inactivation__blaTEM−22", "Beta−lactam___Enzymatic inactivation__blaTEM−21", "Tetracycline___Increased efflux__tet(V)", "Beta−lactam___Enzymatic inactivation__blaTEM−116", "Beta−lactam___Enzymatic inactivation__blaTEM−123", "Quinolone___Increased efflux__qepA2", "Beta−lactam___Enzymatic inactivation__blaTEM−1A", "Tetracycline___Increased efflux__tetB(60)", "Aminoglycoside___Enzymatic modification__aac(3)−IV", "Aminoglycoside___Enzymatic modification__aph(3')−IIIa", "Tetracycline___Increased efflux__tet(H)", "Aminoglycoside___Enzymatic modification__aadA1", "Folate pathway antagonist_____dfrA14", "Lincosamide, Streptogramin A, Pleuromutilin___Target protection__lsa(E)", "Aminoglycoside___Enzymatic modification__ant(3")−Ia", "Tetracycline___Increased efflux__tet(Z)" and *"Aminoglycoside___Enzymatic modification__aadA1b".*

While, the genes that are up-regulated in sample 5 compared to sample 3 are: *"Aminoglycoside___Enzymatic modification__aph(3")−Ib", "Tetracycline___Increased efflux__tet(H)", "Lincosamide___Enzymatic inactivation__lnu(G)", "Tetracycline___Increased efflux__tet(31)", "Aminoglycoside___Enzymatic modification__ant(6)−Ia", "Folate pathway antagonist_____dfrA14", "Lincosamide___Enzymatic inactivation__lnu(F)", "Aminoglycoside___Enzymatic modification__aadA2b", "Folate pathway antagonist_____sul2", "Aminoglycoside___Enzymatic modification__aadA14"* and *"Aminoglycoside___Enzymatic modification__aph(6)−Id".*

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

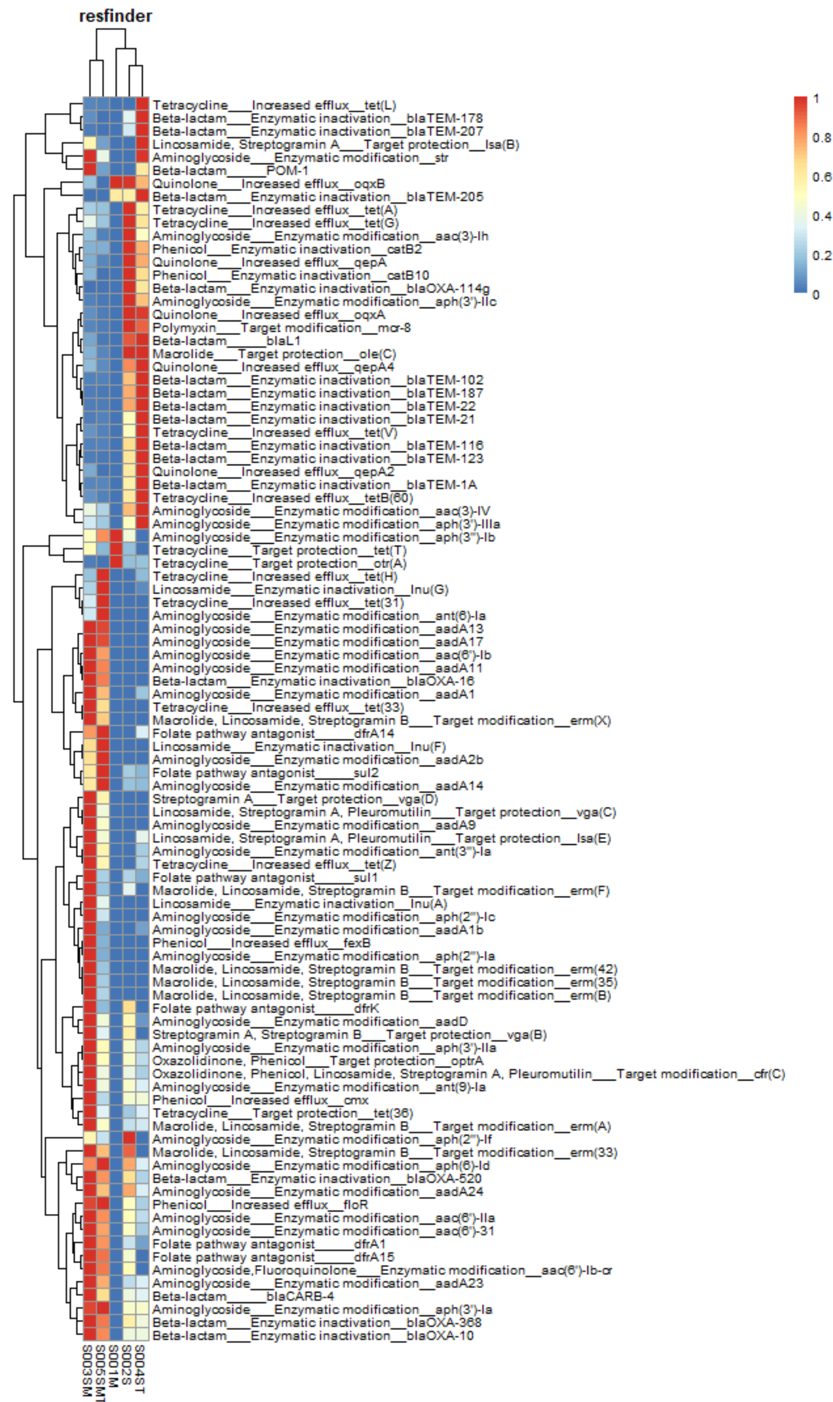


Figure 16 heatmap for resfinder. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

3.4.3 Other differences between the samples

There is one aromatic compound degradation gene that is upregulated in samples with tiamulin compered to control samples, this gene is the "*non____wp_010565636*" [Figure 17].

Also, the genes that are up-regulated in sample 4 compared to sample 2 are: *"Non____YP_004487058", "Non____YP_004155882", "Non____WP_020175902", "Non____YP_006381768", "Non____WP_010565636", "Non____WP_019391011", "Non____WP_016838133", "Non____WP_016427250", "Non____YP_006460639", "Non____WP_007087139", "Non____WP_008297703", "Non____WP_019463199", "Non____WP_007669962", "Non____WP_019521482", "Non____YP_007294207", "Non____WP_007755301", "Non____YP_994980", "Non____YP_001263552", "Non____WP_007569576", "Non____ZP_01011464", "Non____EPZ42916"* and *"Non____WP_017953351".*

While, the genes that are up-regulated in sample 5 compared to sample 3 are: *"Non____NP_246470", "Non____YP_003711113", "Non____WP_004751332", "Non____YP_284572", "Non____YP_001721188", "Hpc____CAB65144", "Non____ZP_05973964", "Non____WP_010565636", "Non____YP_002800264", "Non____YP_003611269", "Non____WP_018151422", "Non____ZP_06714995", "Hpc____AAO17180"* and *"Non____WP_017847306".*

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

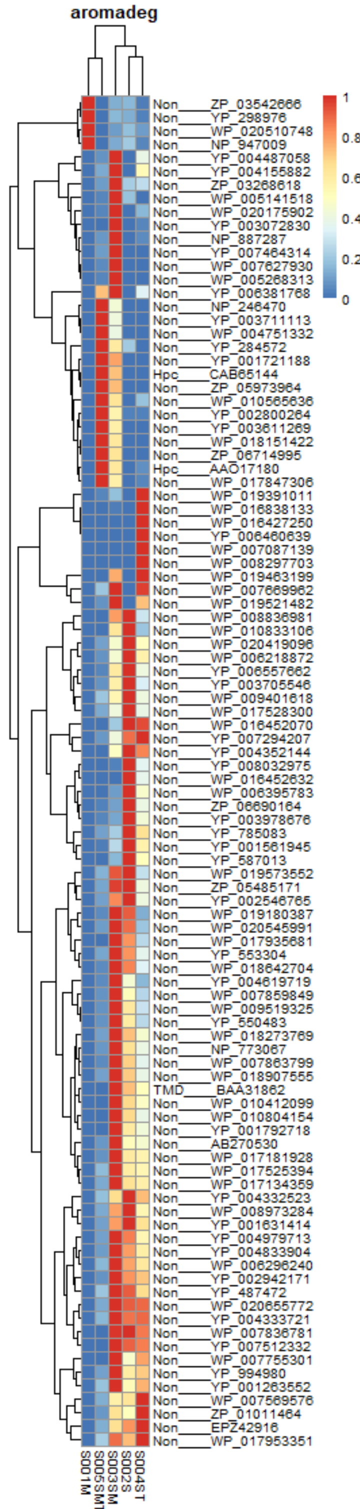


Figure 17 heatmap for aromadeg. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

An increased appearance of the disinfectant-resistant "*sodium deoxycholate (sdc)___evgA*" in the samples with tiamulin and the gene "*arsenic (as)___acr3*", which causes resistance to the metal arsenic, was detected [Figure 18]. The gene "*iron*

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

(fe), manganese (mn)___yfec" is also upregulated in samples in the presence of antibiotics.

Also, the genes that are up-regulated in sample 4 compared to sample 2 are: *"Triclosan, n−hexane, p−xylene___mexF", "Manganese (Mn), Iron (Fe), Cobalt (Co), Zinc (Zn), Nickel (Ni), Copper (Cu), Cadmium (Cd), Gallium (Ga)___fpvA", "Copper (Cu), Sodium acetate___actP", "Nickel (Ni), Cobalt (Co)___nrsD/nreB", "Arsenic (As)___acr3", "Arsenic (As)___arsC", "Copper (Cu), Zinc (Zn), Hydrochloric acid (HCl)___actA", "Arsenic (As)___arsM", "Iron (Fe), Hydrogen Peroxide (H2O2), Plumbagin___ideR", "Triclosan___fabL/ygaA", "Triclosan, n−hexane, p−xylene___mexD", "Chromium (Cr), Iron (Fe), Hydrogen Peroxide (H2O2), 2,6−dichloroindophenol___chrR", "n−hexane, p−xylene___oprM/oprK", "Manganese (Mn), Magnesium (Mg)___mntR", "Chromium (Cr), Methyl Viologen, Menadione___chrB", "Hydrogen Peroxide (H2O2), Benzylkonium Chloride (BAC), Chlorhexidine___cpxA", "Mercury (Hg)___merE", "Mercury (Hg)___merT", "Chromium (Cr)___chrA", "Benzylkonium Chloride (BAC)___norM/pmpM", "Hydrochloric acid (HCl)___gadC/xasA", "Sodium Deoxycholate (SDC)___evgA"* and *"Iron (Fe), Manganese (Mn)___yfeC".*

While, the genes that are up-regulated in sample 5 compared to sample 3 are: *"Manganese (Mn), Iron (Fe), Cobalt (Co), Zinc (Zn), Nickel (Ni), Copper (Cu), Cadmium (Cd), Gallium (Ga)___fpv", "Copper (Cu)___copB", "Arsenic (As)___acr3", "Chromium (Cr), Methyl Viologen, Menadione___chrB", "Hydrogen Peroxide (H2O2), Benzylkonium Chloride (BAC), Chlorhexidine___cpxA", "Mercury (Hg)___merE", "Mercury (Hg)___merT", "Hydrochloric acid (HCl)___gadC/xasA", "Arsenic (As)___pstC", "Sodium Deoxycholate (SDC)___evgA", "Iron (Fe), Manganese (Mn)___yfeD", "Iron (Fe), Manganese (Mn)___yfeB", "Iron (Fe), Manganese (Mn)___yfeA", "Triton X−100___mtrF", "Ethidium Bromide___vexD"* and *"Iron (Fe), Manganese (Mn)___yfeC".*

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

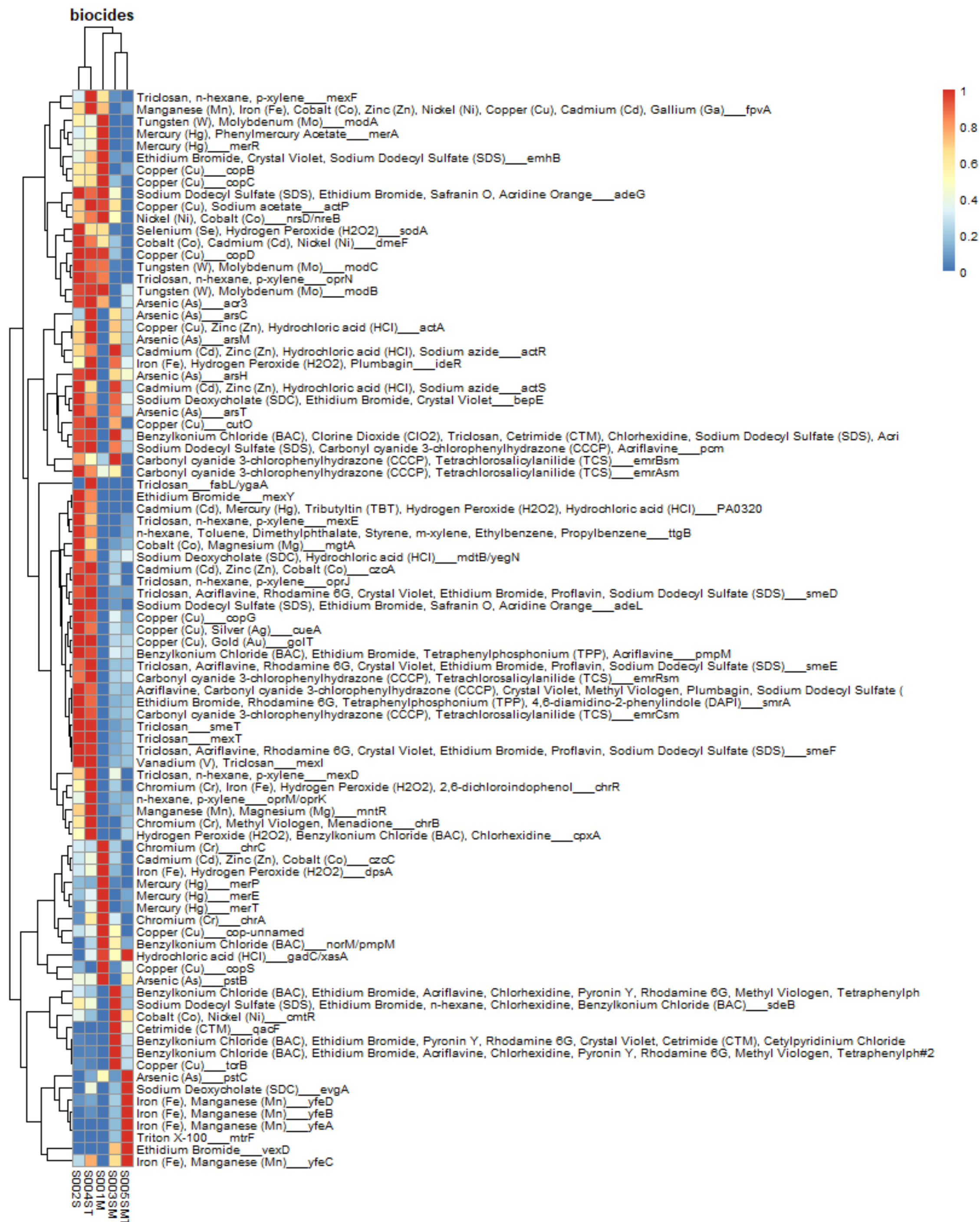


Figure 18 heatmap for biocides. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

In the MGE marker database, there was an increase in the occurrence of the gene "*integrase___integrase*" in the samples with tiamulin compared to the control samples. Also, the genes "*transposase___tn7_transposase_b_(plasmid)*" and "*transposase___multispecies:_heteromeric_transposase_endonuclease_subunit_tnsa*" are upregulated when tiamulin is present [Figure 19].

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Also, the genes that are up-regulated in sample 4 compared to sample 2 are:
"*transposase___IS3−like_element_ISGbe1_family_transposase*",
"*transposase___Tn3_family_transposase_ISMex22*",
"*transposase___MULTISPECIES:_IS481_family_transposase*",
"*transposase___IS982_family_transposase*",
"*transposase___MULTISPECIES:_IS5/IS1182_family_transposase*",
"*recombinase___MULTISPECIES:_recombinase_RecR*",
"*recombinase___recombinase_family_protein*",
"*transposase___IS200/IS605_family_transposase*",
"*recombinase___recombinase_XerD", "recombinase___recombinase_XerC*",
"*integrase___site−specific_integrase*",
"*integrase___MULTISPECIES:_site−specific_integrase", "integrase___integrase*",
"*transposase___IS4_family_transposase*",
"*transposon___TraG_family_conjugative_transposon_ATPase*",
"*transposase___IS1182_family_transposase*",
"*recombinase___MULTISPECIES:_recombinase_family_protein*",
"*transposase___ISAs1_family_transposase*",
"*Transposase___Tn7_transposase_B_(plasmid)*" and
"*transposase___MULTISPECIES:_heteromeric_transposase_endonuclease_subunit_TnsA*"

While, the genes that are up-regulated in sample 5 compared to sample 3 are:
"*integrase___integrase*", "*transposase___IS1595_family_transposase*",
"*transposase___IS3_family_transposase*",
"*Integrase___MULTISPECIES:_integrase*",
"*Transposase___heteromeric_transposase_endonuclease_subunit_TnsA*",
"*transposase___MULTISPECIES:_ISL3−like_element_IS1411_family_transposase*",
"*transposase___Tn7_transposase_B_(plasmid)*",
"*transposase___IS1_family_transposase*" and

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

"*transposase___MULTISPECIES:_heteromeric_transposase_endonuclease_subunit_TnsA*".

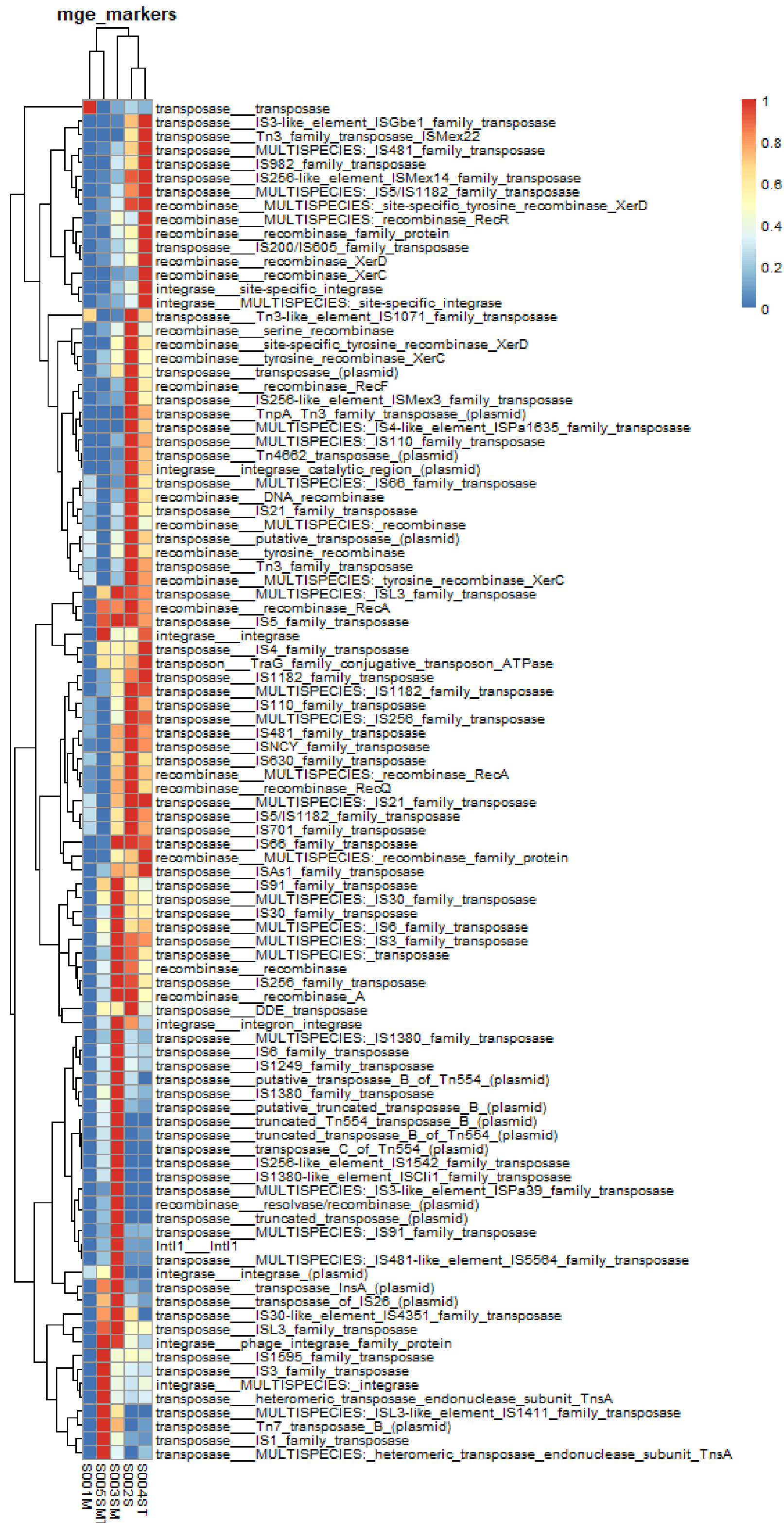


Figure 19 heatmap for MGE markers. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Also, some virulence genes appear to be increased by the addition of tiamulin, those genes are:
“*lps_o−antigen_p._aeruginosa____probable_acetyltransferase_wbpd___lps_o−antigen_p._aeruginosa*”,
“*ype_iv_pili_ai097____camp−regulatory_protein___type_iv_pili_ai097*” and
“*lps_o−antigen_p._aeruginosa____nucleotide_sugar_epimerase/dehydratase_wbpm___lps_o−antigen_p._aeruginosa*” [Figure 20].

Also, the genes that are up-regulated in sample 4 compared to sample 2 are:
“T6SS_SS193____putative_type_VI_secretion_system_protein_TssF___T6SS_SS193”,“Hcp_secretion_island−1_encoded_type_VI_secretion_system_H−T6SS____putative_ABC_transport_system,_substrate−binding_protein___Hcp_secretion_island−1_encoded_type_VI_secretion_system_H−T6SS”,
“Alginate_VF0091____Alginate_regulatory_protein_AlgQ___Alginate_VF0091”,
“type_IV_pili_AI097____catabolite_repression_control_protein___type_IV_pili_AI097", “Capsule_CVF775____IS4_family_transposase_ORF_1___Capsule_CVF775”,
“HSI−I_VF0334____type_IV_secretion_associated_protein_TagR,_positively_regulates_PpkA___HSI−I_VF0334”,
“direct_heme_uptake_system_IA049____Heme−transport_protein,_PhuT___direct_heme_uptake_system_IA049”,
“Heme_utilization__CVF769____hypothetical_protein___Heme_utilization__CVF769”,
“Deoxyhexose_linking_sugar,_209_Da_capping_structure_AI138____two−component_sensor___Deoxyhexose_linking_sugar,_209_Da_capping_structure_AI138”,
“alcaligin_IA010____alcaligin_biosynthesis_enzyme___alcaligin_IA010”,
“Pyochelin_VF0095____dihydroaeruginoic_acid_synthetase_PchE___Pyochelin_VF0095”,
“ornibactin_IA036____iron−hydroxamate_transporter_permease_subunit___ornibactin_IA036”, “pyoverdine_IA001____dipeptidase_precursor___pyoverdine_IA001”,
“Capsule_CVF775____WecC_protein___Capsule_CVF775”,
“ferric_uptake_regulator____transcriptional_repressor_of_iron−responsive_genes_Fur_family___ferric_uptake_regulator”,
“Fibronectin−binding_protein_AI169____malate_synthase_G___Fibronectin−binding_protein_AI169”,
“Trw_type_IV_secretion_system_VF0372____Trw_type_IV_secretion_sysmtem_VirB4−like_trwK_protein___Trw_type_IV_secretion_system_VF0372”,
“T6SS_VF0569____type_VI_secretion_system_contractile_sheath_large_subunit_VipB___T6SS_VF0569”,
“GroELHsp60/Cpn60.2_AI353____chaperonin_GroEL___GroELHsp60/Cpn60.2_AI353”,
“LPS_O−antigen_P._aeruginosa____probable_acetyltransferase_WbpD___LPS_O−antigen_P._aeruginosa”,
“type_IV_pili_AI097____cAMP−regulatory_protein___type_IV_pili_AI097”,
“Capsule_CVF775____VI_polysaccharide_biosynthesis_protein_VipB/tviC___Capsule_CVF775”,
“ClpP_VF0074____ATP−dependent_Clp_protease_proteolytic_subunit___ClpP_VF0074”,
“Phenazines_biosynthesis_CVF536____phenazine_biosynthesis_protein_PhzF,_isomerase___Phenazines_biosynthesis_CVF536”,
“Bcp_pili_AI128____collagen_adhesion_protein___Bcp_pili_AI128”,

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

“Fibronectin−binding_protein_AI230____outer_membrane_protein,_oprE3___Fibronectin−binding_protein_AI230”, *“Hcp_secretion_island−1_encoded_type_VI_secretion_system_H−T6SS____putative_ABC_transport_system,_membrane_protein___Hcp_secretion_island−1_encoded_type_VI_secretion_system_H−T6SS”,* *“ShlA−ShlB_SS130____hemolysin_determinant___ShlA−ShlB_SS130”,* *“Fibronectin−binding_protein_AI182____ebhA_protein___Fibronectin−binding_protein_AI182”, “HSI−2_SS179____hypothetical_protein___HSI−2_SS179”,* *“MgtBC_VF0106____Mg2+_transport_protein___MgtBC_VF0106”,* *“LPS_O−antigen_P._aeruginosa____nucleotide_sugar_epimerase/dehydratase_WbpM___LPS_O−antigen_P._aeruginosa”,* *“HSI−2_SS179____ClpA/B−type_protease___HSI−2_SS179”,* *“HSI−I_VF0334____type_VI_secretion_system_tubule−forming_protein_VipB___HSI−I_VF0334”, “GroELHsp60____60_kDa_chaperonin___GroELHsp60”,* *“Imp_SS195____hypothetical_protein___Imp_SS195”,* *“LPS_CVF383____acyl_carrier_protein___LPS_CVF383”,* *“Hcp_secretion_island−1_encoded_type_VI_secretion_system_H−T6SS____hypothetical_protein___Hcp_secretion_island−1_encoded_type_VI_secretion_system_H−T6SS”* and *“ClpC_VF0072____endopeptidase_Clp_ATP−binding_chain_C___ClpC_VF0072”.*

While, the genes that are up-regulated in sample 5 compared to sample 3 are: *“Heme_biosynthesis_CVF506____glutamate−1−semialdehyde−2,1−aminomutase___Heme_biosynthesis_CVF506”,* *“Capsule_CVF775____Phosphomannomutase___Capsule_CVF775”,* *“Capsule_CVF775____UDP−glucose_4−epimerase___Capsule_CVF775”,* *“LPS_O−antigen_P._aeruginosa____probable_UDP−glucose/GDP−mannose_dehydrogenase_WbpA___LPS_O−antigen_P._aeruginosa”,* *“Capsule_CVF775____WbbJ_protein___Capsule_CVF775”,* *“T6SS_SS194____Type_VI_secretion_system_contractile_sheath_large_subunit_TssC/VipB___T6SS_SS194”,* *“Capsule_CVF775____bifunctional_UDP−N−acetylglucosamine_pyrophosphorylase/glucosamine−1−phosphate_N−acetyltransferase___Capsule_CVF775”,* *“Capsule_CVF775____hypothetical_protein___Capsule_CVF775”,* *“LPS_O−antigen_P._aeruginosa____probable_acetyltransferase_WbpD___LPS_O−antigen_P._aeruginosa”,* *“Proteobactin_IA034____ABC_transporter_permease___Proteobactin_IA034”,* *“LPS_O−antigen_P._aeruginosa____probable_aminotransferase_WbpE___LPS_O−antigen_P._aeruginosa”,* *“EF−Tu_VF0460____elongation_factor_Tu___EF−Tu_VF0460”,* *“Hsp60_VF0159____Hsp60,_60K_heat_shock_protein_HtpB___Hsp60_VF0159”,* *“Capsule_CVF775____VI_polysaccharide_biosynthesis_protein_VipB/tviC___Capsule_CVF775”,* *“Capsule_CVF775____nucleoside−diphosphate_sugar_epimerase___Capsule_CVF775”, “Capsule_CVF775____WecE_protein___Capsule_CVF775”* and *“LPS_O−antigen_P._aeruginosa____nucleotide_sugar_epimerase/dehydratase_WbpM___LPS_O−antigen_P._aeruginosa”.*

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

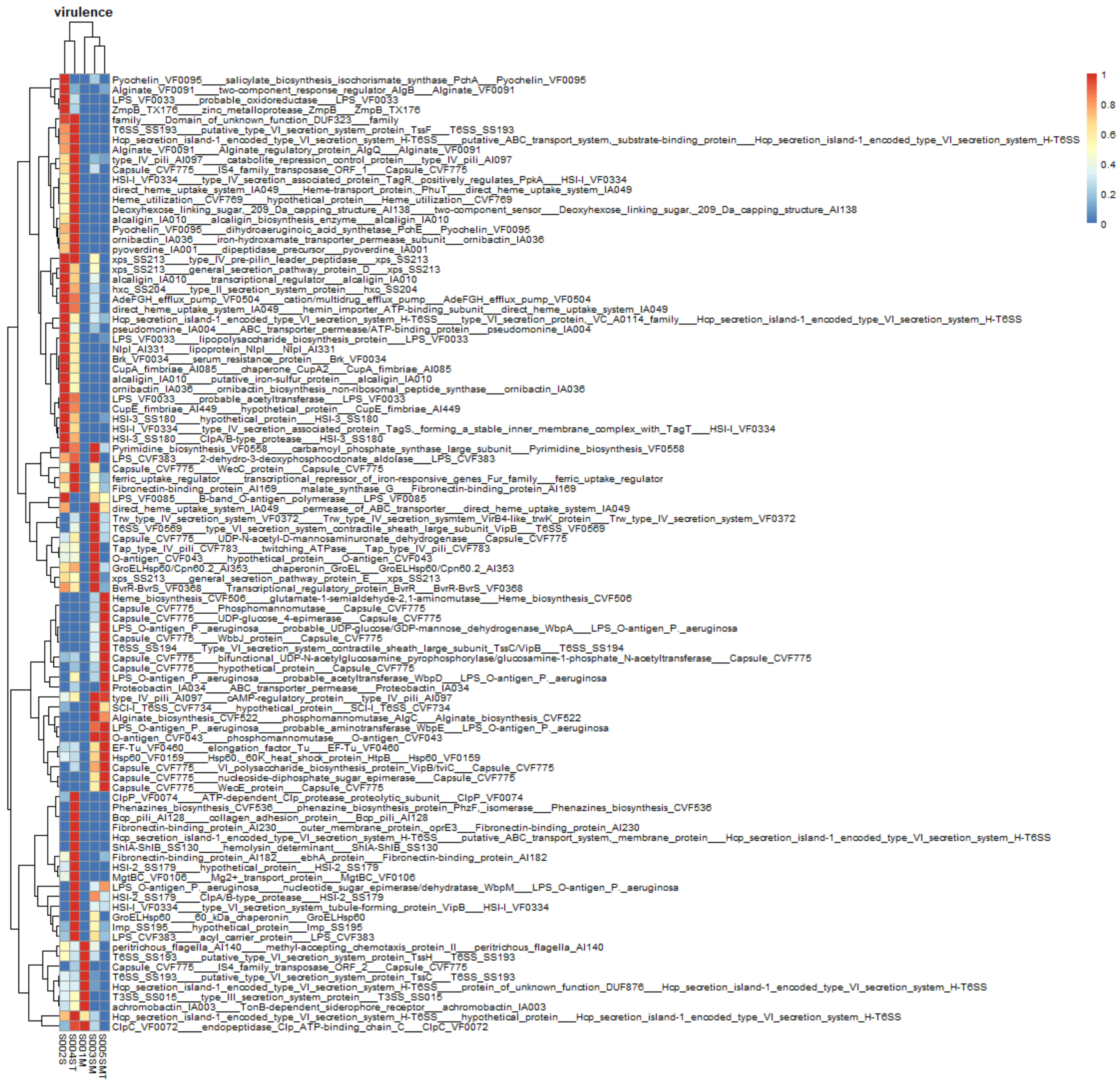


Figure 20 heatmap for virulence. There are 5 different samples: Sample 1 (manure, S001M), sample 2(soil, S002S), sample 3(soil and manure, S003SM), sample 4(soil and tiamulin, S004ST) and sample 5(soil, manure and tiamulin, S005SMT).

Heatmaps from the other databases, with no significant differences for the samples 2 and 4, can be found in “Supplementary III”.

4. Discussion

The alpha diversity indices presented refer to the deepARG database, while the remaining alpha diversity indices for each database are available in Supplementary I. According to these indices, the sample 1 (manure) is obviously characterised by lower variety of resistance genes, a result congruent for all indices, even if the reads are more than twice as many as

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

the other samples. This is due to the fact that in the origin of the manure, the pig's gut, low variation in the environmental conditions can be observed compared to soil. The animal's digestive tract has relatively stable temperature, pH, minerals and nutrients, so it promotes the development of specific microbes that grow well in these conditions and have less gene diversity. The mixture of manure with soil has significantly more resistance genes than the manure itself, confirming both that the addition of soil has a significant effect on the microbial community and, also, that there is the greatest amount and variety of resistance genes in the soil. The ratio of the index values between the other samples (sample 2, 3, 4 and 5) varied for the different alpha indices and the full explanation is given in the results section on alpha indices. The different patterns followed by the different alpha diversity indices are probably due to the fact that they represent marker features of the community of different dominance level, as has been reported the Shannon index represents the least abundant genes, the Simpson and inverse Simpson indices are more representative of the most dominant genes, while the Fisher alpha index is indicative of highly dominant genes (Jost L., 2006). In any case, the common denominator is the greater effect of manure compared to the antibiotic. It is also noteworthy that, when the Fisher's α index is concerned, tiamulin increases the alpha diversity of resistance genes in the soil that did not receive manure while in the case of the soil that received manure the opposite happens in a more pronounced manner. This means that in the presence of an antibiotic, the highly dominant resistance genes increase when manure is absent from the soil, while when manure is present the highly dominant resistance genes are reduced. Finally, the Shannon, Simpson and inverse Simpson indices provide the same results between them, as tiamulin reduces diversity of less/medium abundant resistance genes due to its toxic effect on some microbes, and manure reduces it even more as it has limited range of microorganisms that can be expected to dominate over soil microbes (and their genes). Let's not forget that we are looking at resistance genes in general and not just those targeted by tiamulin, if tiamulin kills microorganisms containing resistance genes to other antibiotics then the effects of tiamulin will be less than the effects of other antibiotics.

As shown in the NMDS plot, samples 2 (soil) and 4 (soil and tiamulin) are clustering together, sample 3 (soil and manure) is grouped with sample 5(soil, manure and tiamulin) and sample 1 is separate from rest. Thus, we can assume that the effect of manure is stronger than the effect of tiamulin. This makes sense because when adding manure, a plethora of different microorganisms end up in the soil, causing changes in the community, its genome and its plasmidome. Also, microbes can produce antibiotics, and this is a pressure factor for the

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

spread of antibiotic resistance genes. In addition, tiamulin only kills the non- resistant microbes, but if resistance genes were *a priori* present in the community, they are likely to be transferred between microbes. It would be useful to keep in mind that resistance to tiamulin is a rare event (Krasucka et al. 2010), as we have mentioned in the introduction (section: "1.3.3 antibiotic resistance").

In the deepARG and resfinder databases, the majority of nine simultaneously upregulated genes in the samples with the tiamulin compared to the controls, have biological functions that can support their increasing after the tiamulin use.

As previously mentioned in the deepARG heatmap, the genes simultaneously upregulated in samples 4 and 5 are : "*multidrug___major_facilitator_superfamily_transporter*", "*peptide___ugd*", "*unclassified___transcriptional_regulatory_protein_cpxr_cpxr*" and "*mls___tlcc*".

The "*multidrug___major_facilitator_superfamily_transporter*", refers to the major facilitator superfamily (MFS), a family of membrane transporters present in bacteria (Pao, Paulsen, & Saier, 1998). One of their targets is the antibiotic tetracycline, which inhibits the 30S subunit of bacterial ribosomes (Roberts M. C., 2005).

The product of "*peptide___ugd*" synthesizes and transports the 4-amino-4-deoxy-L-arabinose (Ara4N) to Lipid A, which results in antimicrobial resistance (Lee et al. 2004). This antibiotic resistance mechanism is relevant against peptide antibiotics and polymyxin-like antibiotics. Both these two classes of antibiotics act through disruption of the bacterial cell membrane (Axelsen P. H., 2008, Breazeale, Ribeiro & Raetz, 2002).

The "*unclassified___transcriptional_regulatory_protein_cpxr_cpxr"* gene codes for the CpxR protein activating the RND efflux pump MexAB-OprM, which has the impact of drug resistance (Tian et al., 2016). As has been mentioned Tiamulin resistance was demonstrated to be associated with: (i) mutations that alter the chemical affinity of tiamulin with the ribosomal target site; (ii) ribosome detoxification mechanisms with pumps that remove tiamulin from the ribosomes; (iii) cell detoxification mechanisms that remove tiamulin form the cell environment (Killeavy, Jogl & Gregory, 2020; Nagy et al., 2023; Bøsling et al., 2003).

The macrolide-lincosamide-streptogramin (MLS) antibiotic group is an important group of translation inhibitors that act on the 50S ribosomal subunit (Tsui et al. 2004). Thus, "mls___tlcc" and "*mls___macb*" are putative antibiotic resistance genes against MLS

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

antibiotics. Although these antibiotics also target on the large subunit of the ribosome like tiamulin, there is no study linking their resistance genes to resistance to tiamulin.

The "*multidrug___adej*" gene is a membrane transporter that confers antibiotic resistance against multiple drugs, some of which target the 30S subunit of ribosome, such as tetracycline (Damier-Piolle et al. 2008, LaPlante et al. 2022). This mechanism is relevant since multidrug resistance transporters, in several cases, include also the transport of pleuromutilins (Schwarz et al., 2016). However, tetracyclines are not necessarily co-transported with pleuromutilins, also the small rRNA subunit that tetracyclines target is different from the large rRNA subunit that tiamulin targets.

From the resfinder database the genes "*folate pathway antagonist_____dfra14*" and the "*tetracycline___increased efflux__tet(h)*" were found to be increased simultaneously in samples 4 and 5. DfrA14 is a dihydrofolate reductase and their resistance mechanism is based on antibiotic target replacement (Márquez et al., 2008). Tetracycline, like tiamulin, inhibits bacterial ribosomal function by binding at the to the 23S rRNA peptidyl transferase centre (LaPlante et al. 2022, Killeavy, Jogl & Gregory, 2020). The "*diaminopyrimidine___dfra14*" and "*tetracycline___teth*" genes also showed interest in the deepARG database, where they also appeared increased in samples containing tiamulin, and they possess similar activity to the two previous genes from resfinder.

Interestingly, we observed an increase of the "*integrase___integrase*" gene in samples with tiamulin compared to the control samples. This MGE marker gene codes for an integrase enzyme which can insert foreign DNA into the bacterial chromosome (Maertens, Engelman, & Cherepanov, 2022). Similarly, the two transposases present in tiamulin samples have similar mechanisms for breaking and rejoining the DNA chain (Hickman & Dyda, 2015).

The addition of tiamulin showed no differences in microbe taxonomy (as shown in detail in the picture of the Supplementary III section) but the genes between samples differed greatly (as discussed at length in Sections 3.4.2 and 3.4.3). This suggests that the antibiotic may have an effect on the microbial community even though its use does not drastically change its composition. Certainly, the use of the focus-mobilome method during sample processing played an important role in these results, as several 16S rRNA genes and bacterial chromosomal material were removed from the samples and this apparently had an effect on the taxonomic analysis. Even if the method we used was not very successful in the removal of chromosomal DNA, it potentially significantly altered the chromosomal DNA presence. The

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

antibiotic is a selective agent between microbes but as we have seen in previous studies of the group the presence of manure strongly influences the community composition (Katsivelou et al. 2023), even in this analysis it is clear that manure influences the microbial community to a greater extent than tiamulin. Nevertheless, important controls, containing the untreated cells would have provided further insights about the method success. The observation of rDNA reduction was based on qPCR, which generally showed the amount of the gene in the samples, ideally, we should also run shotgun analysis on samples that have not been digested with exonuclease to see the qualitative changes caused by the removal of chromosomal replicons. Therefore, it would be advisable next time to use appropriate controls to evaluate the success of the method. The single aromatic compound degradation gene that can be inferred to be induced by tiamulin, as it was differentially present in the samples with tiamulin compared to other samples, could possibly belong to general degraders of aromatic compounds, whose abundance was increased by tiamulin amendment. An assumption could be that tiamulin amendment could have eliminated competitors of these degraders, something that can explain their higher abundance in these samples. For example, if two bacteria coexist in the soil microbial community that are competitive with each other (e.g. due to space limitations), one of which degrades an aromatic compound and the other one happens to be sensitive to tiamulin, then after the application of the antibiotic the abundance of the bacterium that degrades the aromatic compound will increase as its competition will have decreased. This will have the effect of appearing to increase the abundance of the aromatic compound degradation gene when the antibiotic is present. However, this, and other assumptions, need to be investigated further, as correlation does not immediately equal causation. If this case is not random, two further hypotheses are that (i) this specific aromatic degradation gene is present in the same plasmid as the antibiotic resistance genes, or (ii) the presence of this gene maybe help the effect of the antibiotic.

In addition, regarding the heavy metal resistance genes, iron and manganese resistance genes were found in increased abundance in antibiotic amended samples. This is congruent with recent studies which have highlighted co-resistance against heavy metals and antibiotics (Edet, Bassey & Joseph, 2023). But let's not forget that both antibiotics and heavy metals are stressors, so the coexistence of their resistance genes may be associated with generic stress responses of bacterial cells in the community. Some strains (e.g. strain ZC255) that have high resistance to stress in general also have co-tolerance to heavy metals and antibiotics (Zhu et al., 2023).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Genes, that, in one way or another, are associated with the ribosome, are likely to be increased due to tiamulin, which, as mentioned, is also associated with the 50S ribosomal subunit. Also, it may affect efflux pumps, because of their influence on entry or exit of compounds from the cell.

As we understood during the processing, there is no ideal method that can be specifically used to normalize the data for this dataset. For example, we could not use the 16S rRNA genes for this purpose, as our wet-lab approach was supposed to eliminate chromosomal genetic elements that carry the vast majority of 16S rDNA genes. We then realised that not all 16S rRNA gene sequences were eliminated. In that case we could have used the 16S rRNA gene copies for normalization purposes, only if we knew that the level of performance of the chromosomal DNA depletion was consistent among samples. On the other hand, in the case of the centered log-ratio transformation, the logarithm implementation masks the differences to a degree that only pronounced differences are apparent. Finally, lack of replication prevented us from using non-parametric approach based statistical tests. Given the above, we didn't normalize the data, and instead just used their relative abundances for the performance of the presented statistical tests, while we also emphasised on the descriptive presentation of the data.

From the above results we can deduce a possible effect of the antibiotic tiamulin on the mobilome, or more generally on the gene composition of the community. An interesting finding is the existence of genes whose abundance is increased in the presence of Tiamulin. Resistance genes such as *"multidrug___major_facilitator_superfamily_transporter", "peptide___ugd", "unclassified___transcriptional_regulatory_protein_cpxr_cpxr", "mls___tlcc", "diaminopyrimidine___dfra14" and "tetracycline___teth"* were found to increase, most of which are known to confer resistance to antibiotics that have a similar target to Tiamulin (30S ribosomal subunit). Interest was also found in genes such as integrase that helps in the transfer of genetic elements, aromatic compound degradation gene and heavy metal resistance genes. Although the greater influence of manure on the microbial community compared to the antibiotic is evident, this is in agreement with the results of previous studies of the group (Katsivelou et al., 2023).

This experiment has demonstrated the power of shotgun metagenomics in mobilome analysis. It also resurfaced the weaknesses of the current setup. In follow-up studies, it is strongly recommended to include as many controls as possible in the shotgun sequencing analysis, in order to validate the reduction of chromosomal DNA both qualitatively, next to

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

the quantitative approach of the qPCR. Moreover, more antibiotics could be examined, individually or in combination, regarding their possible effect on the microbial community composition and the gene pool. Another possible approach is to take samples directly from the environment, i.e. crop soils treated with pig manure that have been administered tiamulin, in order to study whether the results of this analysis are the same as those found under laboratory conditions, as it is likely that under real cultivation conditions and different soils the influence of the antibiotic will differ. The presence of pathogenic micro-organisms that affect humans and the possibility of their uptake of ARGs could also be tested by parallel shotgun analysis. Another suggestion is to use manure containing antibiotics as a fertiliser in potted plants and then to control the micro-organisms that come into contact with the plant and the possibility of increasing the abundance of ARGs. Finally, follow-up analysis is expected to include tools that exploit the taxonomic dispersal of the identified marker genes with tools like e.g. metaCHIP (Song et al., 2019), in order to gain insights on the current knowledge status of the mobility of the genes identified in the current samples and infer associated risks possibly induced by TIA or manure.

5. References

- Abraham, E. P., & Chain, E. (1988). An enzyme from bacteria able to destroy penicillin. 1940. Reviews of infectious diseases, 10(4), 677–678.
- Alawi, M., Velasco-Torrijos, T. & Walsh, F. (2022). Plasmid-mediated antimicrobial resistance in drinking water. Environmental Advances, 8, 100191, 2666-7657. https://doi.org/10.1016/j.envadv.2022.100191.
- Andermann, T., Antonelli, A., Barrett, R. L., & Silvestro, D. (2022). Estimating Alpha, Beta, and Gamma Diversity Through Deep Learning. Frontiers in plant science, 13, 839407. https://doi.org/10.3389/fpls.2022.839407
- Anderson, M. J., Ellingsen, K. E., & McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. Ecology letters, 9(6), 683–693. https://doi.org/10.1111/j.1461-0248.2006.00926.x
- Anderson-Sprecher, R., Flatman, G. T., & Borgman, L. (1994). Environmental sampling: a brief review. Journal of exposure analysis and environmental epidemiology, 4(2), 115–131.
- Andrade C. (2020). Sample Size and its Importance in Research. Indian journal of psychological medicine, 42(1), 102–103. https://doi.org/10.4103/IJPSYM.IJPSYM_504_19

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Arango, G, Dai, D, Pruden, A, Vikesland, P, Heath, L, Zhang, L (2019). NanoARG: a web service for detecting and contextualizing antimicrobial resistance genes from nanopore-derived metagenomes. Microbiome 7, 88. https://doi.org/10.1186/s40168-019-0703-9
- Arango-Argoty, G., Garner, E., Pruden, A., Heath, L. S., Vikesland, P., & Zhang, L. (2018). DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. Microbiome, 6(1), 23. https://doi.org/10.1186/s40168-018-0401-z
- Aslam, B., Khurshid, M., Arshad, M. I., Muzammil, S., Rasool, M., Yasmeen, N., Shah, T., Chaudhry, T. H., Rasool, M. H., Shahid, A., Xueshan, X., & Baloch, Z. (2021). Antibiotic Resistance: One Health One World Outlook. Frontiers in Cellular and Infection Microbiology, 11, 771510. https://doi.org/10.3389/fcimb.2021.771510
- Axelsen P. H. (2008). A chaotic pore model of polypeptide antibiotic action. Biophysical journal, 94(5), 1549–1550. https://doi.org/10.1529/biophysj.107.124792
- 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., … Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. BMC genomics, 9, 75. https://doi.org/10.1186/1471-2164-9-75
- 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. https://doi.org/10.1371/journal.pone.0128036
- Bender, S. F., Wagg, C., & van der Heijden, M. G. A. (2016). An Underground Revolution: Biodiversity and Soil Ecological Engineering for Agricultural Sustainability. Trends in ecology & evolution, 31(6), 440–452. https://doi.org/10.1016/j.tree.2016.02.016
- Bendtsen, J. D., Nielsen, H., von Heijne, G., & Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. Journal of molecular biology, 340(4), 783–795. https://doi.org/10.1016/j.jmb.2004.05.028
- 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, 289-300
- Bland, C., Ramsey, T. L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N. C., & Hugenholtz, P. (2007). CRISPR recognition tool (CRT): a tool for automatic detection

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

of clustered regularly interspaced palindromic repeats. BMC bioinformatics, 8, 209. https://doi.org/10.1186/1471-2105-8-209

- Boers, S. A., Jansen, R., & Hays, J. P. (2019). Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology, 38(6), 1059–1070. https://doi.org/10.1007/s10096-019-03520-3
- Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R. L., Rebelo, A. R., Florensa, A. F., Fagelhauer, L., Chakraborty, T., Neumann, B., Werner, G., Bender, J. K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B. B., Malhotra-Kumar, S., … Aarestrup, F. M. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. The Journal of antimicrobial chemotherapy, 75(12), 3491–3500. https://doi.org/10.1093/jac/dkaa345
- Bøsling, J., Poulsen, S. M., Vester, B., & Long, K. S. (2003). Resistance to the peptidyl transferase inhibitor tiamulin caused by mutation of ribosomal protein l3. Antimicrobial agents and chemotherapy, 47(9), 2892–2896. https://doi.org/10.1128/AAC.47.9.2892-2896.2003
- Brady, A., & Salzberg, S. L. (2009). Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. Nature methods, 6(9), 673–676. https://doi.org/10.1038/nmeth.1358
- Breazeale, S. D., Ribeiro, A. A., & Raetz, C. R. (2002). Oxidative decarboxylation of UDP-glucuronic acid in extracts of polymyxin-resistant Escherichia coli. Origin of lipid a species modified with 4-amino-4-deoxy-L-arabinose. The Journal of biological chemistry, 277(4), 2886–2896. https://doi.org/10.1074/jbc.M109377200
- Browne, P. D., Kot, W., Jørgensen, T. S., & Hansen, L. H. (2020). The Mobilome: Metagenomic Analysis of Circular Plasmids, Viruses, and Other Extrachromosomal Elements. Methods in molecular biology (Clifton, N.J.), 2075, 253–264. https://doi.org/10.1007/978-1-4939-9877-7_18
- Buchfink, B, Xie, C, Huson, D H (2015). Fast and sensitive protein alignment using DIAMOND. Nat Methods 12, 59-60, https://doi.org/10.1038/nmeth.3176
- Bush, K., & Bradford, P. A. (2016). β-Lactams and β-Lactamase Inhibitors: An Overview. Cold Spring Harbor perspectives in medicine, 6(8), a025247. https://doi.org/10.1101/cshperspect.a025247

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Callahan, Benjamin J., Paul J. McMurdie, Michael J. Rosen, Andrew W. Han, Amy Jo A. Johnson, and Susan P. Holmes. 2016. "DADA2: High-Resolution Sample Inference from Illumina Amplicon Data." Nature Methods 13 (7): 581–83. https://doi.org/10.1038/nmeth.3869.
- Carr, V. R., Shkoporov, A., Hill, C., Mullany, P., & Moyes, D. L. (2021). Probing the Mobilome: Discoveries in the Dynamic Microbiome. Trends in microbiology, 29(2), 158–170. https://doi.org/10.1016/j.tim.2020.05.003
- Castillo, M., Martín-Orúe, S. M., Taylor-Pickard, J. A., Pérez, J. F., & Gasa, J. (2008). Use of mannanoligosaccharides and zinc chelate as growth promoters and diarrhea preventative in weaning pigs: Effects on microbiota and gut function. Journal of animal science, 86(1), 94–101. https://doi.org/10.2527/jas.2005-686
- Chaffin, D. O., & Rubens, C. E. (1998). Blue/white screening of recombinant plasmids in Gram-positive bacteria by interruption of alkaline phosphatase gene (phoZ) expression. Gene, 219(1-2), 91–99. https://doi.org/10.1016/s0378-1119(98)00396-5
- Chan, S., Ng, S., Chan, H. P., Pascoe, E. M., Playford, E. G., Wong, G., Chapman, J. R., Lim, W. H., Francis, R. S., Isbel, N. M., Campbell, S. B., Hawley, C. M., & Johnson, D. W. (2020). Perioperative antibiotics for preventing post-surgical site infections in solid organ transplant recipients. The Cochrane database of systematic reviews, 8(8), CD013209. https://doi.org/10.1002/14651858.CD013209.pub2
- Chao A. (1987). Estimating the population size for capture-recapture data with unequal catchability. Biometrics, 43(4), 783–791.
- Chen, Q-L, Cui, H-L, Su, J-Q, Penuelas, J, Zhu, Y-G (2019). Antibiotic resistomes in plant microbiomes. Trends Plant Sci 24, 530-541, https://doi.org/10.1016/j.tplants.2019.02.010
- Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics (Oxford, England), 34(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560
- Chen, Y., Lun, A. T., & Smyth, G. K. (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Research, 5, 1438. https://doi.org/10.12688/f1000research.8987.2
- Chevreux, B., Wetter, T., & Suhai, S. (1999). Genome Sequence Assembly Using Trace Signals and Additional Sequence Information Computer Science and Biology. Proceedings of the German Conference on Bioinformatics. 99, 45-56.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Cycoń, M., Mrozik, A., & Piotrowska-Seget, Z. (2019). Antibiotics in the Soil Environment-Degradation and Their Impact on Microbial Activity and Diversity. Frontiers in microbiology, 10, 338. https://doi.org/10.3389/fmicb.2019.00338
- Damier-Piolle, L., Magnet, S., Brémont, S., Lambert, T., & Courvalin, P. (2008). AdeIJK, a resistance-nodulation-cell division pump effluxing multiple antibiotics in Acinetobacter baumannii. Antimicrobial agents and chemotherapy, 52(2), 557–562. https://doi.org/10.1128/AAC.00732-07
- Delmont, T. O., Robe, P., Clark, I., Simonet, P., & Vogel, T. M. (2011). Metagenomic comparison of direct and indirect soil DNA extraction approaches. Journal of microbiological methods, 86(3), 397–400. https://doi.org/10.1016/j.mimet.2011.06.013
- Dennis, J.E., Gay, D.M., & Welsch, R.E. (1981). An adaptive nonlinear least-squares algorithm. ACM Transactions on Mathematical Software, 7, 348-368.
- Dohm, J.C., Lottaz, C., Borodina, T., & Himmelbauer, H. (2008) Substantial biases in ultra-short read data sets from high-throughput DNA sequencing, Nucleic Acids Research, 36(16), e105.
- Edet, U. O., Bassey, I. U., & Joseph, A. P. (2023). Heavy metal co-resistance with antibiotics amongst bacteria isolates from an open dumpsite soil. Heliyon, 9(2), e13457. https://doi.org/10.1016/j.heliyon.2023.e13457
- Fisher, R A, Corbet, A S, Williams, C B (1943). The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. J Anim Ecol 12, 42-58, https://doi.org/10.2307/1411
- Gadde, U., Kim, W. H., Oh, S. T., & Lillehoj, H. S. (2017). Alternatives to antibiotics for maximizing growth performance and feed efficiency in poultry: a review. Animal health research reviews, 18(1), 26–45. https://doi.org/10.1017/S1466252316000207
- Garde, S., Chodisetti, P. K., & Reddy, M. (2021). Peptidoglycan: Structure, Synthesis, and Regulation. EcoSal Plus, 9(2), 10.1128/ecosalplus.ESP-0010-2020. https://doi.org/10.1128/ecosalplus.ESP-0010-2020
- Gaynes R. (2017). The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. Emerging Infectious Diseases, 23(5), 849–853. https://doi.org/10.3201/eid2305.161556
- Ghosh, D., Veeraraghavan, B., Elangovan, R., & Vivekanandan, P. (2020). Antibiotic Resistance and Epigenetics: More to It than Meets the Eye. Antimicrobial agents and chemotherapy, 64(2), e02225-19. https://doi.org/10.1128/AAC.02225-19

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Gilbert, J. A., Field, D., Swift, P., Thomas, S., Cummings, D., Temperton, B., Weynberg, K., Huse, S., Hughes, M., Joint, I., Somerfield, P. J., & Mühling, M. (2010). The taxonomic and functional diversity of microbes at a temperate coastal site: a 'multi-omic' study of seasonal and diel temporal variation. PloS one, 5(11), e15545. https://doi.org/10.1371/journal.pone.0015545
- Gingold E. B. (1985). Bacterial transformation. Methods in molecular biology (Clifton, N.J.), 2, 237–240. https://doi.org/10.1385/0-89603-064-4:237
- 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. https://doi.org/10.1101/pdb.prot5368
- Godzik A. (2011). Metagenomics and the protein universe. Current opinion in structural biology, 21(3), 398–403. https://doi.org/10.1016/j.sbi.2011.03.010
- Good, I J (1953). The population frequencies of species and the estimation of population parameters. Biometrika 40, 237-264, https://doi.org/10.1093/biomet/40.3-4.237
- Gould K. (2016). Antibiotics: from prehistory to the present day. The Journal of antimicrobial chemotherapy, 71(3), 572–575. https://doi.org/10.1093/jac/dkv484
- Gregorius, H.R. & Gillet, E.M. (2008). Generalized Simpson-diversity. Ecological Modelling, 211(1-2), 90-96. https://doi.org/10.1016/j.ecolmodel.2007.08.026
- Green, M. R., & Sambrook, J. (2019). Screening Bacterial Colonies Using X-Gal and IPTG: α- Complementation. Cold Spring Harbor protocols, 2019(12), 10.1101/pdb.prot101329. https://doi.org/10.1101/pdb.prot101329
- Gupta N. (2019). DNA Extraction and Polymerase Chain Reaction. Journal of cytology, 36(2), 116–117. https://doi.org/10.4103/JOC.JOC_110_18
- D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., 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. https://doi.org/10.1038/nature10388
- de Nies, L., Busi, S. B., Kunath, B. J., May, P., & Wilmes, P. (2022). Mobilome-driven segregation of the resistome in biological wastewater treatment. eLife, 11, e81196. https://doi.org/10.7554/eLife.81196
- Dib, J. R., Wagenknecht, M., Farías, M. E., & Meinhardt, F. (2015). Strategies and approaches in plasmidome studies-uncovering plasmid diversity disregarding of

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

linear elements?. Frontiers in microbiology, 6, 463. https://doi.org/10.3389/fmicb.2015.00463

- Diebold, P. J., New, F. N., Hovan, M., Satlin, M. J., & Brito, I. L. (2021). Linking plasmid-based beta-lactamases to their bacterial hosts using single-cell fusion PCR. eLife, 10, e66834. https://doi.org/10.7554/eLife.66834
- Domingues, S., da Silva, G. J., & Nielsen, K. M. (2012). Integrons: Vehicles and pathways for horizontal dissemination in bacteria. Mobile genetic elements, 2(5), 211–223. https://doi.org/10.4161/mge.22967
- Duarte, M., Jauregui, R., Vilchez-Vargas, R., Junca, H., & Pieper, D. H. (2014). AromaDeg, a novel database for phylogenomics of aerobic bacterial degradation of aromatics. Database : the journal of biological databases and curation, 2014, bau118. https://doi.org/10.1093/database/bau118
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics (Oxford, England), 27(16), 2194–2200. https://doi.org/10.1093/bioinformatics/btr381
- Escalona, M., Rocha, S., & Posada, D. (2016). A comparison of tools for the simulation of genomic next-generation sequencing data. Nature reviews. Genetics, 17(8), 459–469. https://doi.org/10.1038/nrg.2016.57
- Ewing, Brent, LaDeana Hillier, Michael C. Wendl, and Phil Green. 1998. "Base-Calling of Automated Sequencer Traces UsingPhred. I. Accuracy Assessment." Genome Research 8 (3): 175–85. https://doi.org/10.1101/gr.8.3.175.
- Fang, H., Wang, H., Cai, L., & Yu, Y. (2015). Prevalence of antibiotic resistance genes and bacterial pathogens in long-term manured greenhouse soils as revealed by metagenomic survey. Environmental Science & Technology, 49(2), 1095–1104. https://doi.org/10.1021/es504157v
- Finn D. R. (2024). A metagenomic alpha-diversity index for microbial functional biodiversity. FEMS microbiology ecology, 100(3), fiae019. https://doi.org/10.1093/femsec/fiae019
- Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., & Bateman, A. (2010). The Pfam protein families database. Nucleic acids research, 38(Database issue), D211–D222. https://doi.org/10.1093/nar/gkp985

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Forsberg, K. J., Patel, S., Gibson, M. K., Lauber, C. L., Knight, R., Fierer, N., & Dantas, G. (2014). Bacterial phylogeny structures soil resistomes across habitats. Nature, 509(7502), 612–615. https://doi.org/10.1038/nature13377
- Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., & Dantas, G. (2012). The shared antibiotic resistome of soil bacteria and human pathogens. Science (New York, N.Y.).), 337(6098), 1107–1111. https://doi.org/10.1126/science.1220761
- Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. Genomics, 107(1), 1–8. https://doi.org/10.1016/j.ygeno.2015.11.003
- Helinski D. R. (2022). A Brief History of Plasmids. EcoSal Plus, 10(1), eESP00282021. https://doi.org/10.1128/ecosalplus.esp-0028-2021
- Hickman, A. B., & Dyda, F. (2015). Mechanisms of DNA Transposition. Microbiology spectrum, 3(2), MDNA3–2014. https://doi.org/10.1128/microbiolspec.MDNA3-0034-2014
- Holmes, A. H., Moore, L. S., Sundsfjord, A., Steinbakk, M., Regmi, S., Karkey, A., Guerin, P. J., & Piddock, L. J. (2016). Understanding the mechanisms and drivers of antimicrobial resistance. Lancet (London, England), 387(10014), 176–187. https://doi.org/10.1016/S0140- 6736(15)00473-0
- Hutchings, M. I., Truman, A. W., & Wilkinson, B. (2019). Antibiotics: past, present and future. Current opinion in microbiology, 51, 72–80. https://doi.org/10.1016/j.mib.2019.10.008
- Islam, K. M., Klein, U., & Burch, D. G. (2009). The activity and compatibility of the antibiotic tiamulin with other drugs in poultry medicine--A review. Poultry science, 88(11), 2353–2359. https://doi.org/10.3382/ps.2009-00257
- Jayachandran S. (2018). Pre-antibiotics era to post-antibiotic era. Journal of Indian Academy of Oral Medicine and Radiology 30(2), 100-101. https://journals.lww.com/aomr/Fulltext/2018/30020/Pre_antibiotics_era_to_post_antibioti c_era.2.aspx
- Jechalke, S., Heuer, H., Siemens, J., Amelung, W., & Smalla, K. (2014). Fate and effects of veterinary antibiotics in soil. Trends in microbiology, 22(9), 536–545. https://doi.org/10.1016/j.tim.2014.05.005.
- Jiménez-Jiménez, C., Moreno, V. M., & Vallet-Regí, M. (2022). Bacteria-Assisted Transport of Nanomaterials to Improve Drug Delivery in Cancer Therapy.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Nanomaterials (Basel, Switzerland), 12(2), 288. https://doi.org/10.3390/nano12020288

- Johnson, C. M., & Grossman, A. D. (2015). Integrative and Conjugative Elements (ICEs): What They Do and How They Work. Annual review of genetics, 49, 577–601. https://doi.org/10.1146/annurev-genet-112414-055018
- Jørgensen, T. S., Xu, Z., Hansen, M. A., Sørensen, S. J., & Hansen, L. H. (2014). Hundreds of circular novel plasmids and DNA elements identified in a rat cecum metamobilome. PloS one, 9(2), e87924. https://doi.org/10.1371/journal.pone.0087924
- Jost, L. (2006). Entropy and diversity. Oikos 113, 363-375, https://nsojournals.onlinelibrary.wiley.com/doi/10.1111/j.2006.0030-1299.14714.x
- Jones, B. V., & Marchesi, J. R. (2007). Transposon-aided capture (TRACA) of plasmids resident in the human gut mobile metagenome. Nature methods, 4(1), 55–61. https://doi.org/10.1038/nmeth964
- Jost L (2006). Entropy and diversity. Oikos 113, 363-375.
- Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., & Hattori, M. (2004). The KEGG resource for deciphering the genome. Nucleic acids research, 32(Database issue), D277–D280. https://doi.org/10.1093/nar/gkh063
- Kapoor, G., Saigal, S., & Elongavan, A. (2017). Action and resistance mechanisms of antibiotics: A guide for clinicians. Journal of anaesthesiology, clinical pharmacology, 33(3), 300–305. https://doi.org/10.4103/joacp.JOACP_349_15
- Karaolia, P, Vasileiadis, S, Michael, S G, Karpouzas, D G, Fatta-Kassinos, D (2021). Shotgun metagenomics assessment of the resistome, mobilome, pathogen dynamics and their ecological control modes in full-scale urban wastewater treatment plants. J Hazard Mater 10.1016/j.jhazmat.2021.126387, 126387, 10.1016/j.jhazmat.2021.126387
- Kasimanickam, V., Kasimanickam, M., & Kasimanickam, R. (2021). Antibiotics Use in Food Animal Production: Escalation of Antimicrobial Resistance: Where Are We Now in Combating AMR?. Medical sciences (Basel, Switzerland), 9(1), 14. https://doi.org/10.3390/medsci9010014
- Kasman, L. M., & Porter, L. D. (2022). Bacteriophages. In StatPearls. StatPearls Publishing
- Katsivelou, E., Perruchon, C., Karas, P. A., Sarantidou, A., Pappa, E., Katsoula, A., Ligda, P., Sotiraki, S., Martin-Laurent, F., Vasileiadis, S., & Karpouzas, D. G. (2023).

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Accelerated dissipation, soil microbial toxicity and dispersal of antimicrobial resistance in soils repeatedly exposed to tiamulin, tilmicosin and sulfamethoxazole. The Science of the total environment, 893, 164817. https://doi.org/10.1016/j.scitotenv.2023.164817

- Katz, L., & Baltz, R. H. (2016). Natural product discovery: past, present, and future. Journal of industrial microbiology & biotechnology, 43(2-3), 155–176. https://doi.org/10.1007/s10295- 015-1723-5
- Killeavy, E. E., Jogl, G., & Gregory, S. T. (2020). tiamulin-Resistant Mutants of the Thermophilic Bacterium Thermus thermophilus. Antibiotics (Basel, Switzerland), 9(6), 313. https://doi.org/10.3390/antibiotics9060313
- Kim H. Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restorative dentistry & endodontics, 42(2), 152–155. https://doi.org/10.5395/rde.2017.42.2.152
- Kimura, N. (2018) Chapter 14 - Novel Biological Resources Screened From Uncultured Bacteria by a Metagenomic Method,Metagenomics, Academic Press, 273-288. https://doi.org/10.1016/B978-0-08-102268-9.00014-8
- Knapik, K. (2013). Genetic analysis of bacteriophages from clinical and environmental samples. https://doi.org/10.13140/RG.2.2.29564.08327
- Knight, R., Jansson, J., Field, D., Fierer, N., Desai, N., Fuhrman, J. A., Hugenholtz, P., van der Lelie, D., Meyer, F., Stevens, R., Bailey, M. J., Gordon, J. I., Kowalchuk, G. A., & Gilbert, J. A. (2012). Unlocking the potential of metagenomics through replicated experimental design. Nature biotechnology, 30(6), 513–520. https://doi.org/10.1038/nbt.2235
- Kolde R (2019). pheatmap: Pretty Heatmaps. R package version 1.0.12. https://CRAN.R-project.org/package=pheatmap
- Köninger, J., Lugato, E., Panagos, P., Kochupillai, M., Orgiazzi, A., & Briones, M. J. I. (2021). Manure management and soil biodiversity: Towards more sustainable food systems in the EU. Agricultural Systems, 194, 103251. https://doi.org/10.1016/j.agsy.2021.103251
- Koonin E. V. (2005). Orthologs, paralogs, and evolutionary genomics. Annual review of genetics, 39, 309–338. https://doi.org/10.1146/annurev.genet.39.073003.114725
- Krasucka, D., Mitura, A., Cybulski, W., Kos, K., & Pietroń, W. (2010). tiamulin hydrogen fumarate--veterinary uses and HPLC method of determination in premixes and medicated feeding stuffs. Acta poloniae pharmaceutica, 67(6), 682–685.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Ladd, M., & Bordoni, B. (2023). Genetics, Transposons. In StatPearls. StatPearls Publishing.
- Lapidus, A.L. (2009). Genome Sequence Databases: Sequencing and Assembly, Encyclopedia of Microbiology (Third Edition), Academic Press, https://doi.org/10.1016/B978-012373944-5.00028-6.
- LaPlante, K. L., Dhand, A., Wright, K., & Lauterio, M. (2022). Re-establishing the utility of tetracycline-class antibiotics for current challenges with antibiotic resistance. Annals of medicine, 54(1), 1686–1700. https://doi.org/10.1080/07853890.2022.2085881
- Lasken R. S. (2009). Genomic DNA amplification by the multiple displacement amplification (MDA) method. Biochemical Society transactions, 37(Pt 2), 450–453. https://doi.org/10.1042/BST0370450
- Lau, S. Y., & Zgurskaya, H. I. (2005). Cell division defects in Escherichia coli deficient in the multidrug efflux transporter AcrEF-TolC. Journal of bacteriology, 187(22), 7815–7825. https://doi.org/10.1128/JB.187.22.7815-7825.2005
- Lee, C. T., & Tuljapurkar, S. (2008). Population and prehistory I: Food-dependent population growth in constant environments. Theoretical population biology, 73(4), 473–482. https://doi.org/10.1016/j.tpb.2008.03.001
- Lee, H., Hsu, F. F., Turk, J., & Groisman, E. A. (2004). The PmrA-regulated pmrC gene mediates phosphoethanolamine modification of lipid A and polymyxin resistance in Salmonella enterica. Journal of bacteriology, 186(13), 4124–4133. https://doi.org/10.1128/JB.186.13.4124-4133.2004
- Li, Z., Zhou, J., Liang, H., Ye, L., Lan, L., Lu, F., Wang, Q., Lei, T., Yang, X., Cui, P., & Huang, J. (2022). Differences in Alpha Diversity of Gut Microbiota in Neurological Diseases. Frontiers in neuroscience, 16, 879318. https://doi.org/10.3389/fnins.2022.879318
- Liu, B., Zheng, D., Zhou, S., Chen, L., & Yang, J. (2022). VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic acids research, 50(D1), D912–D917. https://doi.org/10.1093/nar/gkab1107
- Liu, L. F., Rowe, T. C., Yang, L., Tewey, K. M., & Chen, G. L. (1983). Cleavage of DNA by mammalian DNA topoisomerase II. The Journal of biological chemistry, 258(24), 15365–15370.

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Lowe, T. M., & Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic acids research, 25(5), 955–964. https://doi.org/10.1093/nar/25.5.955
- Lukashin, A. V., & Borodovsky, M. (1998). GeneMark.hmm: new solutions for gene finding. Nucleic acids research, 26(4), 1107–1115. https://doi.org/10.1093/nar/26.4.1107
- Maertens, G. N., Engelman, A. N., & Cherepanov, P. (2022). Structure and function of retroviral integrase. Nature reviews. Microbiology, 20(1), 20–34. https://doi.org/10.1038/s41579-021-00586-9
- Mancuso, G., De Gaetano, S., Midiri, A., Zummo, S., & Biondo, C. (2023). The Challenge of Overcoming Antibiotic Resistance in Carbapenem-Resistant Gram-Negative Bacteria: "Attack on Titan". Microorganisms, 11(8), 1912. https://doi.org/10.3390/microorganisms11081912
- Marcon, E., & Hérault, B. (2015). entropart: An R Package to Measure and Partition Diversity. Journal of Statistical Software, 67(8), 1–26. https://doi.org/10.18637/jss.v067.i08
- Márquez, C., Labbate, M., Ingold, A. J., Roy Chowdhury, P., Ramírez, M. S., Centrón, D., Borthagaray, G., & Stokes, H. W. (2008). Recovery of a functional class 2 integron from an Escherichia coli strain mediating a urinary tract infection. Antimicrobial agents and chemotherapy, 52(11), 4153–4154. https://doi.org/10.1128/AAC.00710-08
- Marti, R., Scott, A., Tien, Y. C., Murray, R., Sabourin, L., Zhang, Y., & Topp, E. (2013). Impact of manure fertilization on the abundance of antibiotic-resistant bacteria and frequency of detection of antibiotic resistance genes in soil and on vegetables at harvest. Applied and Environmental Microbiology, 79(18), 5701–5709. https://doi.org/10.1128/AEM.01682-13
- McHardy, A. C., Martín, H. G., Tsirigos, A., Hugenholtz, P., & Rigoutsos, I. (2007). Accurate phylogenetic classification of variable-length DNA fragments. Nature methods, 4(1), 63–72. https://doi.org/10.1038/nmeth976
- Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nature communications, 7, 11257. https://doi.org/10.1038/ncomms11257
- Metzker M. L. (2010). Sequencing technologies - the next generation. Nature reviews. Genetics, 11. https://doi.org/10.1038/nrg2626

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Michael, C. A., Dominey-Howes, D., and Labbate, M. (2014). 'The Antimicrobial Resistance Crisis: Causes, Consequences, and Management'. Frontiers in Public Health 2. https://doi.org/10.3389/fpubh.2014.00145
- Miller, J. R., Koren, S., & Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. Genomics, 95(6), 315–327. https://doi.org/10.1016/j.ygeno.2010.03.001
- Monzoorul-Haque, M., Ghosh, T. S., Komanduri, D., & Mande, S. S. (2009). SOrt-ITEMS: Sequence orthology based approach for improved taxonomic estimation of metagenomic sequences. Bioinformatics (Oxford, England), 25(14), 1722–1730. https://doi.org/10.1093/bioinformatics/btp317
- More, S J (2020). European perspectives on efforts to reduce antimicrobial usage in food animal production. Irish Veterinary Journal 73, 2, 10.1186/s13620-019-0154-4
- Mukherjee, A., Sen, S., & Agarwal, K. (1993). Ciprofloxacin: mammalian DNA topoisomerase type II poison in vivo. Mutation research, 301(2), 87–92. https://doi.org/10.1016/0165-7992(93)90029-u
- Muhamad Rizal, N. S., Neoh, H. M., Ramli, R., A/L K Periyasamy, P. R., Hanafiah, A., Abdul Samat, M. N., Tan, T. L., Wong, K. K., Nathan, S., Chieng, S., Saw, S. H., & Khor, B. Y. (2020). Advantages and Limitations of 16S rRNA Next-Generation Sequencing for Pathogen Identification in the Diagnostic Microbiology Laboratory: Perspectives from a Middle-Income Country. Diagnostics (Basel, Switzerland), 10(10), 816. https://doi.org/10.3390/diagnostics10100816
- 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(Database issue), D190–D195. https://doi.org/10.1093/nar/gkp951
- Munita, J. M., & Arias, C. A. (2016). Mechanisms of Antibiotic Resistance. Microbiology spectrum, 4(2), 10.1128/microbiolspec.VMBF-0016-2015. https://doi.org/10.1128/microbiolspec.VMBF-0016-2015
- Nagy, E. Z., Kovács, Á. B., Wehmann, E., Bekő, K., Földi, D., Bányai, K., Kreizinger, Z., & Gyuranecz, M. (2023). Phenotypic and genetic insights into efflux pump mechanism in Mycoplasma anserisalpingitidis. Frontiers in microbiology, 14, 1216893. https://doi.org/10.3389/fmicb.2023.1216893

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Nguyen, B. T., Chen, Q. L., He, J. Z., & Hu, H. W. (2022). Livestock manure spiked with the antibiotic tylosin significantly altered soil protist functional groups. Journal of hazardous materials, 427, 127867. https://doi.org/10.1016/j.jhazmat.2021.127867
- Nicolaou, K. C., & Rigol, S. (2018). A brief history of antibiotics and select advances in their synthesis. The Journal of antibiotics, 71(2), 153–184. https://doi.org/10.1038/ja.2017.62
- 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 : an international journal for rapid publication of reports on genes and genomes, 15(6), 387–396. https://doi.org/10.1093/dnares/dsn027
- O'Connor, T., Kleemann, R., & Attard, J. (2022). Vulnerable vegetables and efficient fishers: A study of primary production food losses and waste in Ireland. Journal of environmental management, 307, 114498. https://doi.org/10.1016/j.jenvman.2022.114498
- Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2024). vegan: Community Ecology Package. R package version 2.6-6. https://CRAN.R-project.org/package=vegan
- Olivier, N. B. (2014). Ribosome revelations. Nature, 513(7519), 491–492. https://doi.org/10.1038/nature13747
- Ouyang, Y., Li, Q., Kuang, X., Wang, H., Wu, J., Ayepa, E., Chen, H., Abrha, G. T., Zhang, Z., Li, X., & Ma, M. (2021). YMR152W from Saccharomyces cerevisiae encoding a novel aldehyde reductase for detoxification of aldehydes derived from lignocellulosic biomass. Journal of bioscience and bioengineering, 131(1), 39–46. https://doi.org/10.1016/j.jbiosc.2020.09.004
- 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(Database issue), D206–D214. https://doi.org/10.1093/nar/gkt1226

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Pagès, J. M., James, C. E., & Winterhalter, M. (2008). The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. Nature reviews. Microbiology, 6(12), 893–903. https://doi.org/10.1038/nrmicro1994
- Pal, C., Bengtsson-Palme, J., Rensing, C., Kristiansson, E., & Larsson, D. G. (2014). BacMet: antibacterial biocide and metal resistance genes database. Nucleic acids research, 42(Database issue), D737–D743. https://doi.org/10.1093/nar/gkt1252
- Pao, S. S., Paulsen, I. T., & Saier, M. H., Jr (1998). Major facilitator superfamily. Microbiology and molecular biology reviews : MMBR, 62(1), 1–34. https://doi.org/10.1128/MMBR.62.1.1-34.1998
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome research, 25(7), 1043–1055. https://doi.org/10.1101/gr.186072.114
- Partridge, S. R., Kwong, S. M., Firth, N., & Jensen, S. O. (2018). Mobile Genetic Elements Associated with Antimicrobial Resistance. Clinical microbiology reviews, 31(4), e00088-17. https://doi.org/10.1128/CMR.00088-17
- Peng, Y., Leung, H. C., Yiu, S. M., & Chin, F. Y. (2011). Meta-IDBA: a de Novo assembler for metagenomic data. Bioinformatics (Oxford, England), 27(13), i94–i101. https://doi.org/10.1093/bioinformatics/btr216
- Perruchon, C, Katsivelou, E, Karas, P A, Vassilakis, S, Lithourgidis, A A, Kotsopoulos, T A, Sotiraki, S, Vasileiadis, S, Karpouzas, D G (2022). Following the route of veterinary antibiotics tiamulin and tilmicosin from livestock farms to agricultural soils. J Hazard Mater 429, 128293, https://doi.org/10.1016/j.jhazmat.2022.128293
- Poey, M. E., Azpiroz, M. F., & Laviña, M. (2019). On sulfonamide resistance, sul genes, class 1 integrons and their horizontal transfer in Escherichia coli. Microbial pathogenesis, 135, 103611. https://doi.org/10.1016/j.micpath.2019.103611
- Poupaud, M., Putthana, V., Patriarchi, A., Caro, D., Agunos, A., Tansakul, N., & Goutard, F. L. (2021). Understanding the veterinary antibiotics supply chain to address antimicrobial resistance in Lao PDR: Roles and interactions of involved stakeholders. Acta tropica, 220, 105943. https://doi.org/10.1016/j.actatropica.2021.105943
- Pozdeev, G., Mogre, A., & Dorman, C. J. (2021). Consequences of producing DNA gyrase from a synthetic gyrBA operon in Salmonella enterica serovar Typhimurium. Molecular microbiology, 115(6), 1410–1429. https://doi.org/10.1111/mmi.14689

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Prescott J. F. (2014). The resistance tsunami, antimicrobial stewardship, and the golden age of microbiology. Veterinary microbiology, 171(3-4), 273–278. https://doi.org/10.1016/j.vetmic.2014.02.035
- Prüsse, E, Quast, C, Yilmaz, P, Ludwig, W, Peplies, J, Glöckner, F O (2011). SILVA: comprehensive databases for quality checked and aligned ribosomal rna sequence data compatible with ARB. In Handbook of Molecular Microbial Ecology I. eds.John Wiley & Sons, Inc.) 10.1002/9781118010518.ch45, pp 393-398.
- 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. https://doi.org/10.1038/nbt.3935
- R Core Team (2024). R: A Language and Environment for Statistical Computing . R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
- Ravenhall, M., Škunca, N., Lassalle, F., & Dessimoz, C. (2015). Inferring horizontal gene transfer. PLoS computational biology, 11(5), e1004095. https://doi.org/10.1371/journal.pcbi.1004095
- Ribeiro da Cunha, B., Fonseca, L. P., & Calado, C. R. C. (2019). Antibiotic Discovery: Where Have We Come from, Where Do We Go?. Antibiotics (Basel, Switzerland), 8(2), 45. https://doi.org/10.3390/antibiotics8020045
- Rickwood, D., Ford, T., & Graham, J. (1982). Nycodenz: a new nonionic iodinated gradient medium. Analytical biochemistry, 123(1), 23–31. https://doi.org/10.1016/0003-2697(82)90618-2
- Roberts M. C. (2005). Update on acquired tetracycline resistance genes. FEMS microbiology letters, 245(2), 195–203. https://doi.org/10.1016/j.femsle.2005.02.034
- Rodriguez-R, L M (2018). Nonpareil: metagenome coverage estimation and projections for 'Nonpareil'. R package version 3.3.1. 2018, from https://CRAN.R-project.org/package=Nonpareil
- Rodriguez-R, L. M., & Konstantinidis, K. T. (2014a). Estimating coverage in metagenomic data sets and why it matters. The ISME journal, 8(11), 2349–2351. https://doi.org/10.1038/ismej.2014.76
- Rodriguez-R, L. M., & Konstantinidis, K. T. (2014b). Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. Bioinformatics (Oxford, England), 30(5), 629–635. https://doi.org/10.1093/bioinformatics/btt584

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

- Rodriguez-R, L. M., Gunturu, S., Harvey, W. T., Rosselló-Mora, R., Tiedje, J. M., Cole, J. R., & Konstantinidis, K. T. (2018). The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucleic acids research, 46(W1), W282–W288. https://doi.org/10.1093/nar/gky467
- Rolling, T., Zhai, B., Frame, J., Hohl, T. M., & Taur, Y. (2022). Customization of a DADA2-based pipeline for fungal internal transcribed spacer 1 (ITS1) amplicon data sets. JCI insight, 7(1), e151663. https://doi.org/10.1172/jci.insight.151663
- Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., Loiacono, K. A., Lynch, B. A., MacNeil, I. A., Minor, C., Tiong, C. L., Gilman, M., Osburne, M. S., Clardy, J., Handelsman, J., & Goodman, R. M. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Applied and environmental microbiology, 66(6), 2541–2547. https://doi.org/10.1128/AEM.66.6.2541-2547.2000
- Schwarz, S., Shen, J., Kadlec, K., Wang, Y., Brenner Michael, G., Feßler, A. T., & Vester, B. (2016). Lincosamides, Streptogramins, Phenicols, and Pleuromutilins: Mode of Action and Mechanisms of Resistance. Cold Spring Harbor perspectives in medicine, 6(11), a027037. https://doi.org/10.1101/cshperspect.a027037
- Shannon, C E (1948). A mathematical theory of communication. Bell System Technical Journal 27, 379–423 and 623–656
- Siefert J. L. (2009). Defining the mobilome. Methods in molecular biology (Clifton, N.J.), 532, 13–27. https://doi.org/10.1007/978-1-60327-853-9_2
- Silva, B. S. O., Coutinho, F. H., Gregoracci, G. B., Leomil, L., de Oliveira, L. S., Fróes, A., Tschoeke, D., Soares, A. C., Cabral, A. S., Ward, N. D., Richey, J. E., Krusche, A. V., Yager, P. L., de Rezende, C. E., Thompson, C. C., & Thompson, F. L. (2017). Virioplankton Assemblage Structure in the Lower River and Ocean Continuum of the Amazon. mSphere, 2(5), e00366-17. https://doi.org/10.1128/mSphere.00366-17
- Simpson, E H (1949). Measurement of diversity. Nature 163, 688-688, https://doi.org/10.1038/163688a0
- Smalla, K., Jechalke, S., & Top, E. M. (2015). Plasmid Detection, Characterization, and Ecology. Microbiology spectrum, 3(1), 10.1128/microbiolspec.PLAS-0038-2014. https://doi.org/10.1128/microbiolspec.PLAS-0038-2014
- Song, W., Wemheuer, B., Zhang, S., Steensen, K., & Thomas, T. (2019). MetaCHIP: Community-level horizontal gene transfer identification through the combination of

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

best-match and phylogenetic approaches. Microbiome, 7(1), 36. https://doi.org/10.1186/s40168-019-0649-y

- Spohn, S.N., Young, V.B., (2018). Physiology of the Gastrointestinal Tract: Chapter 32 - Gastrointestinal Microbial Ecology With Perspectives on Health and Disease (Sixth Edition), Academic Press, 737-753, https://doi.org/10.1016/B978-0-12-809954-4.00032-3.
- Stanton, I C, Murray, A K, Zhang, L, Snape, J, Gaze, W H (2020). Evolution of antibiotic resistance at low antibiotic concentrations including selection below the minimal selective concentration. Communications Biology 3, 467, 10.1038/s42003-020-01176-w
- Sun, Y. M., Qiu, T. L., Gao, M., Shi, M. M., Zhang, H., F., & Wang, X. M. (2019). Inorganic and organic fertilizers application enhanced antibiotic resistome in greenhouse soils growing vegetables. Ecotoxicology and Environmental Safety, 179, 24–30. https://doi.org/10.1016/j.ecoenv.2019.04.039
- Swann MM. (1969) Joint Committee on the Use of Antibiotics in Animal Husbandry and Veterinary Medicine: report presented to Parliament by the Secretary of State for Social Services. HM Stationery Office. 791(1), 1525-31. https://api.parliament.uk/historic- hansard/commons/1969/nov/20/use-of-antibiotics-in-animal-husbandry
- Tansirichaiya, S., Mullany, P., & Roberts, A. P. (2016). PCR-based detection of composite transposons and translocatable units from oral metagenomic DNA. FEMS microbiology letters, 363(18), fnw195. https://doi.org/10.1093/femsle/fnw195
- Taş, N., de Jong, A. E., Li, Y., Trubl, G., Xue, Y., & Dove, N. C. (2021). Metagenomic tools in microbial ecology research. Current opinion in biotechnology, 67, 184–191. https://doi.org/10.1016/j.copbio.2021.01.019
- Thierauf, A., Perez, G., & Maloy, A. S. (2009). Generalized transduction. Methods in molecular biology (Clifton, N.J.), 501, 267–286. https://doi.org/10.1007/978-1-60327-164-6_23
- Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics - a guide from sampling to data analysis. Microbial informatics and experimentation, 2(1), 3. https://doi.org/10.1186/2042-5783-2-3
- Thomas, T., Rusch, D., DeMaere, M. Z., Yung, P. Y., Lewis, M., Halpern, A., Heidelberg, K. B., Egan, S., Steinberg, P. D., & Kjelleberg, S. (2010). Functional genomic signatures

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

of sponge bacteria reveal unique and shared features of symbiosis. The ISME journal, 4(12), 1557–1567. https://doi.org/10.1038/ismej.2010.74

- Tian, Z. X., Yi, X. X., Cho, A., O'Gara, F., & Wang, Y. P. (2016). CpxR Activates MexAB-OprM Efflux Pump Expression and Enhances Antibiotic Resistance in Both Laboratory and Clinical nalB-Type Isolates of Pseudomonas aeruginosa. PLoS pathogens, 12(10), e1005932. https://doi.org/10.1371/journal.ppat.1005932
- Tickle, T. L., Segata, N., Waldron, L., Weingart, U., & Huttenhower, C. (2013). Two-stage microbial community experimental design. The ISME journal, 7(12), 2330–2339. https://doi.org/10.1038/ismej.2013.139
- Tsui, W. H., Yim, G., Wang, H. H., McClure, J. E., Surette, M. G., & Davies, J. (2004). Dual effects of MLS antibiotics: transcriptional modulation and interactions on the ribosome. Chemistry & biology, 11(9), 1307–1316. https://doi.org/10.1016/j.chembiol.2004.07.010
- Uddin, T. M., Chakraborty, A. J., Khusro, A., Zidan, B. R. M., Mitra, S., Emran, T. B., Dhama, K., Ripon, M. K. H., Gajdács, M., Sahibzada, M. U. K., Hossain, M. J., & Koirala, N. (2021). Antibiotic resistance in microbes: History, mechanisms, therapeutic strategies and future prospects. Journal of infection and public health, 14(12), 1750–1766. https://doi.org/10.1016/j.jiph.2021.10.020
- Vandecraen, J., Chandler, M., Aertsen, A., & Van Houdt, R. (2017). The impact of insertion sequences on bacterial genome plasticity and adaptability. Critical reviews in microbiology, 43(6), 709–730. https://doi.org/10.1080/1040841X.2017.1303661
- Van Epps, A., & Blaney, L. (2016). Antibiotic Residues in Animal Waste: Occurrence and Degradation in Conventional Agricultural Waste Management Practices. Current Pollution Reports, 2(3), 135–155. https://doi.org/10.1007/s40726-016-0037-1
- Virolle, C., Goldlust, K., Djermoun, S., Bigot, S., & Lesterlin, C. (2020). Plasmid Transfer by Conjugation in Gram-Negative Bacteria: From the Cellular to the Community Level. Genes, 11(11), 1239. https://doi.org/10.3390/genes11111239
- von Wintersdorff, C. J., Penders, J., van Niekerk, J. M., Mills, N. D., Majumder, S., van Alphen, L. B., Savelkoul, P. H., & Wolffs, P. F. (2016). Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. Frontiers in microbiology, 7, 173. https://doi.org/10.3389/fmicb.2016.00173
- Wang, M., & Kong, L. (2019). pblat: a multithread blat algorithm speeding up aligning sequences to genomes. BMC bioinformatics, 20(1), 28. https://doi.org/10.1186/s12859-019-2597-8

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

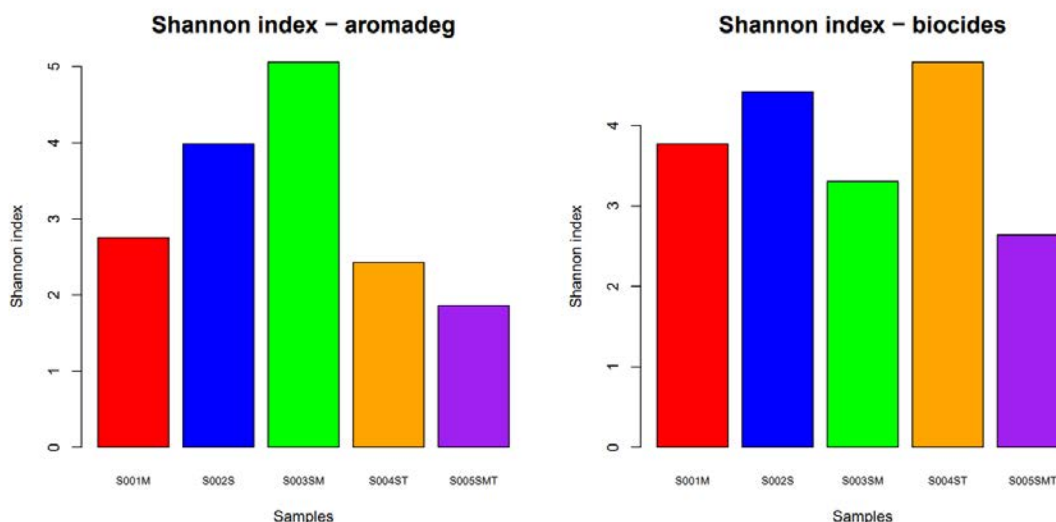
- Wesolowska-Andersen, A., Bahl, M. I., Carvalho, V., Kristiansen, K., Sicheritz-Pontén, T., Gupta, R., & Licht, T. R. (2014). Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. Microbiome, 2, 19. https://doi.org/10.1186/2049-2618-2-19
- WHO. (2017). Antibacterial agents in clinical development. 16(1). 113-123. https://www.jpiamr.eu/app/uploads/2021/09/IAU_Antibacterial-Agents-in-Clinical-Development_webfinal_-2017_09_19.pdf
- Willis A. D. (2019). Rarefaction, Alpha Diversity, and Statistics. Frontiers in microbiology, 10, 2407. https://doi.org/10.3389/fmicb.2019.02407
- Wilson D. N. (2014). Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nature reviews. Microbiology, 12(1), 35–48. https://doi.org/10.1038/nrmicro3155
- Wommack, K. E., Bhavsar, J., & Ravel, J. (2008). Metagenomics: read length matters. Applied and environmental microbiology, 74(5), 1453–1463. https://doi.org/10.1128/AEM.02181-07
- Xu, C, Kong, L, Gao, H, Cheng, X, Wang, X (2022). review of current bacterial resistance to antibiotics in food animals. Front Microbiol 1310.3389/fmicb.2022.822689
- Yao, S., & Liu, T. (2018). Analysis of differential gene expression caused by cervical intraepithelial neoplasia based on GEO database. Oncology letters, 15(6), 8319–8324. https://doi.org/10.3892/ol.2018.8403
- Yoneyama, H., & Katsumata, R. (2006). Antibiotic resistance in bacteria and its future for novel antibiotic development. Bioscience, biotechnology, and biochemistry, 70(5), 1060– 1075. https://doi.org/10.1271/bbb.70.1060
- Yu, C. S., Lin, C. H., Lin, Y. J., Lin, S. Y., Wang, S. T., Wu, J. L., Tsai, M. H., & Chang, S. S. (2020). Clustering Heatmap for Visualizing and Exploring Complex and High-dimensional Data Related to Chronic Kidney Disease. Journal of clinical medicine, 9(2), 403. https://doi.org/10.3390/jcm9020403
- Zanghellini A. (2014). de novo computational enzyme design. Current opinion in biotechnology, 29, 132–138. https://doi.org/10.1016/j.copbio.2014.03.002
- Zhao, R., Feng, J., Liu, J., Fu, W., Li, X., & Li, B. (2019). Deciphering of microbial community and antibiotic resistance genes in activated sludge reactors under high selective pressure of different antibiotics. Water research, 151, 388–402. https://doi.org/10.1016/j.watres.2018.12.034

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

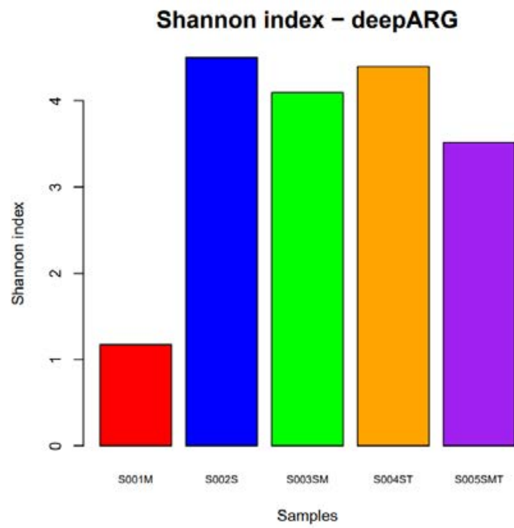
- Zhao, S., Guo, Y., Sheng, Q., & Shyr, Y. (2014). Advanced heat map and clustering analysis using heatmap3. BioMed research international, 2014, 986048. https://doi.org/10.1155/2014/986048
- Zheng, D., Yin, G., Liu, M., Hou, L., Yang, Y., Van Boeckel, T. P., Zheng, Y., & Li, Y. (2022). Global biogeography and projection of soil antibiotic resistance genes. Science advances, 8(46), eabq8015. https://doi.org/10.1126/sciadv.abq8015
- Zhu, B. K., Chen, Q. L., Chen, S. C., & Zhu, Y. G. (2017). Does organically produced lettuce harbor higher abundance of antibiotic resistance genes than conventionally produced? Environment International, 98, 152–159. https://doi.org/10.1016/j.envint.2016.11.001
- Zhu, C., & Yu, J. (2009). Nonmetric multidimensional scaling corrects for population structure in association mapping with different sample types. Genetics, 182(3), 875–888. https://doi.org/10.1534/genetics.108.098863
- Zhu, H. S., Liang, X., Liu, J. C., Zhong, H. Y., Yang, Y. H., Guan, W. P., Du, Z. J., & Ye, M. Q. (2023). Antibiotic and Heavy Metal Co-Resistant Strain Isolated from Enrichment Culture of Marine Sediments, with Potential for Environmental Bioremediation Applications. Antibiotics (Basel, Switzerland), 12(9), 1379. https://doi.org/10.3390/antibiotics12091379

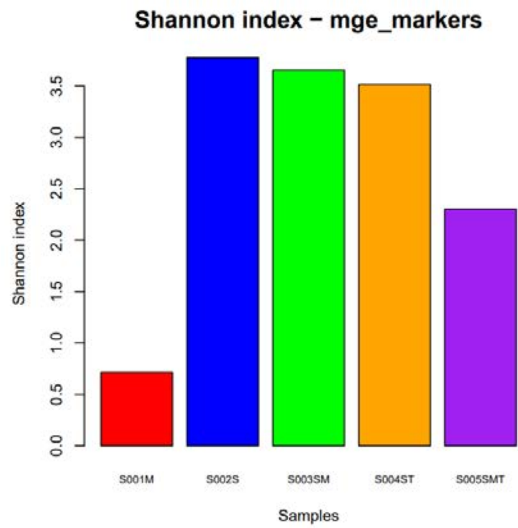
6. Supplementary

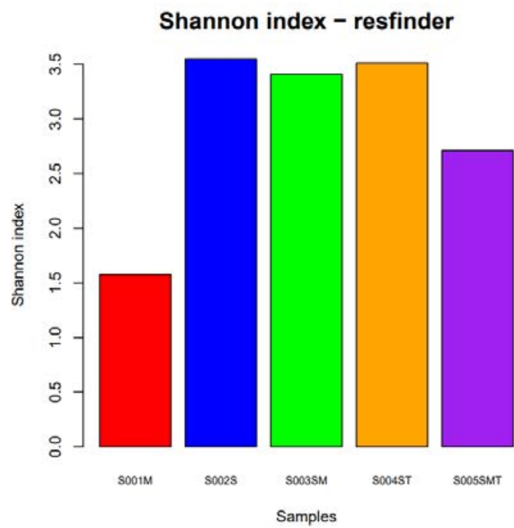
6.1 Supplementary I

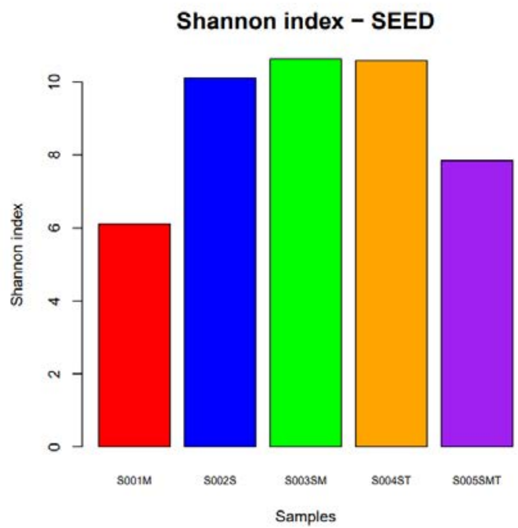


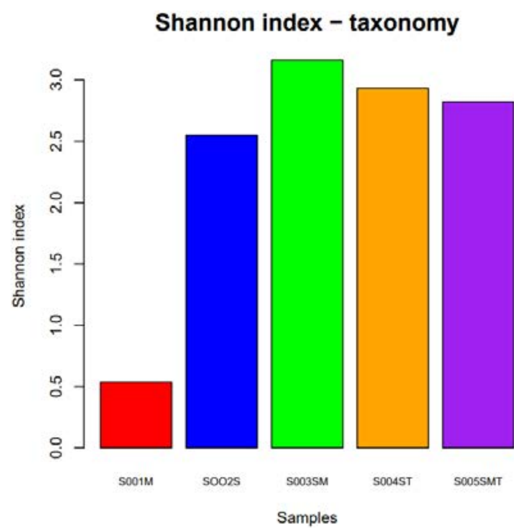
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

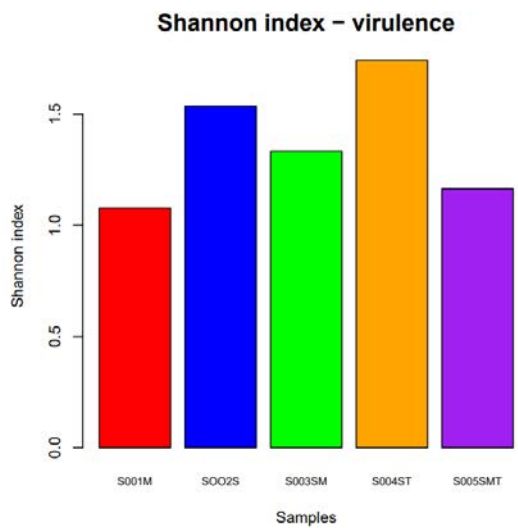




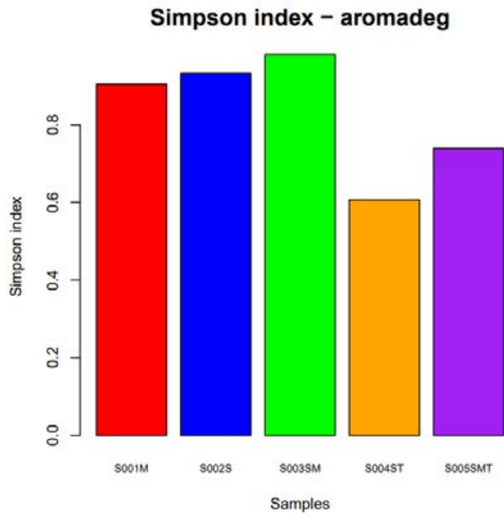


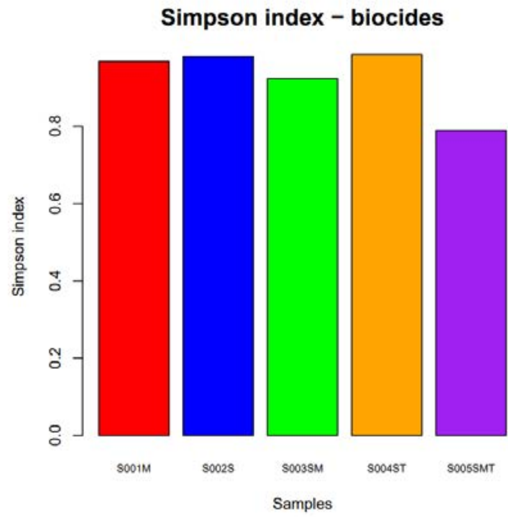


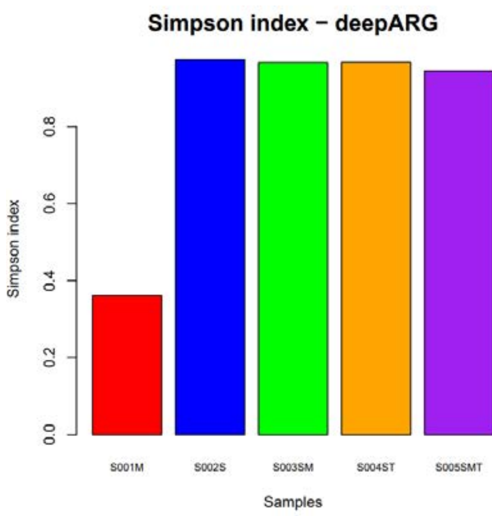


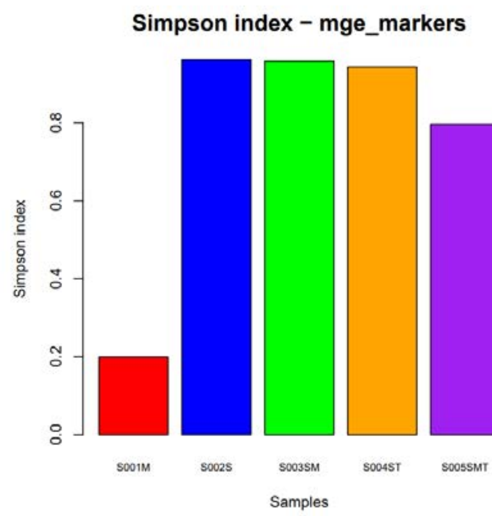


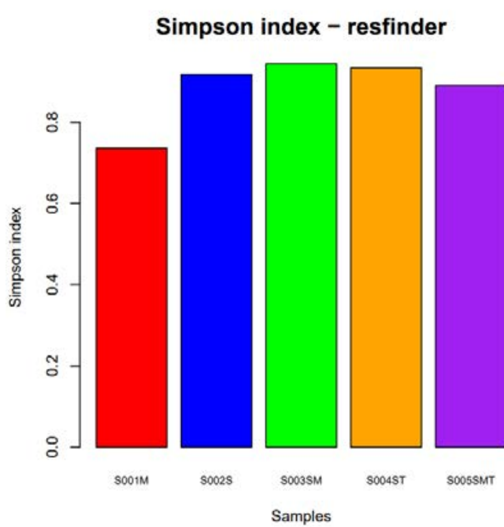
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

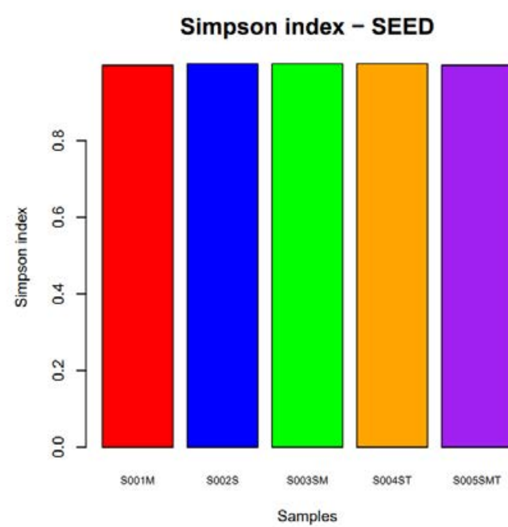




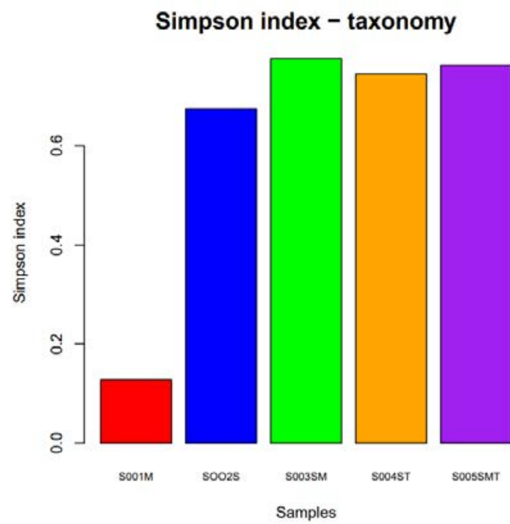


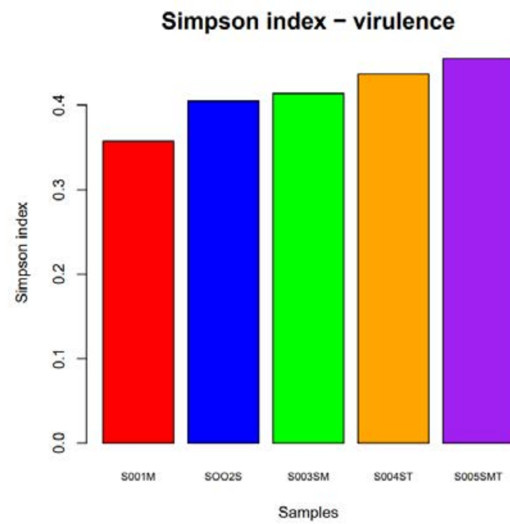


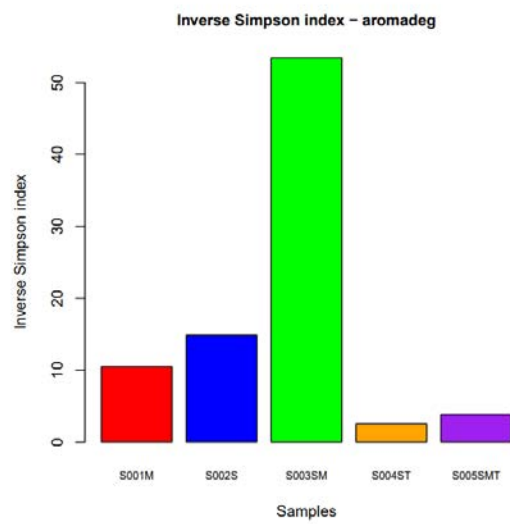


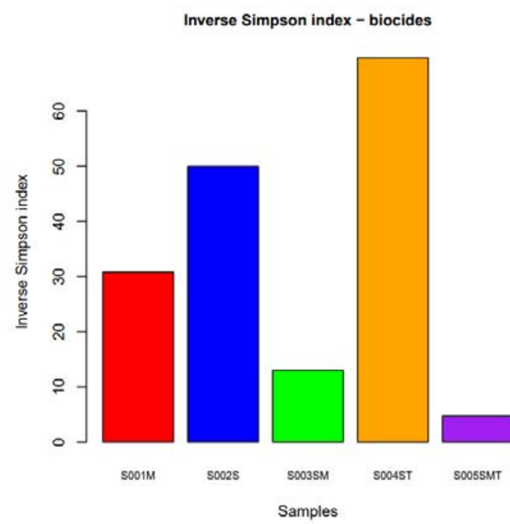


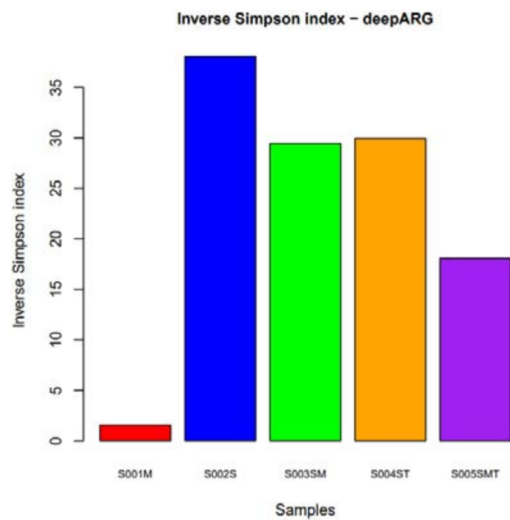
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

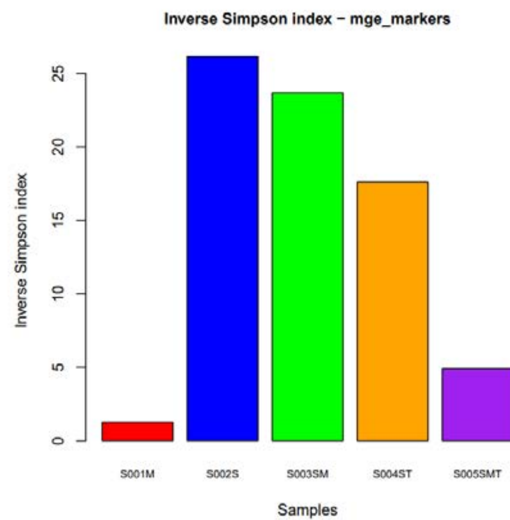




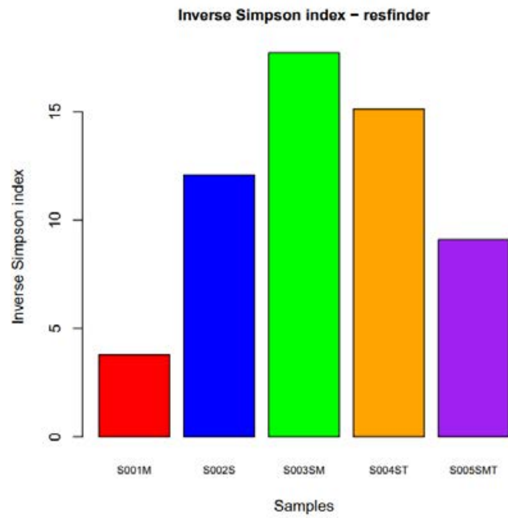


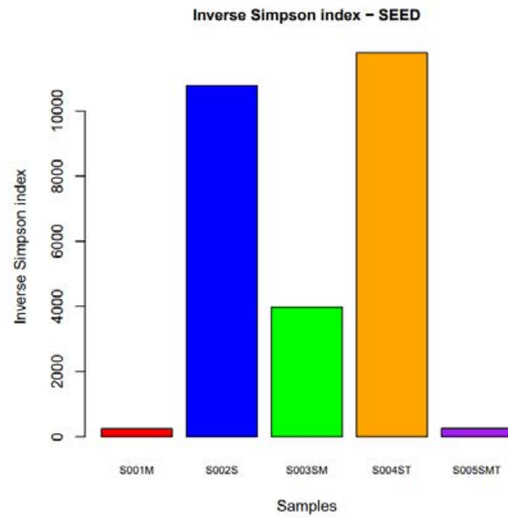


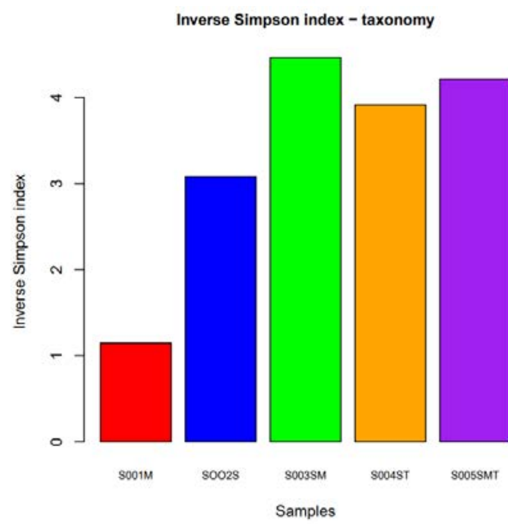


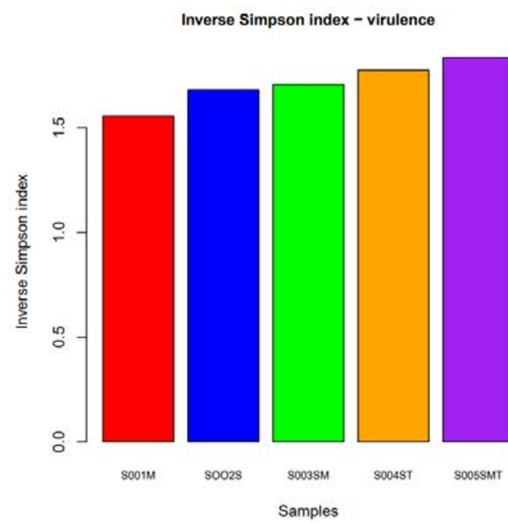


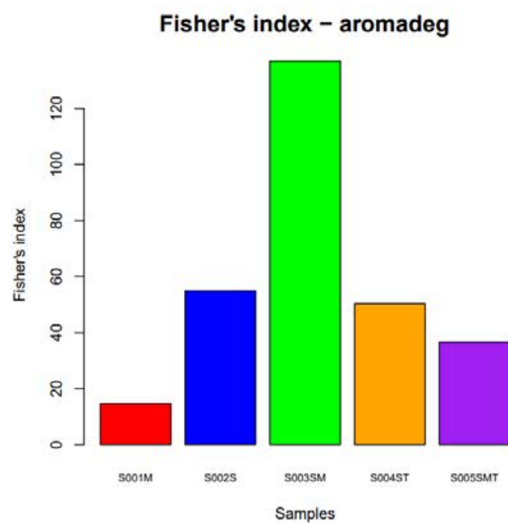
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

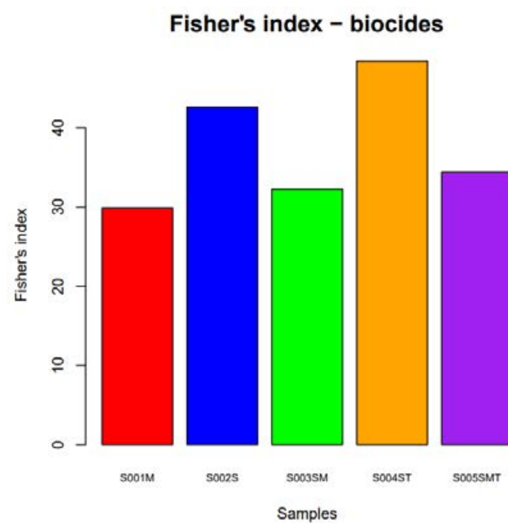




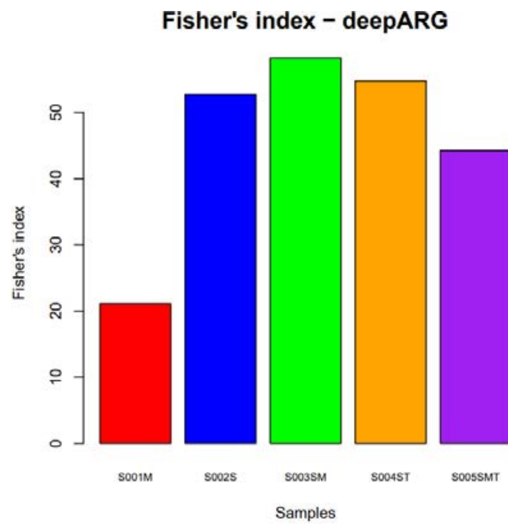


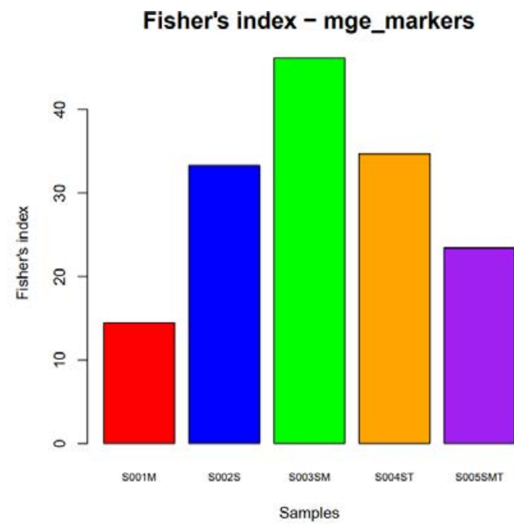


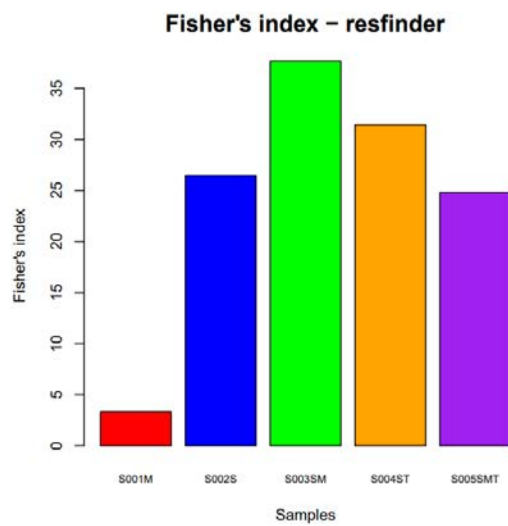


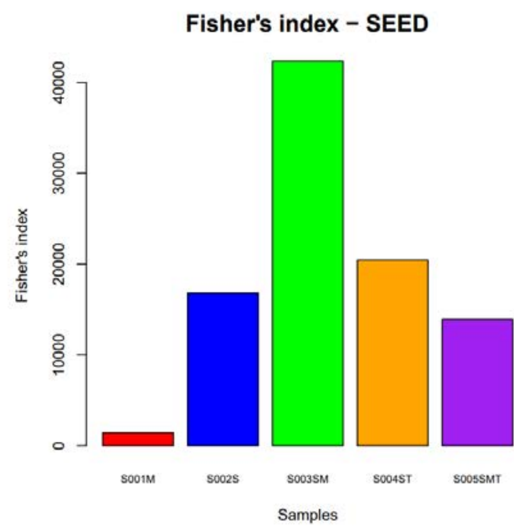


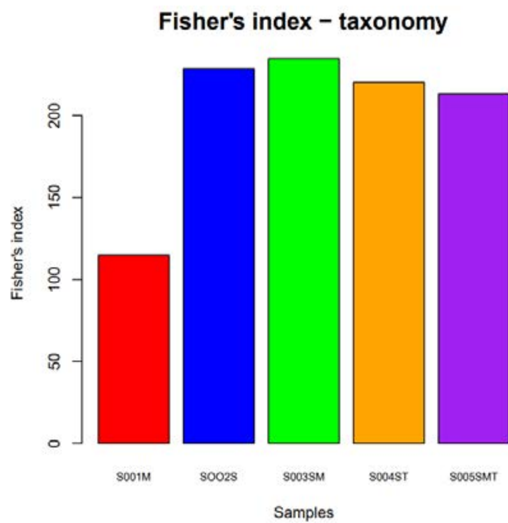
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

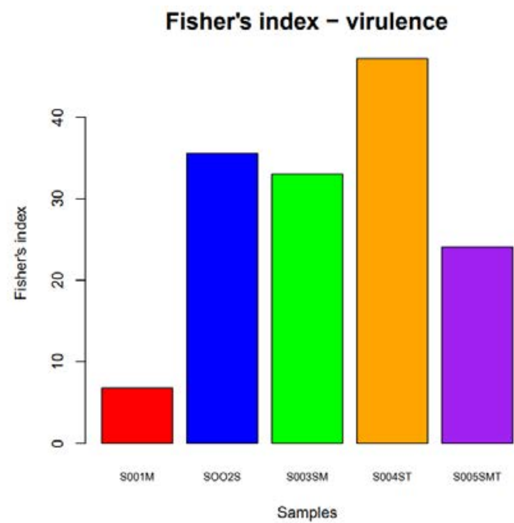




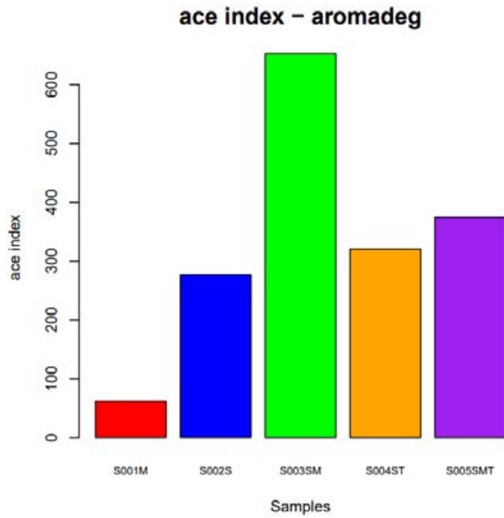


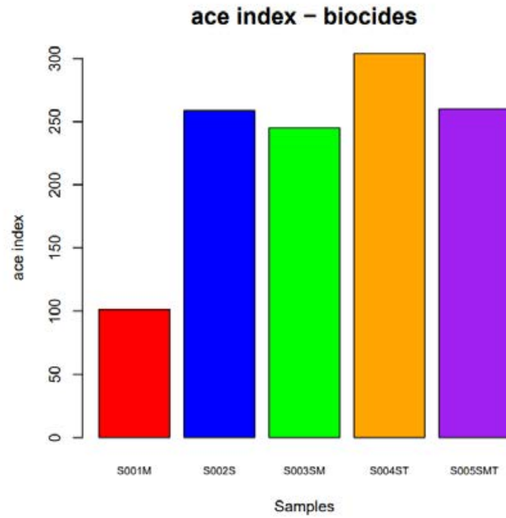


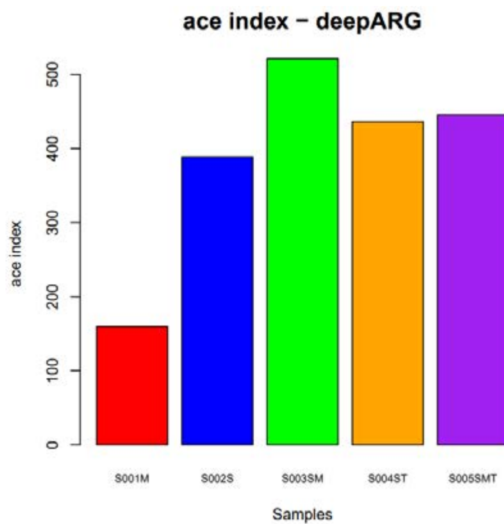


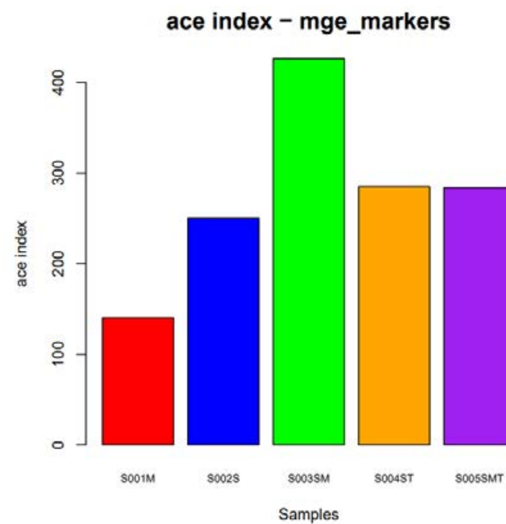


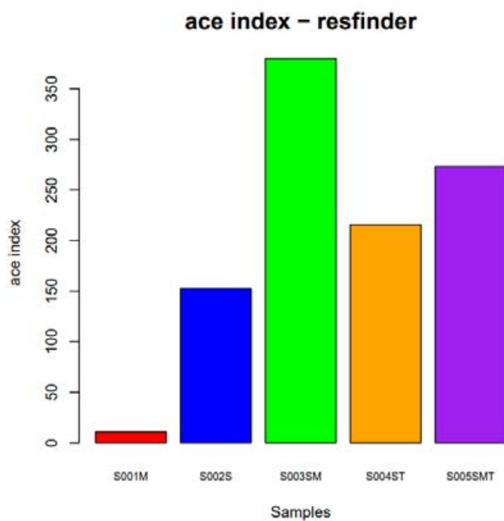
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

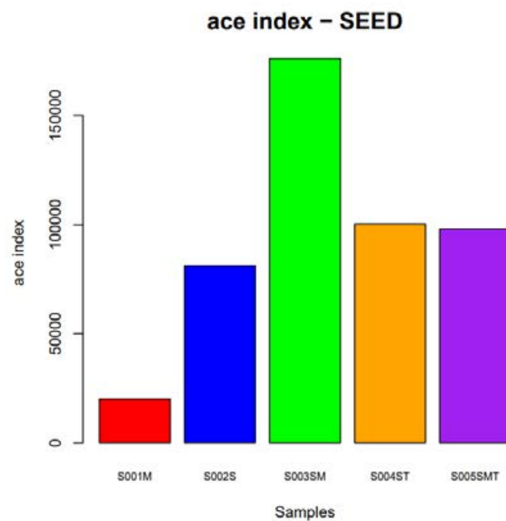




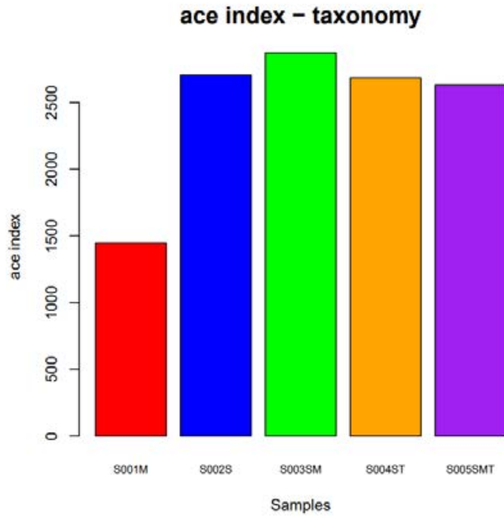


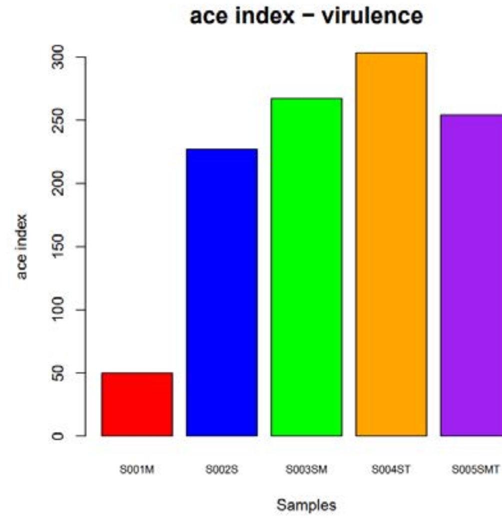


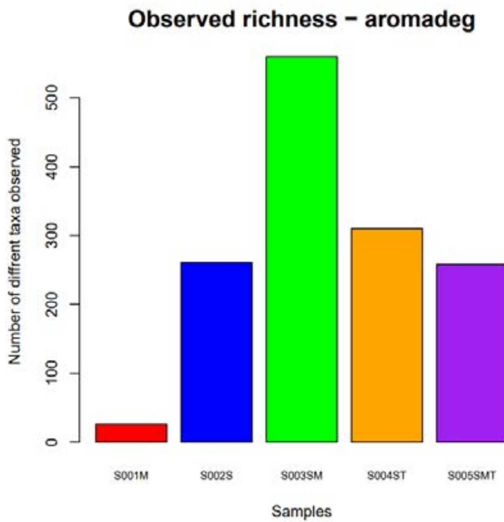


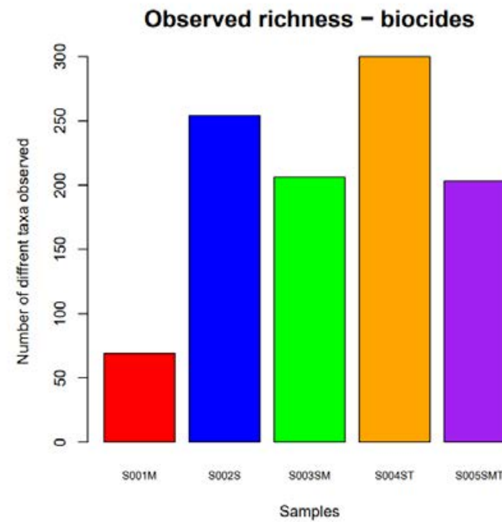


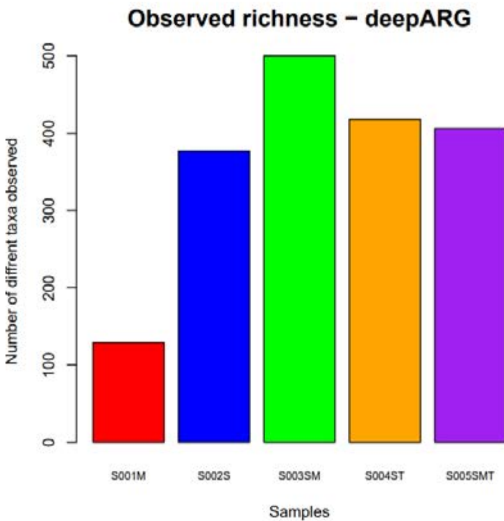
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

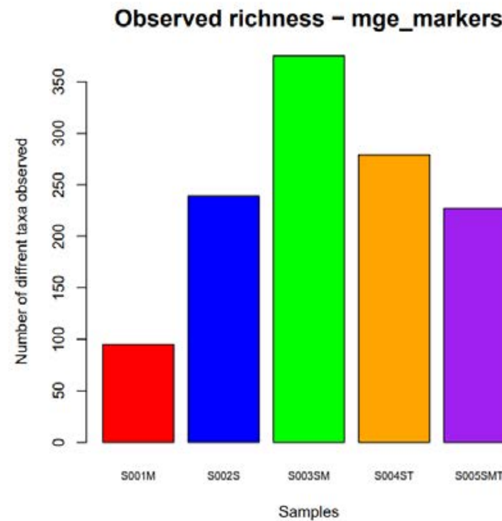




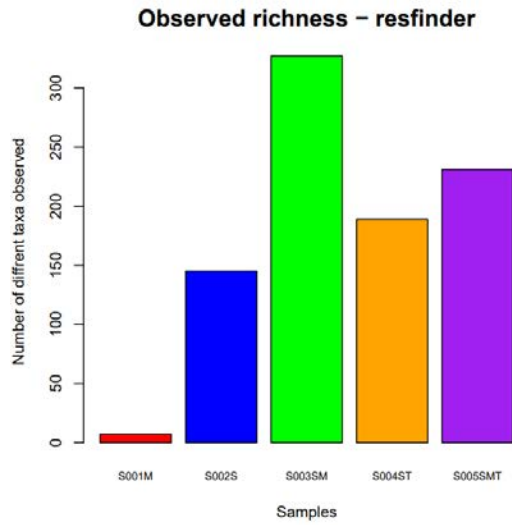


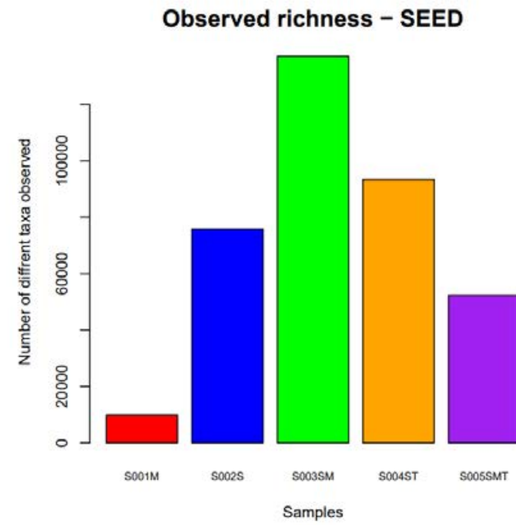


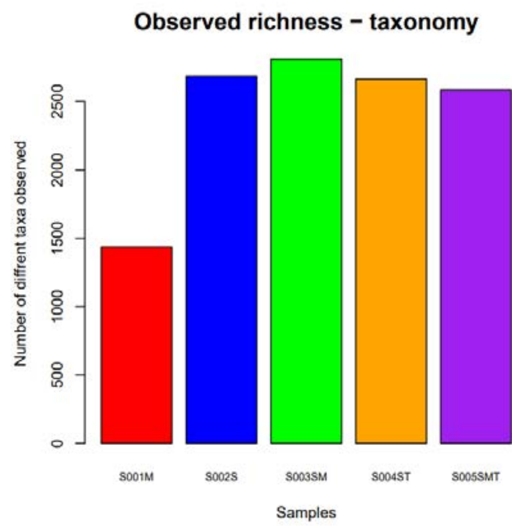


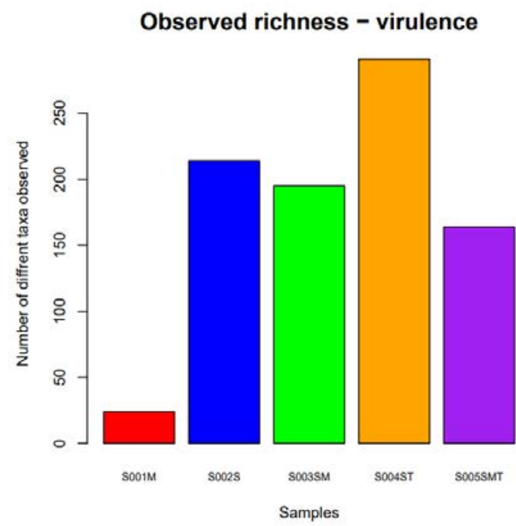


Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

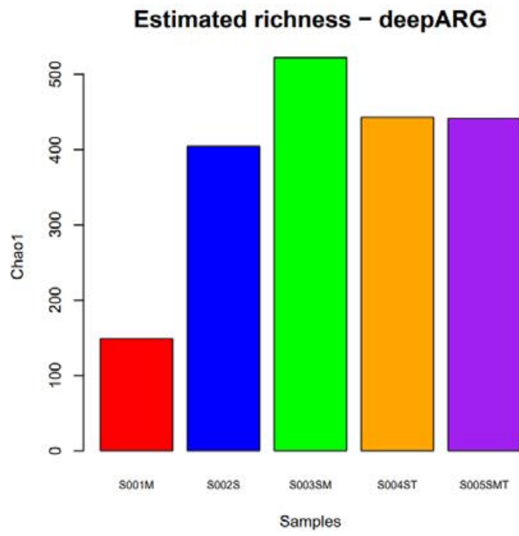


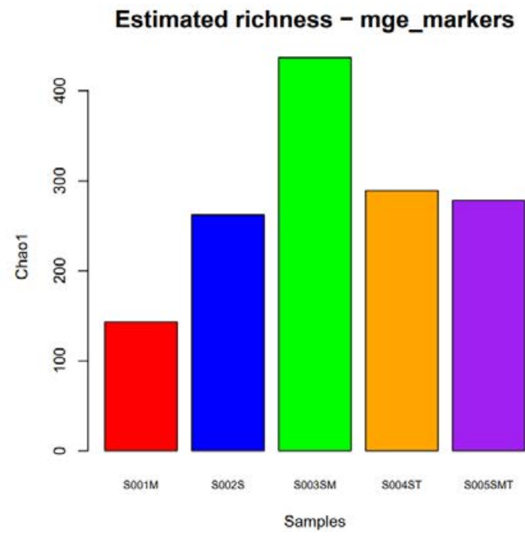


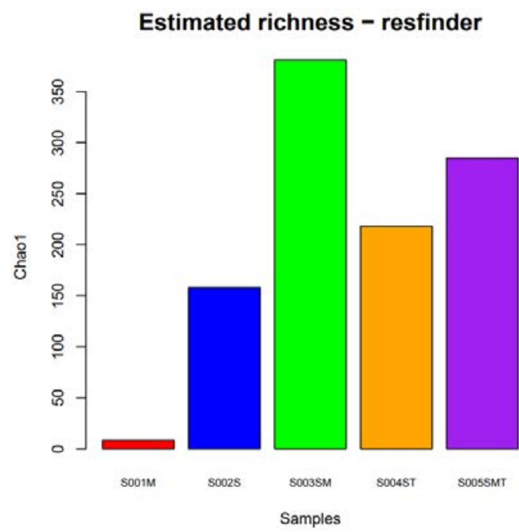


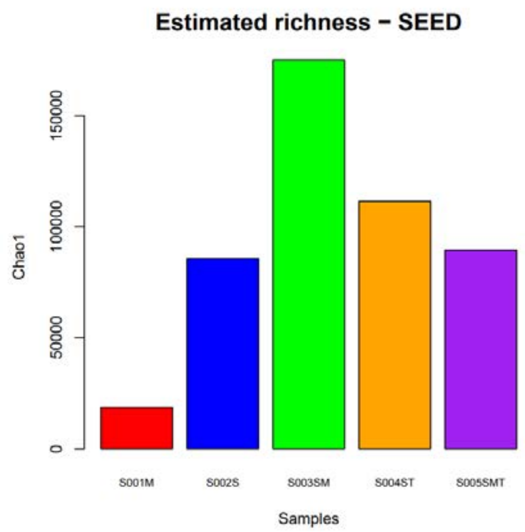


Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

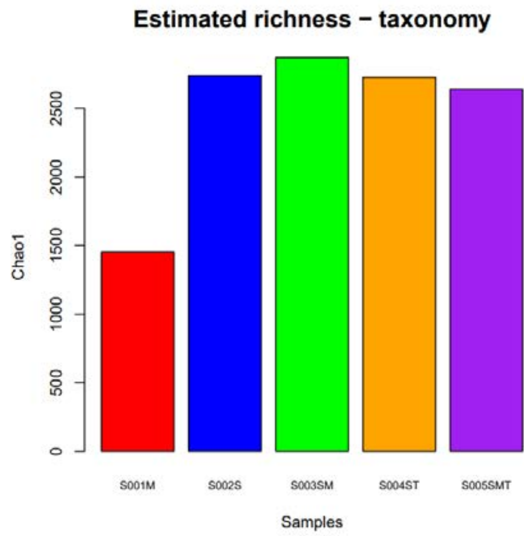


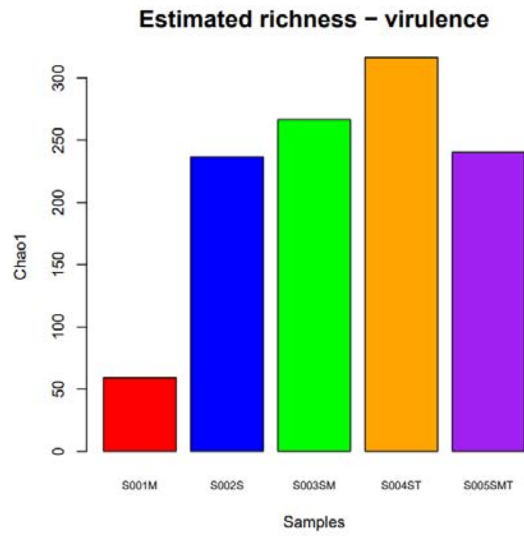


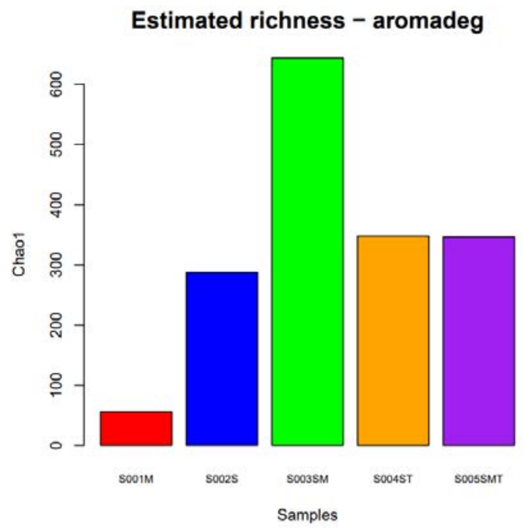


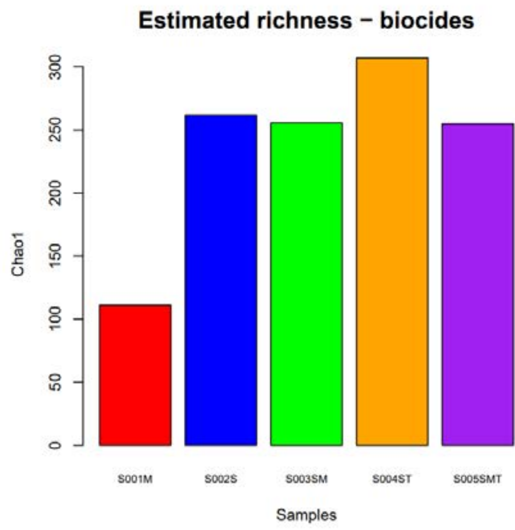


Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

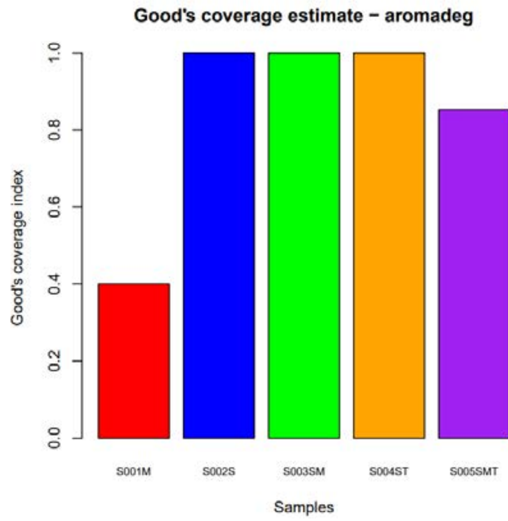


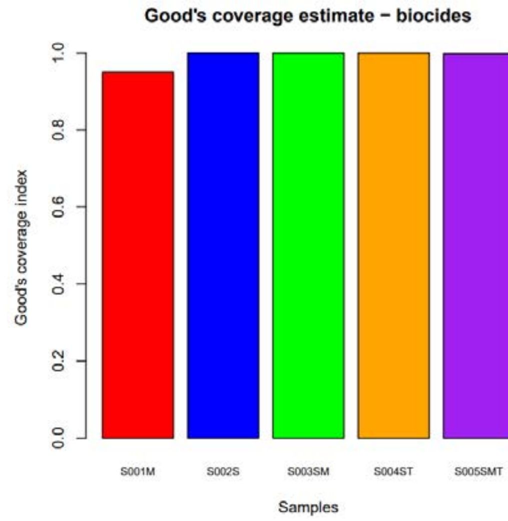


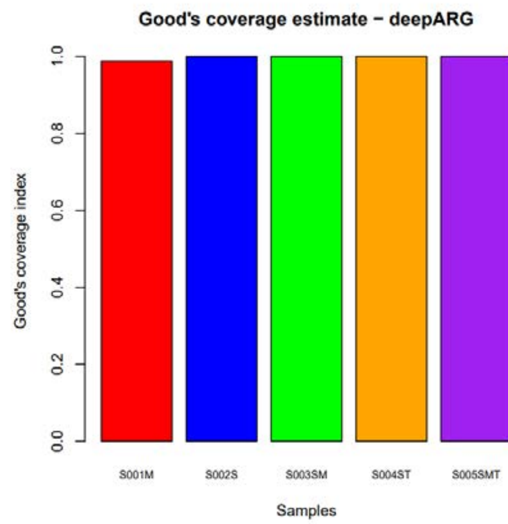


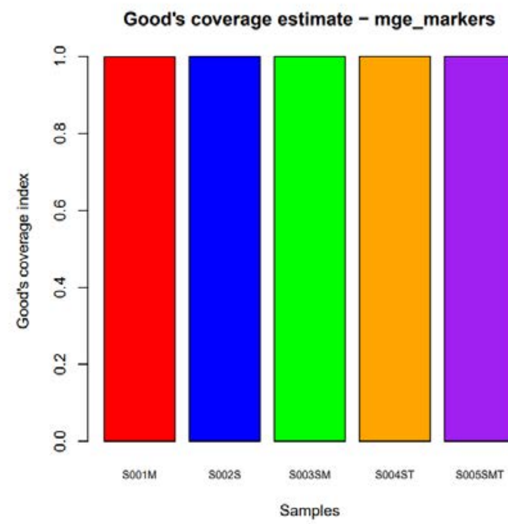


Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

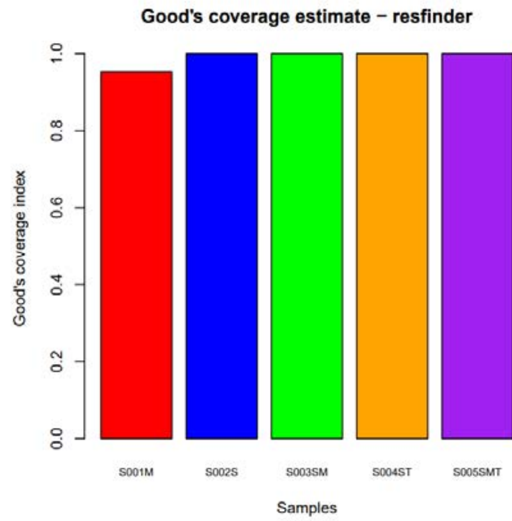


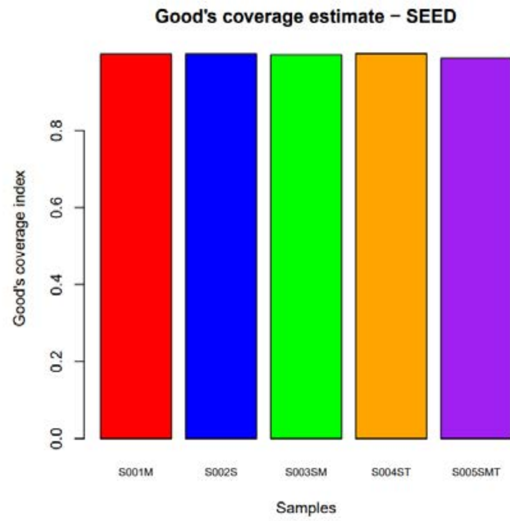


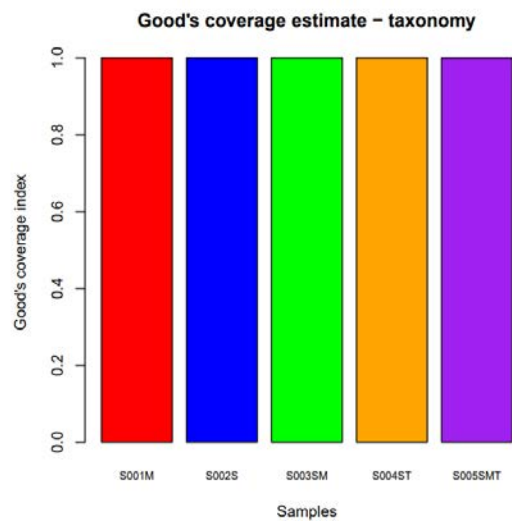


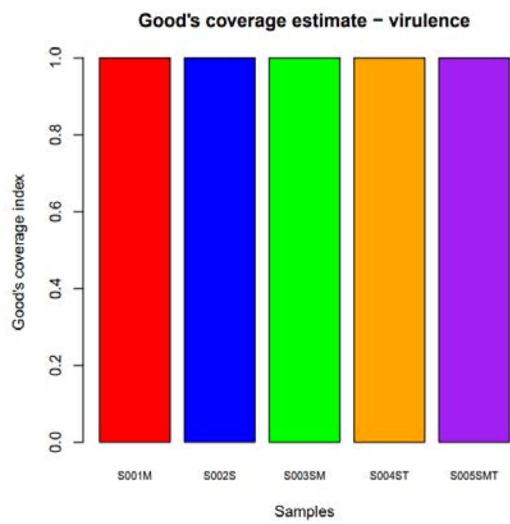


Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89





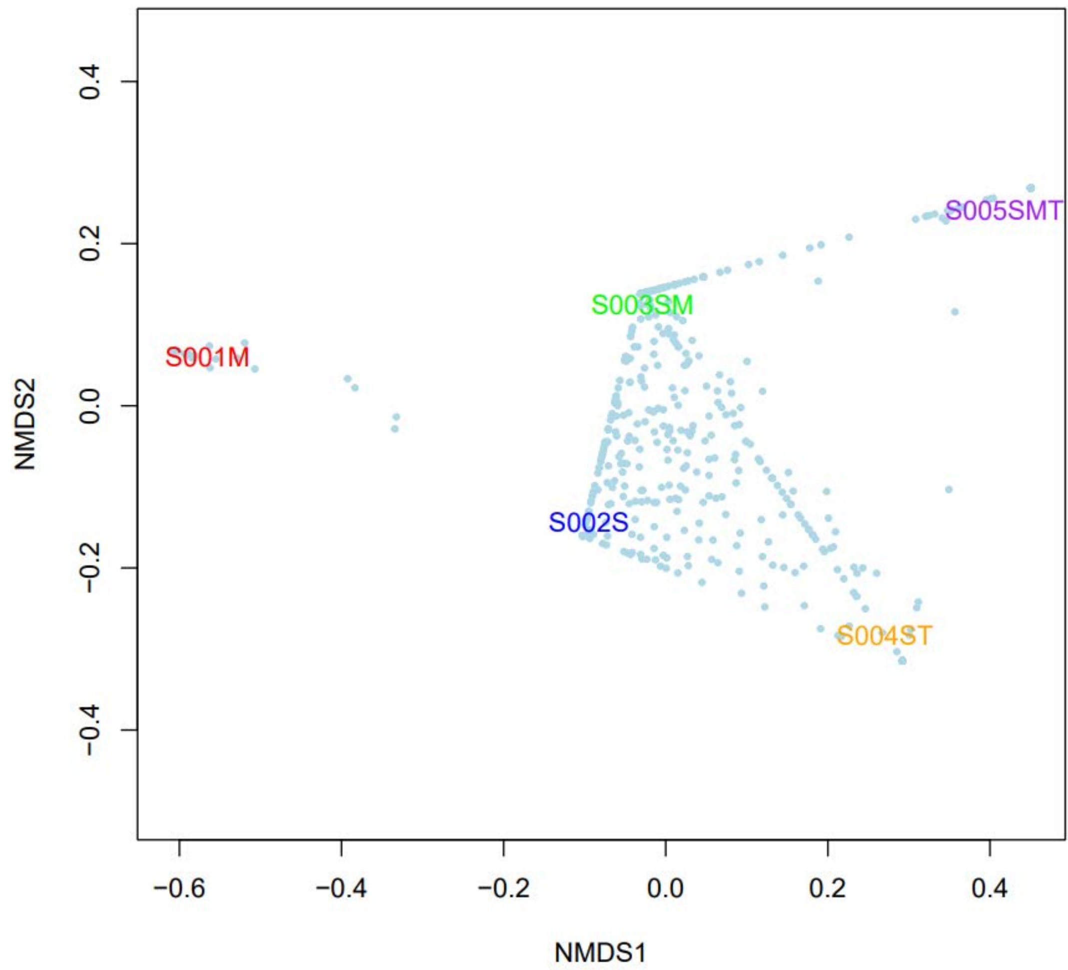




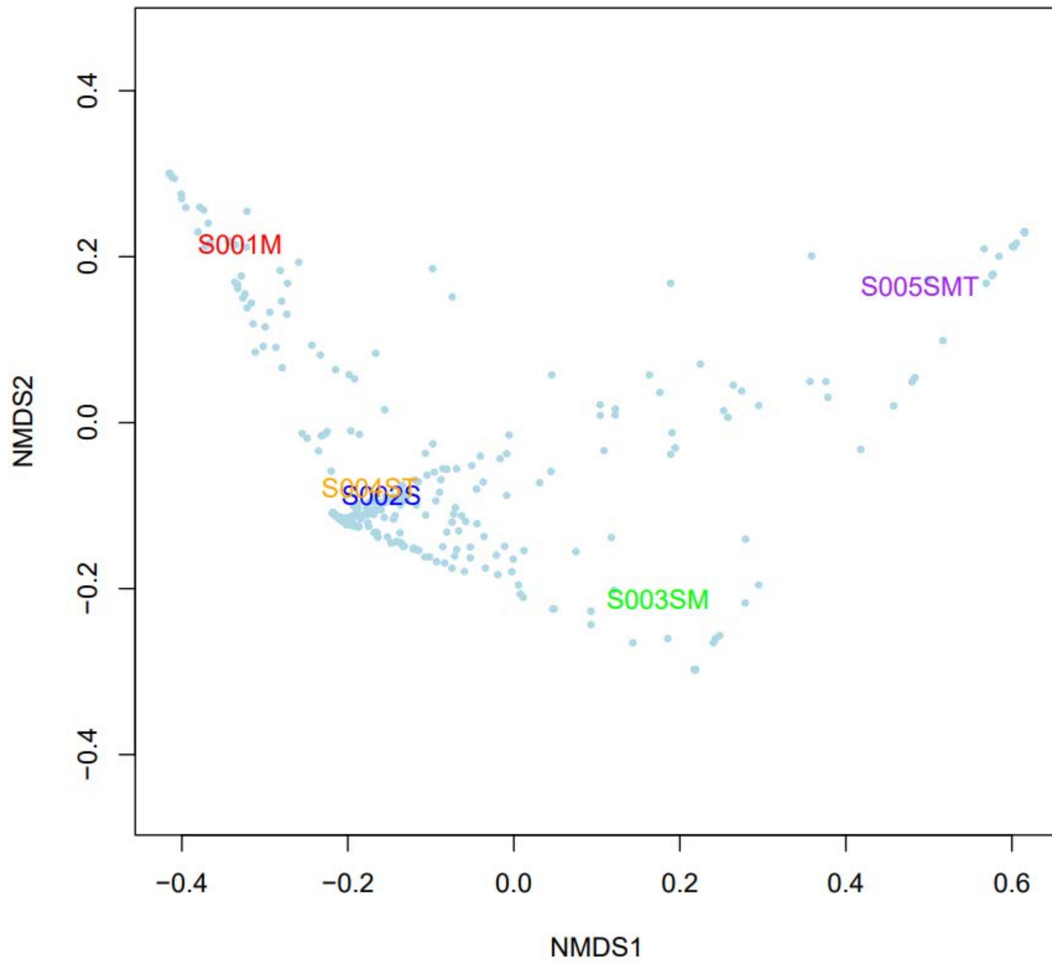
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

6.2 Supplementary II

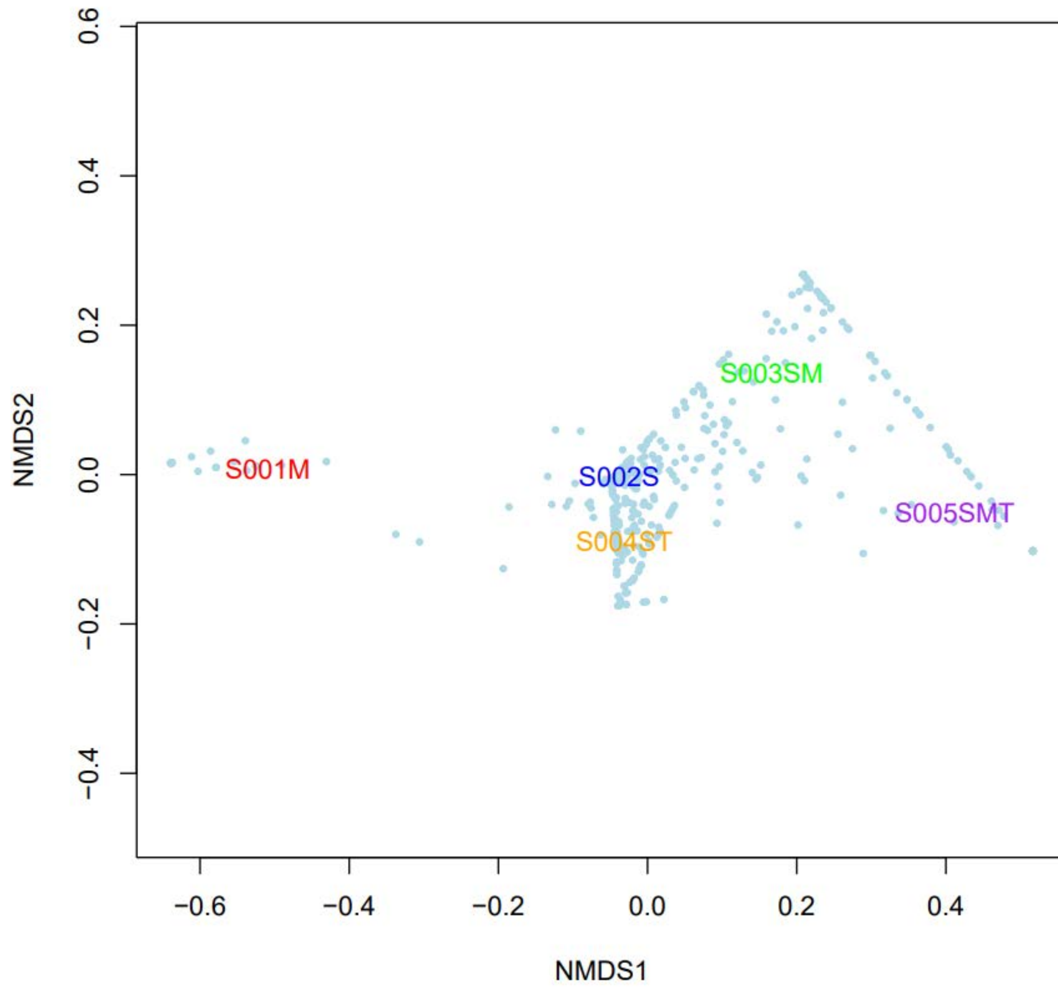
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89



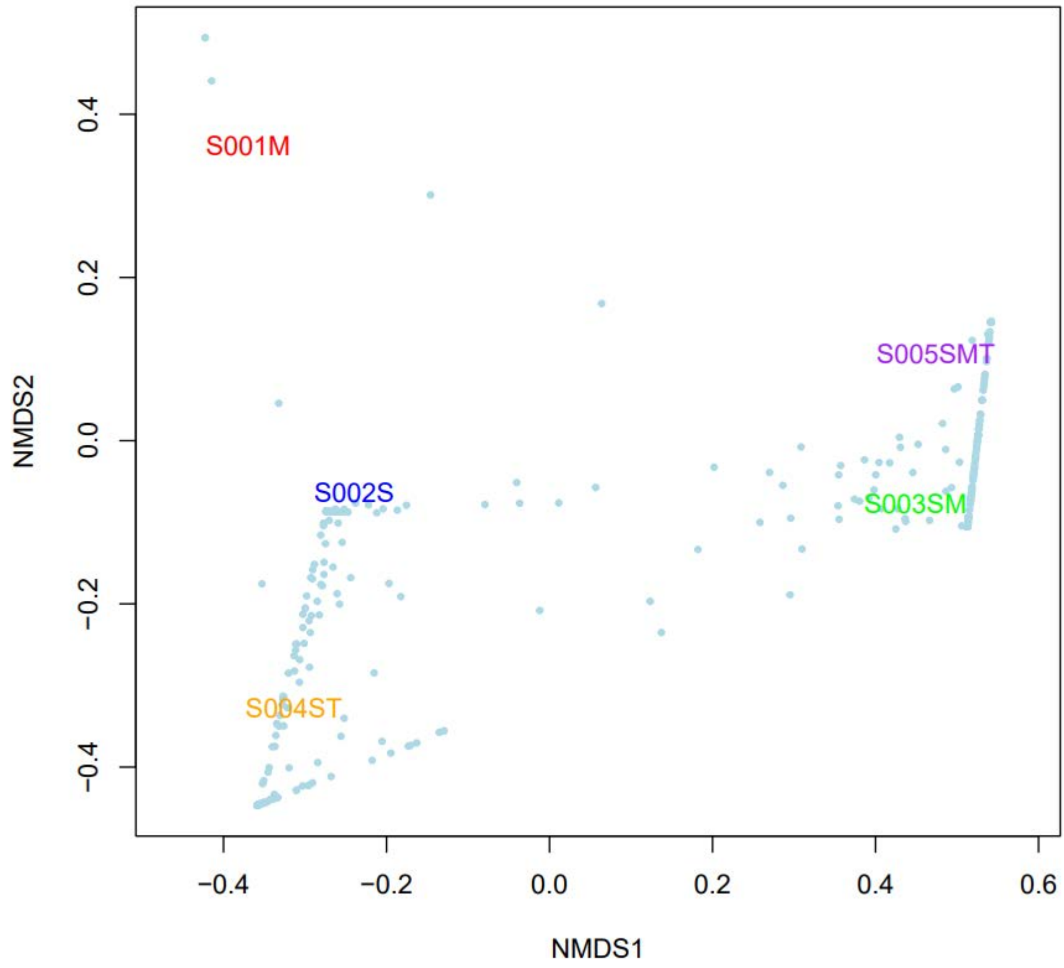
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89



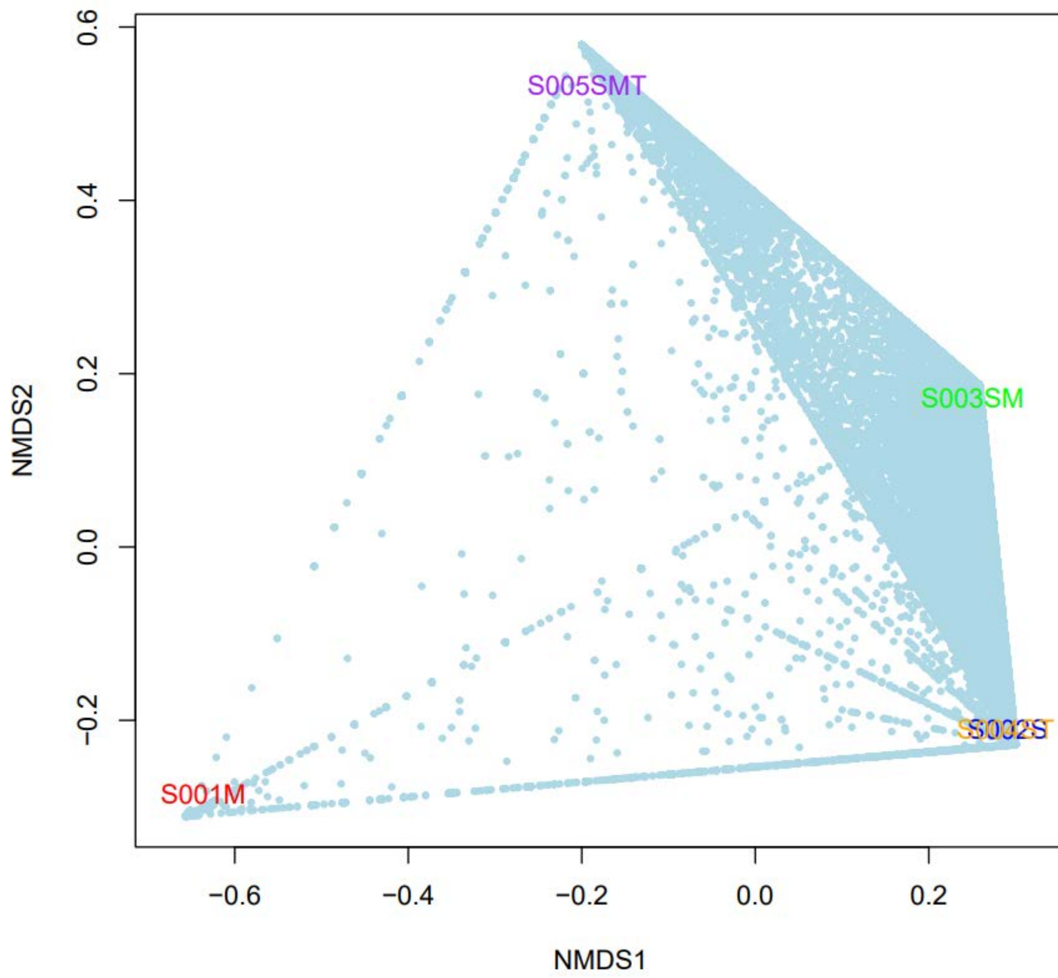
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89



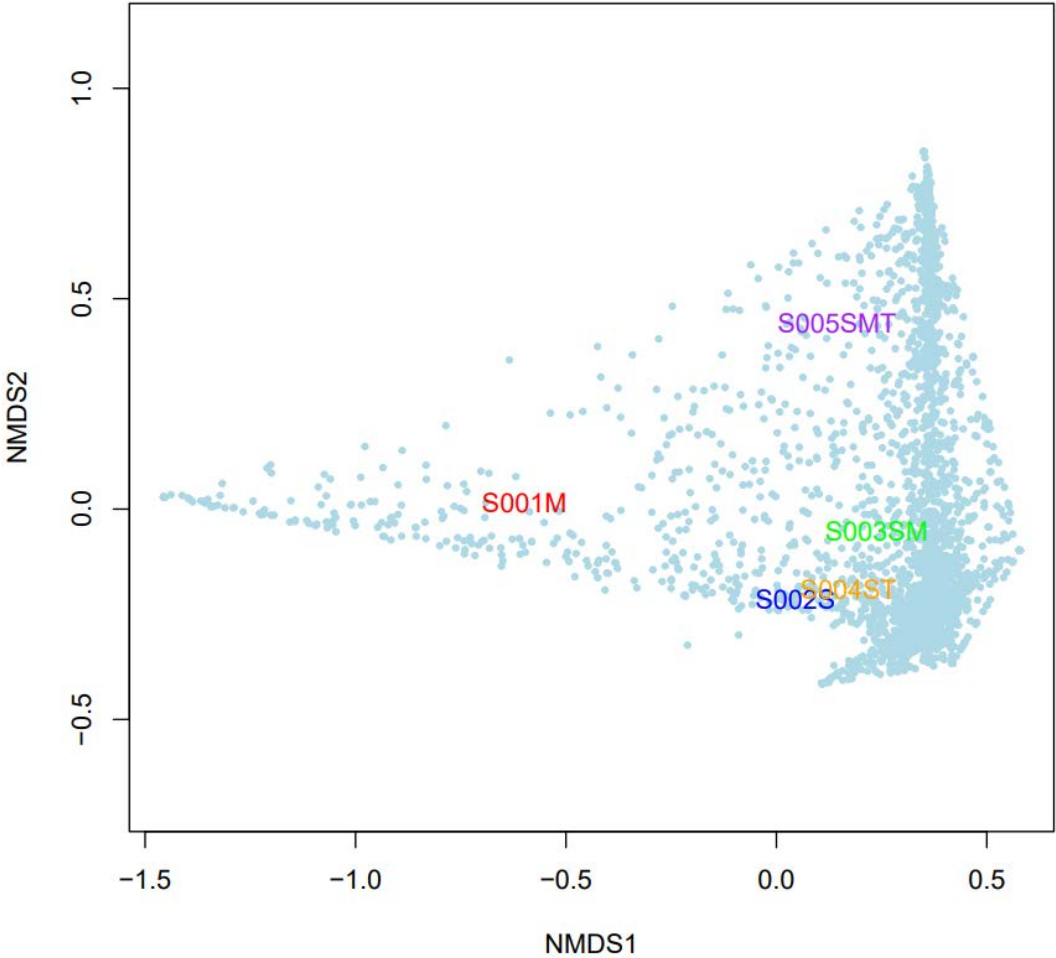
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89



Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

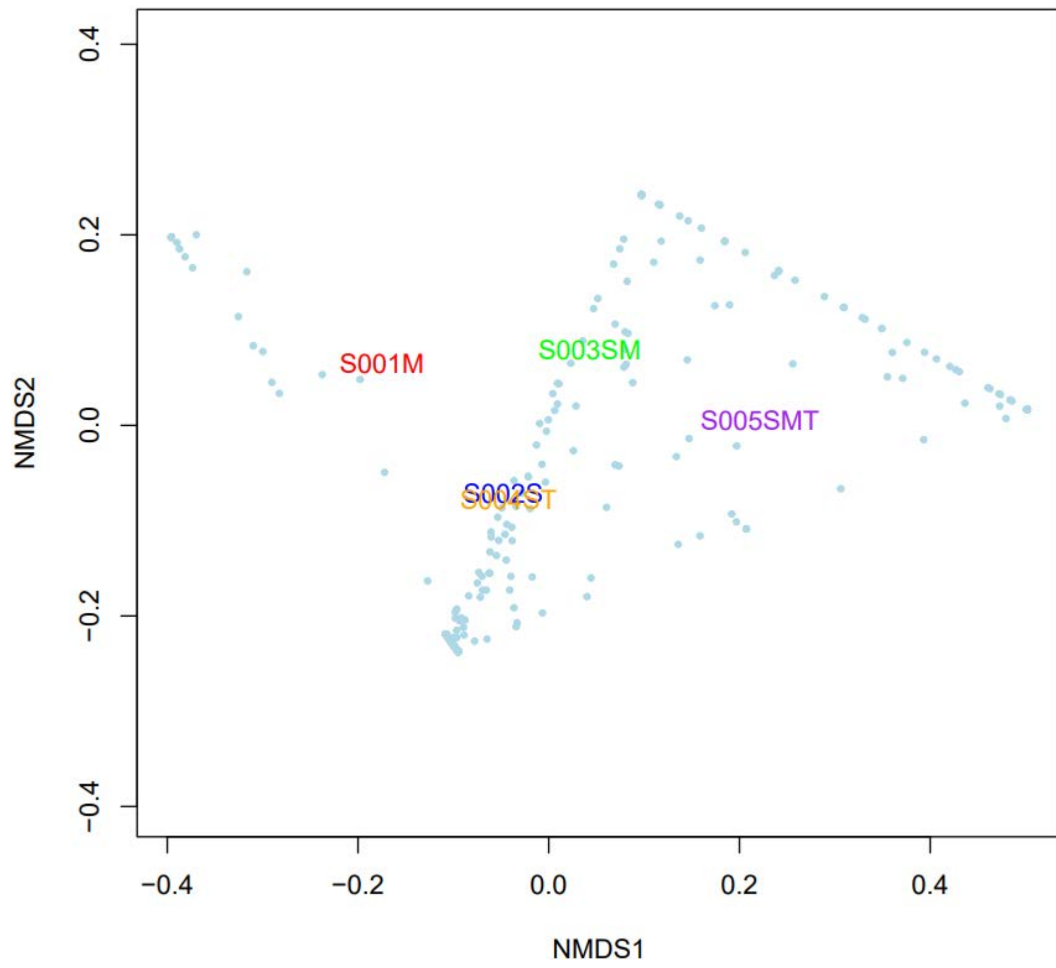


Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89



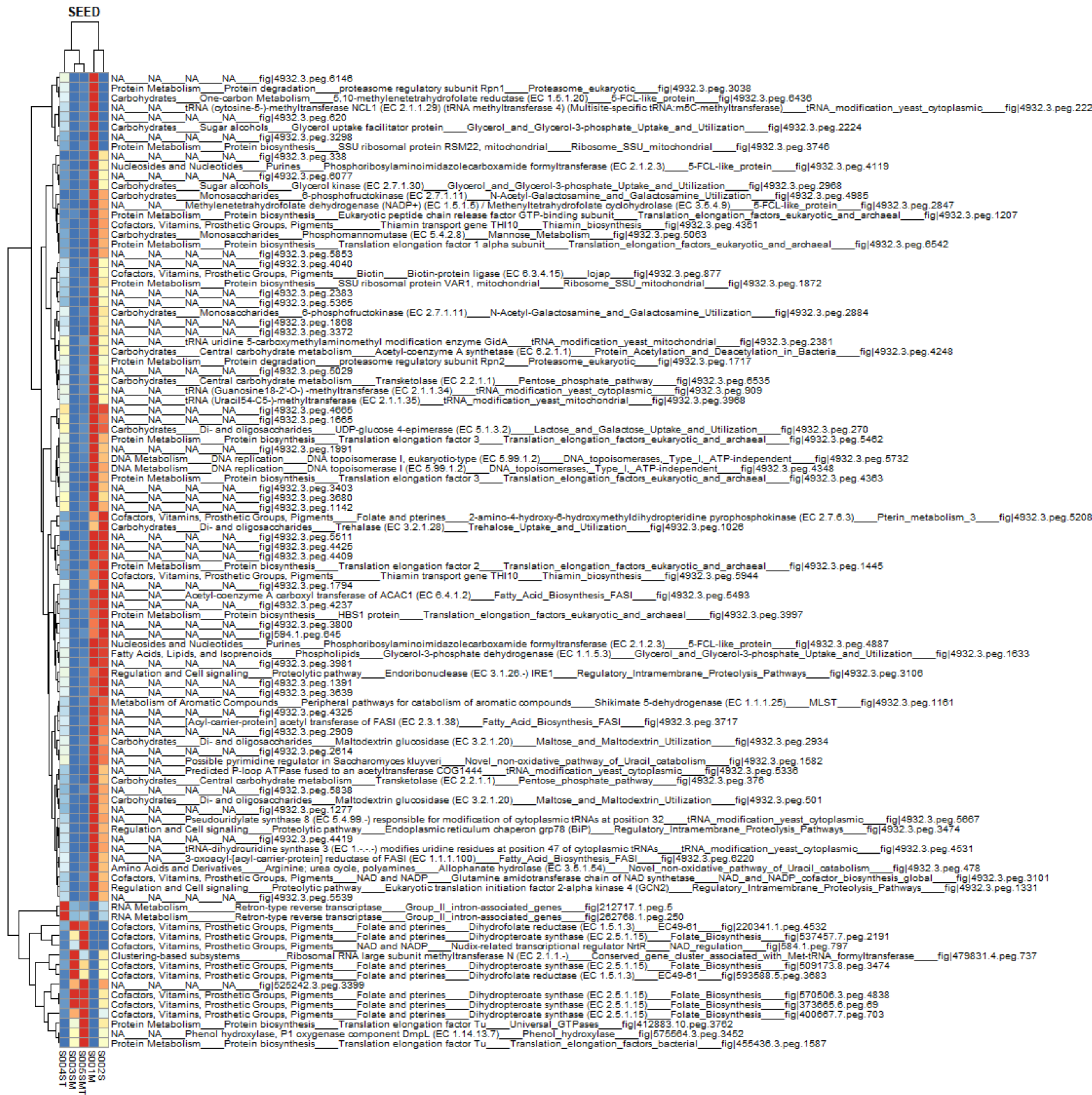
Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

NMDS Plot − virulence

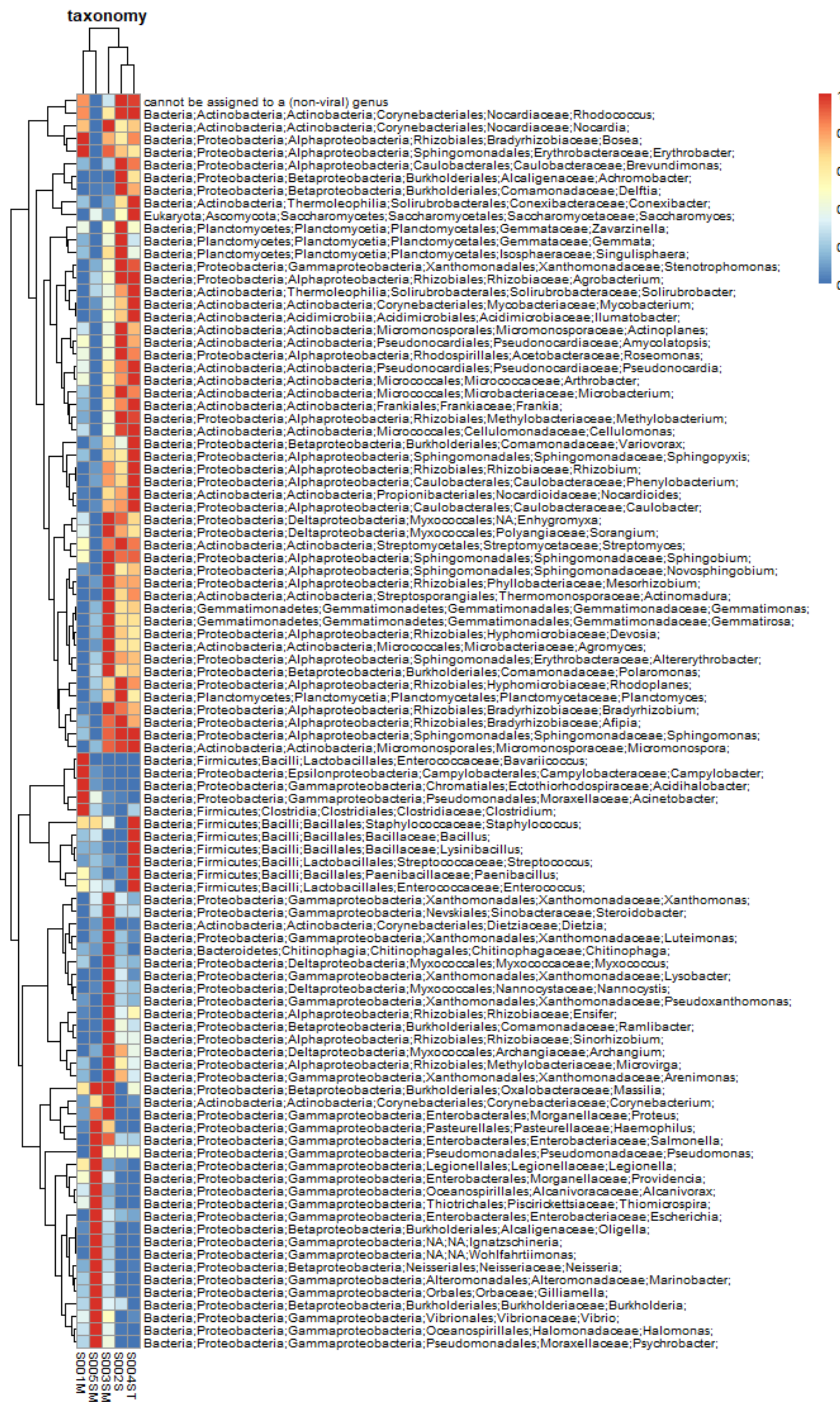


Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

6.3 Supplementary III



Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89



Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89

Institutional Repository - Library & Information Centre - University of Thessaly
15/05/2026 13:25:28 EEST - 18.97.14.89