Hungarian University of Agriculture and Life Sciences

Characterization of the microbiota of irrigation water used in vegetable farming applying MALDI-TOF MS

DOI: 10.54598/003490

Doctoral thesis

Botond Bendegúz Surányi

Budapest

2023

Doctoral School

Name: Doctoral School of Food Science

Field: Food Science

Head: Livia Simon Sarkadi, DSc

Department of Nutrition Science

Institute of Food Science and Technology

Hungarian University of Agriculture and Life Sciences

Supervisors:

Csilla Mohácsi-Farkas, PhD

Department of Food Microbiology, Hygiene and Safety, Institute of Food Science and Technology, Hungarian University of Agriculture and Life Sciences

Tekla Engelhardt, PhD

Digital Food Chain Education, Research, Development and Innovation Institute, University of Veterinary Medicine Budapest

The applicant met the requirement of the Ph.D. regulations of the Hungarian University of

Agriculture and Life Sciences and the thesis is accepted for the defense process.

.....

Head of the Doctoral School

....

Supervisor Supervisor

TABLE OF CONTENTS

Abbreviations

ANOVA - ANalysis of VAriance

API- Analytical profile index

BHI - Brain Heart Infusion

CDC - Centers for Disease Control and Prevention

CMBT - 5-chloro-2-benzothiazolethiol

DA - Discriminant Analysis

DAEC - diffusely adhering *E. coli*

ddNTPS - dideoxynucleoside triphosphates

DNTPS - deoxyribonucleotide triphosphates

EAEC - enteroaggregative *E. coli*

EAST-1 - enteroaggregative heat-stable toxin 1

EFSA - European Food Safety Authority

EHEC - enterohemorrhagic *E. coli*

EIEC - enteroinvasive *E. coli*

EMB - Eosin methylene blue agar

EPEC - enteropathogenic *E. coli*

ETEC - enterotoxigenic *E. coli*

EtOH - Absolute Ethanol

FA - ferulic acid (trans-4-hydroxy-3-methoxycinnamic acid)

HCCA/CHCA - α-Cyano-4-hydroxycinnamic acid

LT - heat-labile toxin

MALDI-TOF MS - Matrix-assisted laser desorption/ionization time of flight mass spectrometry

MDR - MultiDrug-Resistant

MRSA - Methicillin-resistant *Staphylococcus aureus*

NAAT - Nucleic Acid Amplification Tests

NMEC - neonatal meningitis-causing *E. coli*

PCA - Principal Component Analysis

PCR - polymerase chain reaction

Pet - Plasmid-encoded toxin

R2A - Reasoner's 2A agar

SepEC - septicemic *E. coli*

ShET1 - *Shigella* enterotoxin 1

ST - heat-stable toxin

STEC - Shigatoxigenic *Escherichia coli*

TFA - Trifluoracetic acid

TSA - Trypticase Soy Agar

UPEC - uropathogenic *E. coli*

VRBD - Violet Red Bile Dextrose agar

WHO - World Health Organization

XLD - Xylose Lysine Deoxycholate agar

YEA - Yeast Extract Agar

1 INTRODUCTION

In order to intensify and guarantee the agricultural productivity and thereby to be able to feed the world's rapidly growing population, the accessibility of high-quality irrigation water has become very important. In parallel the limited water resources and unpredictability of precipitation lead to an escalated usage of poorly characterized sources of water, which is directly linked to a higher prevalence of foodborne diseases. Recycled and microbiologically non-characterized waters are increasingly applied as irrigation water in cultivation in order to cope with water limitation due to climate change and to support rapid population growth. Crops can be contaminated with potentially harmful microorganisms at any of the several steps in the food production chain, during primary production, at processing stage and during preparation as in each step water plays a crucial role. At farm level, one of the major sources of food-borne pathogens is insufficient quality irrigation water which can be contaminated by sewage overflows, polluted storm- and agricultural runoffs or even by fecal pollution of wildlife (Gu et al. 2013; Uyttendaele et al. 2015). Moreover, serious bacterial pathogens such as *Listeria monocytogenes*, verotoxigenic *Escherichia coli*, *Salmonella* spp., *Escherichia coli* O157:H7 are able to survive and even grow in contaminated irrigation water (Cevallos-Cevallos et al. 2014; Falardeau et al. 2017), while the reported numbers of food- and waterborne outbreaks are also increasing. According to the European Food Safety Authority (EFSA) growing numbers of outbreaks, cases, hospitalizations and deaths related to food of non-animal origin were observed (European Food Safety Authority 2019). Leafy vegetables irrigated with contaminated water are considered to be a common cause of human gastroenteritis, due to the presence of microbial pathogens. In the United States, 22.8%-46% of foodborne illnesses were related to fresh produce such as fruits, fungi, leafy vegetables, sprout vegetables in the period of 1998-2008 (Uyttendaele et al. 2015). Moreover, Turner et al. (2019) analyzed two decades (1996-2016) in the US where they observed 46 outbreaks causing 2240 confirmed cases, where the affected food matrices were romaine lettuce and spinach. In addition, according to the EFSA, 31 outbreaks related to vegetables and juices causing 626 cases. Moreover, 48 waterborne outbreaks were connected to the consumption of tap and well water resulting 1969 cases in 2019 (EFSA 2019, EFSA 2021a). Besides, the interest in consuming fresh produce has been exponentially grown due to the promotion of healthy eating involving the consumption of 5-7 portion of such foods per day (Betts 2014).

Therefore, identifying and characterizing the microorganisms or even the complete bacterial community of irrigation water used for food production and its environment can prevent the increasing numbers of the cases. In order to ensure that, fast and reliable detection and identification of food- and waterborne bacteria should be an attainable option. Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS), a rapid microbiology technique involving laser energy absorbing matrix to create ions from larger molecules, has been applied in the field of microbiology for its fast, accurate and inexpensive nature. Although mass spectrometry based microbial identification dates back to the 1970s, the potential and usage of MALDI-TOF MS have only been realized in the last decades. Its use in microbiology is mainly centered around microbial diagnosis in clinical bacteriology but it is increasingly utilized not only in environmental bacteriology to identify food- and waterborne bacteria but to detect antibiotic resistance in bacteria and for bacterial strain typing as well. As identifying bacteria by 16S rRNA gene sequencing, the golden standard of bacterial identification, requires trained personnel and lengthy processes not to mention its higher cost of identification therefore it is not intended to use for fast identification. For this reason, MALDI-TOF MS can be a promising tool for environmental monitoring of the irrigation water used for food production and its environment.

2 OBJECTIVES

The aim of my PhD thesis is to form a comprehensive picture about the bacterial quality of irrigation waters and its surrounding environment in Hungary. To achieve this, MALDI-TOF MS was used to identify bacteria from different environmental matrices such as ground water, running water, lakes, manures and vegetables. Besides, some technical attributes of MALDI-TOF MS and the best culture media to identify waterborne bacteria by this technique were also analyzed. Furthermore, the efficacy of MALDI-TOF MS regarding identifying waterborne bacteria was tested and compared with Sanger sequencing. In addition, monitoring of irrigation water via culture-dependent and culture independent techniques were also performed.

To achieve these objectives, I set the following tasks:

- Bacteria were isolated and identified from different samples from the food production chain with MALDI-TOF MS (well, running- and still water used for irrigation, vegetables and manure).
- Testing sample preparation methods of MALDI-TOF MS of extended direct transfer and direct transfer procedure to identify waterborne bacteria.
- Examining different culture media to identify waterborne bacteria by MALDI-TOF MS.
- Applying multivariate statistical methods to differentiate bacterial strains using MALDI-TOF MS.
- Analyzing the effect of culture media via multivariate statistical methods on the mass spectra of bacterial strains.
- Testing the efficacy of MALDI-TOF MS regarding waterborne isolates against Sanger sequencing.
- Analyzing the whole (cultivable and uncultivable) bacterial microbiome of irrigation water by both MALDI-TOF MS and 16S rRNA amplicon sequencing.

3 LITERATURE OVERVIEW

3.1 Foodborne outbreaks and foodborne pathogens

3.1.1 Foodborne outbreaks

Foodborne illness is any illness which is the consequence of consuming food contaminated by pathogenic bacteria, viruses, parasites, prions or containing bacterial or fungal toxins. Symptoms usually include vomiting, fever and diarrhea resulting dehydration. Foodborne illnesses or foodborne outbreaks - when at least two people infected from the same source of food - are burden on the healthcare and on the economy as well. As such food recalls due to *Listeria* contamination and deaths related to it cost about 2.8 billion dollars annually in the USA. Moreover, the usually occurring 1 million cases and 378 deaths related to *Salmonella* infections cost about 3 billion dollars in the USA (Bhunia 2018).

In the EU, the most commonly reported zoonotic infection was campylobacteriosis with 120 946 cases, followed by salmonellosis (52 702), yersiniosis (5668), shiga-toxigenic *E. coli* (STEC) infections (4446) and listeriosis (1876) in 2020 (European Food Safety Authority 2021b).

According to EFSA (2021b), 23 outbreaks were related to the consumption of food of non-animal origin. 12 of them were caused by 'vegetables and juices and other products thereof' with 32.1 cases/outbreak, these outbreaks resulted significantly more cases than 'foods of animal origin' (14 cases/outbreak). Bacteria or toxins involved in the former cases were *Bacillus cereus* (three outbreaks), *Clostridium botulinum* and *Clostridium perfringens* (two outbreaks each), *Salmonella* Kedougou, norovirus and *Clostridium parvum* with one outbreak each as well as lectin with two outbreaks. 'Fruit, berries and juices and other products thereof' were observed in two outbreaks caused by *Salmonella* Enteritidis and *Salmonella* Muenchen, in Poland and Germany, respectively with the latter outbreak involving 161 cases with 37 hospitalizations (European Food Safety Authority 2021b). Regarding water related outbreaks 9% of the 11 strong evidence based *Campylobacter* involving outbreaks were categorized into 'tap water, including well water'. Moreover, regarding the 34 Shigatoxin-producing *E. coli* outbreaks, 5 were considered to be strong evidence based of which 40% were related to 'tap water, including well water'. In addition, 13% of the 15 strong evidence based *C. perfringens* related outbreaks, while 33% of the 6 strong evidence based *C.*

botulinum related outbreaks were related to 'vegetables and juices and other products thereof'. In 2020, 35 waterborne outbreaks, decreased by 27.1%, have been reported in the EU of which 26 were related to the consumption of 'tap water, including well water', 8 were connected to 'drinks, including bottled water' while 1 outbreak was unspecified. In Hungary, the most common zoonotic infectious agent was *Campylobacter* with 4461 cases (45.7 notification rate), followed by *Salmonella* (4461 cases, 30.3 notification rate), *Listeria* (32 cases, 0.33 notification rate), STEC (8 cases, 0.08 notification rate) (European Food Safety Authority 2021b). Therefore, as it can be seen from the aforementioned outbreaks, to provide safe food production from farm to fork, monitoring and characterizing the bacterial diversity of irrigation water used for crops produced for human consumption are inevitable to help preventing and reducing foodborne infections.

3.1.2 Major foodborne pathogens

3.1.2.1 Bacillus cereus

Bacillus cereus, first described as *Vibrio subtilis* as a soil-borne organism by Christian Gottfried Ehrenberg in 1835, is a member of *B. subtilis* group of the *Bacillaceae* family and the genus *Bacillus*. Nowadays, it has been utilized as a Gram-positive model organism for genetics and physiology studies. *B. cereus* and other species of genus *Bacillus* are ubiquitous and widely distributed in nature including air, dust, soil, water, plants, animals and humans (Bhunia 2018). Its optimal growing temperature is 30 °C and grows well forming rod-shaped colonies on Trypticase Soy Agar (TSA), Columbia Blood Medium and Nutrient Agar (Reimer et al. 2022). *Bacillus cereus* is a serious pathogen due to its abundance in the environment. Moreover, this species is able to survive in different conditions by producing biofilms, filaments, or endospores or entering into the viable but nonculturable (VBNC) state (Cayemitte et al. 2022). In the EU, *B. cereus* was the fifth causal agent in terms of foodborne outbreaks while it is estimated that this species is responsible for 1.4%-12% of foodborne outbreaks globally (European Food Safety Authority 2019; Grutsch et al. 2018). Food poisoning of *B. cereus* can be emetic, caused by the release of the potent peptide toxin cereulide, or diarrheal syndrome, caused by proteinaceous enterotoxins (e.g., hemolysin BL, nonhemolytic enterotoxin, and cytotoxin K). The former type of food poisoning is generally associated with pasta, rice dishes, beef, poultry, milk pudding, vanilla sauce and infant formulas while the latter one is linked to meat, fish, soups, dairy products, vegetables such as corn or mashed potato. *B. cereus* can also contaminate dried

products such spices, milk powder and cereal products as its endospores are difficult to inactivate due to their resistance to heat, dehydration, radiation and disinfectants. Furthermore, *Bacillus anthracis*, a closely related species to *B. cereus* and another member of the *B. cereus* group, is a serious agent in bioterrorism causing anthrax, a life-threatening systemic disease (Bhunia 2018; Jovanovic et al. 2021).

3.1.2.2 Listeria genus

The genus *Listeria* containing rod-shaped, facultative anaerobe Gram-positive bacteria comprising 17 species with the most important one being *Listeria monocytogenes*. It was first isolated by E.G.D. Murray in 1926 from a rabbit. The principle of *Listeria* species classification is based on the serological reactions of O- and H-antigen with specific antisera. *L. monocytogenes* has at least 13 serotypes classified into four lineages (Luque-Sastre et al. 2018). It is the causative agent of listeriosis, an either invasive or non-invasive disease, usually involving gastroenteritis and fever while mainly affecting immunocompromised people, elderly, pregnant women, neonates and human immunodeficiency virus (HIV)-infected people (Matereke and Okoh 2020).

Due to its ubiquitous nature, species of *Listeria* can be isolated from soil, decaying vegetation, silage, sewage while also naturally occurring in the intestines of animals. Sources of infections include meat (mainly pork), dairy products, fruits and vegetables with hot dogs, salad, smoked fish, milk, soft cheese and products made from unpasteurized milk are usually involved in outbreaks. Unheated or undercooked meat products are particularly on the risk of being the vehicle of listeriosis infections (Bhunia 2018; Matereke and Okoh 2020).

L. monocytogenes is a mesophilic bacterium with an optimum growth temperature of 37 °C. Culture media recommended to cultivate this species involve Columbia Blood Medium, Brain Heart Infusion (BHI) or Brain Heart Agar while selective culture media such as Palcam or Listeria selective agar containing nalidixic acid and trypaflavine can be used for differentiation. Moreover, *L. monocytogenes* expresses beta hemolysin to destruct red blood cells therefore help differentiate it from *Corynebacterium* (Funke et al. 1997; Reimer et al. 2022).

3.1.2.3 Staphylococcus aureus

The species of the genus *Staphylococcus* live commensal on human and animal skin, nostrils, respiratory and genital tracts as part of its natural flora. However, the species

Staphylococcus aureus, an opportunistic pathogen, can cause serious, invasive and fatal infections. *S. aureus* of animal origin live close contact with several bacteria within the same animal host, therefore exchanging genetic material among it and other Gram-positive bacteria can be realized. *S. aureus* can be transmitted to other animals and humans via droplet transmission/infection e.g., coughing, sneezing or even aerosol. Therefore, it is not surprising that a wide range of resistance genes of staphylococci are common in both animals and humans (Wendlandt et al. 2013; Schwarz et al. 2018). Previously, infections of *Staphylococcus aureus* were treated by methicillin (β-lactam antibiotic), however some strains have developed resistance against it, therefore those are called methicillin-resistant *S. aureus* (MRSA) which is a particular public health concern. Staphylococci strains have developed two ways to cope with β-lactam antibiotics, one of them is enzymatic inactivation by *blaZ*- or *blaARL* encoded β-lactamases while the other one is targeting site replacement by the gene products of the *mecA*, *mecB* and *mecC* genes (Bhunia 2018; Schwarz et al. 2018). *blaZ* gene can be found in most MRSA and methicillin-susceptible (but penicillin/ampicillin-resistant) *S. aureus* from food producing animals e.g., bovine, donkey, poultry (Fessler et al. 2010; Gharsa et al. 2012; Wendlandt et al. 2019). In the US alone, *S. aureus* caused skin and soft tissue infections are also common (nearly half million people yearly) while foodborne outbreaks involving the bacteria causes an estimated 241 000 illnesses yearly (Bhunia 2018). *S. aureus* grows well on Columbia blood medium, Trypticase soy agar forming clusters of coccal, grape-like colonies with its optimal growth temperature defined at 37 °C. Baird-Parker agar, culture medium utilized for the selective isolation of staphylococci species, is commonly used to enumerate coagulase-positive staphylococci in the food industry (Reimer et al. 2022). Species of the genus *Staphylococcus* can cause skin infections (boil, carbuncle, and furuncle) but even more severe, life-threatening infections such as life-threatening endocarditis, toxic shock syndrome, sepsis and pneumonia. *S. aureus* is also a causative agent of food poisoning with severe vomiting and cramping, in some cases coupled with diarrhea. It is of significance in veterinary medicine and animal husbandry due to its role in causing mastitis in cows and joint infection in animals, and in poultry, leading to edema and arthritis (Bhunia 2018; Alnakip et al. 2019). According to Centers for Disease Control and Prevention (CDC), foods implicated in foodborne diseases caused by *S. aureus* include milk and milk products as well as cream-containing foods and foods needing hand preparations such as salads and sandwiches (Centers for Disease Control and Prevention 2018).

3.1.2.4 Campylobacter genus

The curved, rod-shaped or s-shaped genus, *Campylobacter*, first isolated by Theodor Escherich (1886), belongs to the Gram-negative Campylobacteraceae family. This genus comprises serious foodborne pathogens and is one of the leading causes of human gastroenteritis globally. In fact, in the period of 2016-2019 the number of cases were over 220 000 each year in the EU (European Food Safety Authority 2021b). Meanwhile in the US, approximately 1.3 million cases per year can be related to *Campylobacter* infections while the incidence of such cases was 17.8 per 100 000 population with 8964 confirmed cases in 2021 (Centers for Disease Control and Prevention 2021). Species of this genus, such as *Campylobacter jejuni*, cause campylobacteriosis of which the main source of infection is undercooked meat with 20-30% of these infections comes from chicken meat consumption (Acheson and Allos 2001; Skarp et al. 2016). Campylobacteriosis, a zoonotic infection, is usually a mild, noninflammatory self-limiting diarrhea without the need of specific antibiotics. However, in more serious cases it can take weeks to recover from the bloody diarrhea caused by *Campylobacter* spp. Infection of *C. jejuni* can also cause Guillain-Barré syndrome, a condition characterized by rapid muscle weakness with damaged peripheral nervous system (Bhunia 2018; Shen et al. 2018).

Mammals and birds are the main reservoirs of *Campylobacter*, even antibiotic resistant species were found in the faeces of wild birds. In addition, it can be found in poultry, rabbits, sheep, horses, cows, pigs, even in vegetables and water (Molina-Lopez et al. 2011; Sanad et al. 2013; Wieczorek and Osek 2013). In the previous years, the most common vehicles for campylobacteriosis were broiler meat and raw milk in the EU (European Food Safety Authority 2021b).

C. jejuni is a microaerophilic species with the optimal growing temperature of 37 °C while it is advised to cultivate on Trypticase soy agar/broth with defibrinated sheep blood. There are several selective culture media to isolate *C. jejuni* including CAMP, Skirrow agars (Corry et al. 2003; Reimer et al. 2022).

3.1.2.5 Escherichia coli

Escherichia coli, first reported by Theodor Escherich in 1885 (Escherich 1885), is a Gram-negative motile rod inhabiting the intestinal tract of both humans and animals. However, it is not only part of the human and animal gut microbiota, but it has been used as a model organism to gain knowledge in bacterial physiology, metabolism,

genetic regulation, signal transduction, and the cell wall structure and function. *E. coli* is suggested to be cultivated on Nutrient Agar and TSA while its optimum growth temperature is 37 °C (Reimer et al. 2022). *E. coli* can be classified, according to its pathogenicity, into two groups with one being of those which cause infections of the gastrointestinal tract while the other group affects the kidney, urinary tract, brain and circulatory system causing septicemia. Serotypes of the first group include: enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), Shiga toxin-producing *E. coli* (STEC) and their most pathogenic subset, enterohemorrhagic *E. coli* (EHEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusely adhering *E. coli* (DAEC). The other group contains septicemic *E. coli* (SepEC), uropathogenic *E. coli* (UPEC) and neonatal meningitis-causing *E. coli* (NMEC). Several *E. coli* can produce different enterotoxins infecting the gastrointestinal tract. ETEC can release (i) heat-labile toxin (LT; LT-I and LT-II) and (ii) heat-stable toxin (ST; STa and STb) which can be inactivated by high temperature (Glenn et al. 2007; Pál 2013; Bhunia 2018). STEC produces Shiga toxin while EAEC produces ShET1 (Shigella enterotoxin 1), Pet (plasmid-encoded toxin) and enteroaggregative heat-stable toxin 1 (EAST-1) (Ruan et al. 2012).

Fecal contamination of water and food is also observed by the enumeration of *E. coli* or coliforms. Transmission of *E. coli* between humans and/or animals can involve several pathways such as direct contact, contact with animal excretion or via the food chain (Poirel et al. 2018). In the food chain, fresh produce can be contaminated by poor quality irrigation water as *E. coli* is one of the several pathogens that is able to not only survive but grow in contaminated irrigation water as well. Moreover, its effective transmission and internalization is also shown through spraying with contaminated water into lettuce, and after attaching it can survive for a longer period of time while its removal is cumbersome (Solomon et al. 2002; Berger et al. 2010). In addition, *E. coli*, along with *Salmonella* and *Listeria monocytogenes*, is one of the major causes of foodborne illnesses, hospitalizations and its serotype O157:H7 is the causative agent of 36% of the approximately 265 000 Shiga toxin-producing *E. coli* (STEC) infections each year in the US (Turner et al. 2019). In the EU, STEC caused 4446 human cases, in the period of 2016-2020. (European Food Safety Authority 2021b). Therefore, its presence in irrigation water poses a risk for any produce consumed raw especially leafy greens. Other foods of animal origin e.g., foods of animal (cattle, poultry, lamb, swine) origin such ground beef, uncooked sausages,

fermented salami, raw milk, yogurt, raw milk cheese are associated with STEC outbreaks. It can also be found in apple cider, on fruits and vegetables (mainly sprouts, salad) (Bhunia 2018).

3.1.2.6 Salmonella genus

Salmonella, another rod-shaped Gram-negative bacterial genus, consists of two species, *Salmonella enterica* and *Salmonella bongori*. The genus is named after Daniel Elmer Salmon who first isolated the bacteria from a pig in 1885. *Salmonella* is one of the leading causes of gastrointestinal diseases globally. This genus colonizes the gastrointestinal tract of different animal hosts such as pigs, cattle, poultry, dogs, cats, birds or even reptiles (Hoelzer et al. 2011; Molina-Lopez et al. 2011). Moreover, *Salmonella enterica* can even be observed in the microbiome healthy human colon with a prevalence of 3-7% (Todar 2012). Humans usually can get infected via the fecal oral route by the ingestion of contaminated foods such as seeded vegetables, eggs, poultry, beef, pork, fruits or dairy products. Moreover, the genus *Salmonella* is able not only to survive but grow in contaminated irrigation water as well (Cevallos-Cevallos et al. 2014). *Salmonella* can be distinguished according to the conditions caused by them as Typhoidal and Nontyphoidal *Salmonella*. The former group contains the Thypoid fever causing *Salmonella enterica* serotypes such as Typhi and Paratyphoid fever-causing *Salmonella* such as Paratyphi A, Paratyphi B and Paratyphi C. The group of Nontyphoidal *Salmonella* comprises *Salmonella enterica* serovars Typhimurium, Enteritidis, Newport, Heidelberg to name a few of which infections can be categorized into non-invasive e.g., food poisoning and invasive e.g., bloodsteam infections (Bhunia 2018; McDermott et al. 2018). The number of serotypes of *S. enterica* subsp. *enterica* are exceed 2000 while the identification of those is based on the surface antigens (somatic O, flagellar H and capsular Vi) (Pál 2013).

The optimal growing temperature of *Salmonella* is 37 °C and it grows well on Columbia Blood Medium, Trypticase Soy Yeast Extract Medium and Trypto Casein Soja Agar (Reimer et al. 2022). Selective culture media for the genus *Salmonella* include Eosin methylene blue (EMB), MacConkey and Xylose Lysine Deoxycholate agar (XLD) agars.

3.2 Microbial identification techniques

3.2.1 Traditional microbial identification techniques

Traditionally, bacteria have been identified by microbiological methods e.g., assessing morphological and biochemical attributes of the isolates. These culture-based methods include the use of specific, appropriate culture media for bacteria and staining methods e.g., Gram-staining. Due to the validation of these traditional methods, their reproducibility is a great advantage. Moreover, their cost-efficient and affordable nature made the conventional identification methods a staple in many hospitals and laboratories. However, traditional methods also have disadvantages, one of them is that those can only be used for microorganisms with generally known growth requirements. Another drawback of traditional methods is that those can take up to 5 days as sample collecting, preparation, data interpretation and comparison prolong the process. It usually takes 24-72 hours to perform and evaluate these tests which causes loss of crucial time in choosing the appropriate treatment. Furthermore, the differentiation of closely related species is not possible with these methods (Kelley 2017; Żukowska 2021). However, newer versions of traditional identification methods such as Analytical Profile Index (API) are able to identify bacteria in a fast and efficient way. API (bioMérieux, Marcy l'Etoile, France), invented by Pierre Janin, is a microbiology identification method based on biochemical tests. It is used for the quick identification of wide range of bacteria and yeasts (API Reference Guide 2019; Topić Popović et al. 2021).

3.2.2 Principles of MALDI-TOF MS bacterial identification

MALDI-TOF MS (matrix assisted laser desorption ionization-time of flight mass spectrometry) has become a popular technique in microbiological identification. In fact, it has revolutionized microbiological diagnostics and identification due to its fast and accurate nature. Moreover, its application significantly reduced the time needed for diagnosis in healthcare facilities therefore contributing to produce successful treatments for patients (Żukowska 2021). MALDI-TOF MS is a soft ionization technique that generates protein mass spectra mainly containing the m/z (mass to charge ratio) and the intensity values of the ribosomal proteins of the isolate. Ribosomal proteins are ancient, conserved molecules thus used for the identification of bacteria. The matrix solution is mixed with the analyzed isolate for ionization to form protein mass spectra with specific molecular weight (2-21 kDA) ranges. The

identification of the microbes is based on the detection of mass signals (Carbonelle et al. 2011; Topić Popović et al. 2021). Identification of the isolates can be performed by comparing the PMF (Protein Mass Fingerprint) of the measured microbe to databases containing PMFs, or by pairing the masses of the identified biomarkers of unknown organisms using proteomic databases. Matching PMF means comparing the mass spectrum of the isolated microbe with databases containing mass spectra of known microbes (Singhal et al. 2015; Ashfaq et al. 2022).

3.2.3 MALDI-TOF MS databases

Two of the largest systems based on mass spectral microbial identification are BioTyper® (Bruker Daltonics GmbH & Co, Bremen, Germany) and VITEK® MS Plus (bioMérieux, Marcy l'Etoile, France) (Singhal et al. 2015). Another system by Shimadzu Corporation (Kyoto, Japan), in a cooperation with bioMérieux was also developed.

Database of MALDI Biotyper is based on the Main Spectra Concept meaning that the reference library entries are stored in the system as Main Spectra (MSP). The MSPs are built on multiple measurements of a defined strain to ensure the variability of the organism. To define unknown isolates, a score for each isolate computed by counting signals in its mass spectrum is compared to the reference mass spectra and vice versa as well as correlating signal intensities of matched signals of mass spectra. The three scores acquired via such calculation are multiplied and normalized to 1 000 and log transformed. The log-score of in the range of 2-3 is considered to be reliable species identification while the score from 1.7 to 1.99 is accepted to genus level, scores below 1.7 is recognized as not reliable identification. The latest MALDI Biotyper library contains PMFs of 4274 unique bacterial species from 704 genera (2022) (Freiwald and Sauer 2009; Welker and Moore 2011; Bruker Daltonics GmbH & Co. KG 2022).

The database of VITEK MS comprises 1095 bacterial and 221 fungal species with 15556 strains used for its development. Moreover, the database is built with multiple strains for each organism to include intra-species diversity and provide highly confident identification (bioMérieux SA 2022). VITEK MS (bioMérieux SA) identifies isolates by applying a computed identification matrix called Advanced Spectrum Classifier comprising a dataset of more than 25 000 binned reference spectra. To identify an isolate, its spectrum is paired to the identification matrix with each spectra receiving a weight according to its frequency within the species and

among all other species in the database as well. That is followed by calculating the probability functions for each species associating with matching spectra to the identification matrix. Then the spectrum of the isolate is paired to the identification matrix and the summed bin weights for each species converting into a probability (Welker and Moore 2011; Garner et al. 2014).

Another database, Saramis developed by AnagnosTec/bioMérieux utilizes the concept of SuperSpectra as identifying reference spectra that were calculated by weighting peaks in a consensus spectra based on their specificity for different taxonomic levels such as genus and species. Consensus spectra are calculated from several isolates of a taxon by opting conserved mass signals. To identify an isolate, its mass spectrum is searched against SuperSpectra and the sum of peak weights is computed for matching mass signals. The sum of peak weights transformed into a confidence value which values over 80% are accepted as significant (Benagli et al. 2011; Welker and Moore 2011).

3.2.4 The application of MALDI-TOF MS for bacterial identification

MALDI-TOF MS has been mostly used for the identification of clinically relevant pathogens. It has been reported that this system, either Bruker's Biotyper or bioMérieux's VITEK 2 MS, can effectively identify clinical isolates even at species level in the range of 85.6%-98.2% (Van Veen et al. 2010; Faron et al. 2015; Hou et al. 2019; Chung et al. 2021). Garner et al. (2014) demonstrated the efficacy of MALDI-TOF MS (VITEK MS) regarding identifying clinical isolates of anaerobic Gram-negative bacteria with correct identifications of 91.7% at species level while Faron et al. (2015) achieved 98.2% correct identifications of aerobic Gram-negative bacteria at species level using Bruker's Biotyper. However, Schulthess et al. (2016) identified 53.7% of clinical isolates of fastidious Gram-negative rods. The correct identifications of Gram-positive clinical isolates are somewhat lower compared to the correct identifications of Gram-negative clinical isolates as Rychert et al. (2013) obtained 92.8%, while Garner et al. (2014) achieved 91.7% correct identifications for Gram-positive bacteria applying VITEK MS. In contrast, Chung et al. (2021) obtained 98% correct identifications of Gram-positive cocci using Bruker's Biotyper. However, Schulthess et al. (2014) achieved 77.6% correct identification of Gram-positive cocci using Biotyper. Nonetheless, according to Bizzini et al. (2010) 98% of routine clinical isolates can be identified at genus level and 90% at species level while only 1% are incorrectly identified using MALDI-TOF MS. In the study of Jamal et al. (2013)

97.2% and 94.7% correct identifications of clinical isolates of Gram-positive cocci were obtained by comparing VITEK MS and Biotyper.

Besides identifying clinically relevant pathogens, MALDI-TOF MS has also been increasingly utilized in environmental research as well. Some of the application areas of MALDI-TOF MS in environmental microbiology include microbial ecology, food microbiology, environmental biotechnology, agriculture and plant sciences as well. Strejcek et al. (2018) applied MALDI-TOF MS to identify microbes found in soils and sediments obtaining 92% and 35% correct genus and species level identification. Kopcakova et al. (2014) utilized MALDI-TOF MS to identify the microflora from waste disposal sites with an identification rate lower than 20% at species level. Moreover, MALDI-TOF MS was also used to identify specific copper resistant microorganisms from soil and water with an identification result of 97% at genus level (Avanzi et al. 2017). Furthermore, El-Nemr et al. (2019) used MALDI-TOF MS to identify bacteria isolated from a market area (e.g., vegetables, soil, air and hand palms of fresh produce handlers) at species level with 41% correct identifications. In another study, Pandey et. al (2019) identified 4.92% and 59% of psychrotolerant bacteria isolated from high altitude soil at species and genus level, respectively. Topić Popović et al. (2022) identified 184 of 321 (57%) Gram-negative bacteria isolated from water and fish samples at species level. Furthermore, Uchida-Fuji et. al (2020) showed the potential of MALDI-TOF MS in environmental microbiology as the authors were able to identify 86.2% of bacteria isolated from horses and their environment at species level.

Moreover, MALDI-TOF MS has been used in environmental studies to identify medically relevant staphylococci from air samples taken from schoolrooms (Fox et al. 2011), assess the bacterial community in a drinking water treatment plant (Sala-Comorera et al. 2017), monitor water quality in the water industry and examine the hygiene facilities (Laukova et al. 2019), identify and characterize human pathogens isolated from chicken meat and water (Elbehiry et al. 2019), isolate and screen seawater microorganisms (Ashfaq et al. 2019) and to monitor groundwater being used as drinking water (Jancova et al. 2020). Furthermore, the application of MALDI-TOF MS has also been studied in food microbiology to identify bacteria responsible for beer spoilage (Turvey et al. 2016), identify and differentiate foodborne pathogenic bacteria (Illikoud et al. 2019), investigate the presence of bacteria in honeys (Pomastowski et al. 2019).

3.2.5 MALDI-TOF MS sample preparation techniques

MALDI-TOF MS identification involves the ionization of the isolate for which the matrix solution is mixed with the sample. Different types of matrices can be used for MALDI-TOF MS identification. One of the most commonly used matrices is α-cyano-4-hydroxycinnamic acid (HCCA/CHCA), which is also the validated matrix (IVD HCCA) of Bruker Biotyper. The most frequently used matrices for bacterial identification include sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid), 2,5-dihydroxybenzoic acid, 5-chloro-2-benzothiazolethiol (CMBT), ferulic acid (trans-4-hydroxy-3-methoxycinnamic acid, FA), 2-(5-hydroxyphenylazo) benzoic acid (Pennanec et al. 2010; Topić Popović et al. 2021). Three types of MALDI-TOF MS sample preparation methods can be differentiated. The fastest and simplest one is the direct transfer procedure. This method involves adding a smear colony directly onto the target plate and immediately overlaying the samples with the previously chosen matrix or the recommended one by the manufacturer. Another method, extended direct transfer procedure, involves adding a smear colony directly onto the target plate and adding 70% formic acid onto the sample. After air-drying, the chosen matrix can be added onto sample. The third method, full extraction method, starts with adding the colonies of an isolate into HPLC grade water and absolute ethanol. Then the sample should be centrifuged and the supernatant should be removed. After this step, 70% formic can be added and mixed. After it is mixed, 100% acetonitrile should be added and the sample should be centrifuged again. Lastly, the supernatant can be placed onto target plate and after air-drying, the matrix can be added. Nevertheless, certain chemical or physical treatments can be performed before MALDI-TOF MS identification such as applying Trifluoracetic acid (TFA) to inactivate bacterial endospores (Drevinek et al. 2012) or 70% ethanol to inactivate non-spore forming bacteria (*Brucella* spp., *Francisella tularensis*, *Neisseria meningitidis*, *Burkholderia pseudomallei*) (Cunningham and Patel 2015).

Several authors have investigated the impact of different sample preparation methods on the confidence of bacterial identification. Pascale et al. (2020) found no significant differences regarding sample preparation methods identifying *Legionella* species from water distribution systems of man-made environments (hospitals, hotels, healthcare facilities, companies). However, Veelo et al. (2014) found that the identification scores of both Gram-positive and -negative bacterial strains (*Peptoniphilus ivorii*, *Propionibacterium acnes*, *Bifidobacterium longum*, *Bifidobacterium dentium*,

Atopobium minutum, *Actinomyces meyeri*, *Actinomyces graevenitzii*, *Actinomyces israellii*, *Eggerthella lenta*) derived from strain collection and clinical samples could be improved by the extended direct transfer procedure. Moreover, according to Alatoom et al. (2011), full extraction method is superior to direct transfer procedure to identify clinical isolates of Gram-positive bacteria. Barcelos et al. (2019) obtained similar results when compared extended direct transfer procedure (on-plate extraction) to full extraction to identify mastitis causing bacteria. According to the study of Schulthess et al. (2014) either formic acid on-target overlay or tube-based extraction significantly increased genus and species identification rates of Gram-positive rods.

3.2.6 Sanger sequencing

A new era has come in the microbial identification with the invention of the polymerase chain reaction (PCR), which is the basis of Nucleic Acid Amplification Tests (NAATs). Therefore, the 16S rRNA gene sequencing is considered to be the "gold-standard" of bacterial identification. The 16S rRNA gene, about 1500 nucleotides length, with the 16S rRNA itself is part of the 30S small subunit of prokaryotic ribosomes. Among its several functions, protein synthesis and structural role are the most important ones. Due to its ubiquity and conserved nature in the microorganisms, the 16S rRNA serves as the most widely applied target for phylogenetic studies involving bacteria and archaea (Woese et al. 1975; Church et al. 2020; Żukowska 2021). Moreover, it is the most commonly used molecular target for genus and species level identification in laboratories, not only due to the ubiquity of 16S, but also because the abundance of available data related to it (Clinical Laboratory Standards Institute 2018).

Sanger sequencing or Chain Termination Sequencing starts with a PCR utilizing short oligonucleotide primers to synthesize complementary amplicons to the template. It is followed by the secondary cycle sequencing of the amplicon in which a thermostable DNA polymerase, a primer designed to anneal to the template nucleic acid and small amounts of the required double-stranded DNA template are needed. Then, deoxyribonucleotide triphosphates (DNTPS; dATP, dTTP, dGTP, and dCTP), labeled with individual fluorescent markers of different spectra, are added to the reaction along with four chain-terminating dideoxynucleoside triphosphates (ddNTPS; ddATP, ddTTP, ddGTP, and ddCTP). Only one of the four dideoxynucleotides is added to each reaction, while the other added nucleotides are ordinary ones. DNA is synthesized by DNA polymerase using ddNTPs effecting termination of sequence elongation. After

that, different length DNA strands are developed with different fluorescently labeled ddNTP (A, T, C, or G) by the cycle sequencing reaction (Sanger et al. 1977; Church et al. 2020). The fluorescence cycle sequencing reaction produces a single-stranded DNA (ssDNA) fragment mixture which is loaded into a polyacrylamide gel capillary located in an automated genetic analyzer in which electrophoresis separates the fragments. The fragments are read by a fluorometric detector to create an electropherogram of the DNA sequence (Applied Biosystems 2009; Wallis and Morrell 2011).

PCR coupled with Sanger sequencing, targeting the 16S rRNA gene and pairing the gene sequences of isolates with classified references in generally known databases is now a routinely applied method to identify bacteria. However, the need of trained laboratory personnel, high costs and being time-consuming, make it difficult and inadequate for rapid identification (Alnakip et al. 2019; Żukowska 2021).

3.2.7 16S ribosomal RNA databases

The golden standard of bacterial identification and classification are based on pairing the 16S rRNA gene sequences of isolates with classified references in generally known databases. Raw 16S sequences of isolates need editing which include reviewing and trimming poor-quality data from the 3’ and 5’ of sequences. To identify the sequences, a fast rapid search algorithm can be run. As such BLAST (basic local alignment search tool), developed by the NCBI (https://blast.ncbi.nlm.nih.gov/), search is used to match the sequences to certain species or genera against one of the reference databases. Several factors can influence the success of a BLAST search including match accuracy, match length, match consistency and match differentiation. Match accuracy shows the degree of similarity, which is the higher the better, between the sequence to be identified and the matching reference sequence. Previously, an isolate with a sequence similarity of >98.5% (Fox et al. 1992; Clarridge 2004; Church et al. 2020) could be accepted to be assigned to a certain species, however newer results suggest a similarity threshold of >98.65% for the same purpose (Janda and Abbott 2007; Kim et al. 2014; Strejcek et al. 2018).

Matching reference sequences should preferably encompass the full sequence of the isolate or the longest alignment possible. Pairwise or multiple alignment can also be carried out in case of mismatches at the edges of sequences. However, when an equal number of mismatches found in sequences, longer matches should be considered to

identify the isolate. By reviewing matching consistency of BLAST results, the best-matching references should include sequences with the same species name (given that numerous entries of the species being identified included in the database) or with the same genus name (given that the database comprises low number of entries of the same species). Match differentiation shows the extent of the difference of the isolate from the next closest species. Therefore, the list of matching reference sequences should include closely related species as well (Church et al. 2020).

The edited sequences are searched against available databases of 16S rRNA which should include good quality, representative sequences for all species. Databases can be differentiated by its type of curation e.g., curated and noncurated databases with the former including manually curated ones and via algorithms. One of the biggest and well-known databases is GenBank (NCBI) hosting all previously published sequences with high-quality coverage and recurrent updates, but it contains redundant entries with limited curation. It currently comprises more than 21,000,000 entries of bacterial 16S sequences derived from clinical and environmental settings. However, partial or entire genomes containing 16S sequences are being added. Another widely known database is SILVA (Max Plank Institute for Marine Microbiology), with around 5,000,000 16S rRNA genes. Other databases such as Greengenes (Second Genome Inc., University of Colorado, and University of Queensland) with around 1,200,000 16S rRNA genes, and RDP (Ribosomal Database Project) containing 3,356,809 16S rRNA sequences, are all manually added and curated ones (Quast et al. 2013; Church et al. 2020). EzBioCloud, is also a manually curated database with quality-controlled sequences used for bacterial identification, contains 66303 bacterial and archeal 16S rRNA sequences (Yoon et al. 2017).

3.2.8 Next-generation sequencing

In order to differentiate the new sequencing methods, developed in the 1990s, from the earlier ones such as Sanger sequencing, the term “next-generation sequencing” (NGS) was born. The development of NGS made it possible to sequence the whole genome of a chosen organism at once. Nowadays, several different methods by different manufacturers are available such as Single-molecule real-time sequencing (Pacific Biosciences), Ion semiconductor (Ion Torrent sequencing), Pyrosequencing (454), Sequencing by synthesis (Illumina), Combinatorial probe anchor synthesis (cPAS-BGI/MGI), Sequencing by ligation (SOLiD sequencing), GenapSys Sequencing. All of which are able to perform millions of reads per run. Depending on

the method, from 75 up to millions of base pairs (bp) of read lengths are available with a usual read length of 200-400 bp. Besides creating thousands or millions of sequences by parallelizing the sequencing process, these new methods were allowed to reduce the cost of the sequencing significantly (Hall 2007; Grada and Weinbrecht 2013). The cost per 1 million bases varies from $5 to 950 depending on the methods, however the cheapest systems include Single-molecule real-time sequencing (Pacific Biosciences) ($7.2-$43.3), Sequencing by synthesis (Illumina) ($5-$150) and Nanopore Sequencing (Oxford) ($7-$100) (Liu et al. 2012; Quail et al. 2012). The application of NGS and whole genome sequencing (WGS) has been a promising tool in pathogen identification, detection, metagenomic analyses, tracking of antimicrobial resistance, taxonomic classification of novel bacterial species and microbiome analysis (Didelot et al. 2012; Dunne et al. 2012, 2017; Church et al. 2020). One of the most significant functions of NGS and WGS is to help understanding and exploring the microbiome of clinically relevant sites of the human body. Thus, several studies analyzed the microbiome of nasal and fecal specimens (Nakamura et al. 2009), salivary glands (Lazarevic et al. 2010), the gut (Greenblum et al. 2012; Panek et al. 2018), lower and upper respiratory tract (Willner et al. 2009), vagina (Virtanen et al. 2019; Sirichoat et al. 2021), urinary tract (Bi et al. 2019) and skin (Delaleu et al. 2021).

However, NGS has been utilized not only in the field of medicine but in metagenomics, studying DNA in the field of agriculture, biotechnology, ecology and environmental remediation. Thus, this novel technology has been used to investigate the bacterial community of various aqueous environments such as activated sludge ecosystems (Saunders et al. 2016), freshwater ecosystems (Iliev et al. 2017, Jesser and Noble 2018), urban surface waters (Jin et al. 2018), heavily polluted rivers (Wu et al. 2019) or cooling water systems (Pinel et al. 2020).

3.3 Microbial diversity of environmental samples

3.3.1 Lakes and running water

Environmental matrices such as water, soil, manure contain a wide range of microbial species. These microorganisms contribute to ecology due to their ability of recycling nutrients, converting chemical elements in various ways e.g., cycles of nitrogen, phosphorus, sulphur and carbon, fixing nitrogen, providing nutrients for plant growth, degrading environmental pollutants, detoxifying or inhibiting other pathogenic

bacteria (Ashfaq et al. 2022). Moreover, the composition of microbial communities is highly depending on the analyzed environmental samples.

The biological diversity of aquatic ecosystems is heavily affected by the microbiological composition of different water bodies, such as reservoirs which can have an influence on the neighboring ecosystems. Therefore, different lakes can have similar or greatly different microbiome around the world. Avanzi et al. (2017) investigated the microbial composition of a wastewater lake in a copper mining area in Brazil. In that study, most of the isolates belonged to genus *Pseudomonas*, a widely spread genus in water bodies, and *Enterobacter* while isolates of *Stenotrophomonas* and *Ralstonia* were also present but were not so frequently cultivated. Similarly, species of *Pseudomonas* such as *Pseudomonas aeruginosa, Pseudomonas fluorescens, Pseudomonas oleovorans, Pseudomonas putida* were dominant in Lake Baikal, Russia while other isolates such as *Acinetobacter calcoaceticus*, *Bacillus megaterium* and *Micrococcus luteus* were also identified (Babich et al. 2021). Genus *Aeromonas*, another ubiquitous genus in aqueous ecosystems, was isolated with high frequency from lakes in Greece with the most dominant species were *Aeromonas bestiarum, Aeromonas salmonicida* and *Aeromonas veronii* (Anagnostopoulos et al. 2023). Besides genus *Aeromonas*, isolates of *Serratia fonticola* were also dominant in the Greek lakes. Moreover, Tekebayeva et al. (2023) isolated and identified *Arthrobacter histidinolovorans*, *Pseudomonas extremorientalis* and *Shewanella baltica* from a salty lake in Kazakhstan using MALDI-TOF MS. In the study of Gupta et al. (2022) the cultivated microbial composition of lakes in India also showed the dominance of genus *Pseudomonas* with species of *Pseudomonas tolaasii*, *Pseudomonas rhodesiae* and *Pseudomonas fluorescens* while species such as *Arthobacter polychromogenes, Bacillus licheniformis, Janthinobacterium lividum, Ralstonia eutropha, Staphylococcus aureus* were also frequent.

Pinar-Méndez et al. (2022) analyzed the microbial quality of drinking water treatment plant revealing that the analyzed samples of groundwater and river water comprised 24 different genera as *Aeromonas* (22%) and *Pseudomonas* (32%) were the most dominant genera identified by MALDI-TOF MS, respectively. Topić Popović et al. (2022) also found that genus *Aeromonas* (53.4%) was the most frequent genus in riverine freshwater and its sediment followed by *Acinetobacter* (10.5%), *Pseudomonas* (6.5%), *Providencia* (4.0%), *Shewanella* (4.0%), *Enterobacter* (3.6%) and *Proteus* (3.6%). Suzuki et al. (2018) monitored coliforms, the fundamental indicators of water

quality for fecal pollution, in sewage, river water, and ground water founding that the coliform bacterial flora were different as the dominant coliform genera were *Klebsiella*, *Enterobacter* and *Serratia*, respectively. Furthermore, in the study of Thompson et al. (2023) changes in the microbiome in Paraopeba River (Minas Gerais, Brazil) was observed due to a collapsed dam as the presence of metal-indicating bacteria (*Acinetobacter*, *Bacillus*, *Novosphingobium* and *Sediminibacterium*) and possible indicators of faecal contamination (*Bacteroides*, *Feaecalibacterium*, *Citrobacter*, *Enterobacter*, *Enterococcus* and *Escherichia*) were also increased.

3.3.2 Vegetables and Manure

Plants also host a wide variety of microorganisms as such its microbiome contains bacteria, fungi, archaea or viruses as well. Microorganisms form communities on plants and influence the health and productivity the crops as well (Ashfaq et al. 2022). The microbiome of raw eaten vegetables is of particular importance due to the fact that those vegetables are usually ingested without further processing or treatments.

The presence of Gram-negative bacteria such as *Pseudomonas marginalis* and *Pantoea agglomerans* on *Chicorum endiva* (chicory) salads was already shown previously (Nguyen-The and Prunier 1989). Moreover, serious foodborne pathogens such as *Listeria monocytogenes*, *Salmonella* spp. and *Escherichia coli* O157:H7 were also found on fresh-cut fruits and vegetables (Abadias et al. 2008). Furthermore, El-Nemr et al. (2019) identified common ubiquitous bacterial genera such as *Pseudomonas*, *Staphylococcus*, *Enterococcus*, *Bacillus* and *Escherichia*/*Shigella*/*Enterobacter* from a fresh produce market involving samples of cucumber, green onion, lettuce, parsley, and tomato. Moreover, Patz et al. (2019) found diverse bacterial composition in the phyllosphere of different plants. In that study genus *Bacillus* was found on the leaves of *Trigonella foenum-graecum* (fenugreek) and *Thymus vulgaris* (thyme) while isolates of *Staphylococcus* were found on *Coriandrum sativum* (coriander) and thyme. Isolates of *Enterococcus faecium* were also found on fenugreek and *Cichorium endivia* (endive lettuce). Santos et al. (2020) analyzed the microbial quality of raw eaten vegetables including carrot, cabbage, lettuce, mixed vegetables and spinach using MALDI-TOF MS. The most frequently occurring genera and species in that study were *Enterobacter* (*E. asburiae*, *E. cancerogenus*, *E. cloacae*, *E. kobei* and *E. ludwigii*), *Pantoea* (*P. agglomerans* and *P. ananatis*), *Rahnella* (*R. aquatilis*) and *Lelliottia* (*L. amnigena*) as those were found in 25.9%, 9.6%, 9% and 6.6% of the samples, respectively. Artimová et al. (2023)

analyzed microbial communities on freshly consumed leafy vegetables and small berries revealing that the most frequently detected and isolated bacterial species were *Pantoea agglomerans* and *Klebsiella oxytoca*, while several species from the genera *Enterobacter*, *Citrobacter*, *Serratia*, and *Raoultela* were also identified by MALDI-TOF MS.

Manure is commonly used to restore or enhance the fertility of soil as the application of it involves adding organic matter e.g., nitrogen, phosphorus or potassium to the soil. One of the three types of manures is animal manure which contains mostly animal feces. The other two are compost, a mix of decomposed plants and animal feces, and green manure which includes crops grown for increasing the organic matter content of soil. Animal manure includes the application of the feces of pigs, cattle, horses, turkeys or chicken. It is known that the gut microbiota of animals, to some extent, are more similar between closely related animals, however its composition varies considerably due to diet and its ingredients, host genetics and the structure of the gastrointestinal system (Borda-Molina et al. 2018; de Jonge et al. 2022).

In a study of Gorliczay et al. (2021) the most dominant bacterial genera in poultry manure were *Bacillus*, *Lysinibacillus* and *Pseudomonas*. Several strains of *E. coli* including O157:H7, *Proteus vulgaris* (Bae et al. 2022) and even carbapenem-resistant *Acinetobacter baumannii* have been isolated and identified from swine manure using MALDI-TOF MS (Hrenovic et al. 2019). Hamame et al. (2022) analyzed the microbial composition of different animals' manure regarding colistin-resistant bacteria. Among analyzed animals, pigs had the highest prevalence of colistin-resistant bacteria with an abundance of intrinsically colistin-resistant bacteria. The feces of chicken contained mostly Gram-negative bacteria (96%) with *Proteus mirabilis* and *P. vulgaris* being the most abundant species. The samples from cattle contained mostly Gram-negative bacteria with the most frequent species were *P. vulgaris* and *E. coli*. Regarding pigs, the distribution of Gram-positive and -negative bacteria were almost identical (53%; 47%) while the most abundant species were *Providencia heimbachae* and different species of *Proteus* (*P. vulgaris*, *P. hauseri*, *P. mirabilis*, *P. penneri*) while *Lactobacillus brevis*, *Lactobacillus curvatus* were also common.

4 MATERIALS AND METHODS

4.1 Microorganisms

4.1.1 Bacterial strains used for testing the culture media's impact on mass spectra

The impact of culture media on the protein mass fingerprint (PMF) of isolates was tested on Gram-negative and -positive bacteria. The bacterial strains, used to test culture media's impact on the PMF of Gram-negative bacteria, were *E. coli* DSM 11250 and *E. coli* ATCC 13706. Both *E. coli* strains were obtained from the Department of Food Science and Technology (DLWT), University of Natural Resources and Life Sciences, Vienna, Austria.

The former was isolated from human feces, while the latter is a strain used for water testing. Furthermore, *S. aureus* ATCC 25923 and *S. aureus* ATCC 43300 were used to test culture media's impact on the PMF of Gram-positive bacteria. The former is a quality control strain, for both identification and media testing while the latter, a methicillin- and oxacillin-resistant strain, used in susceptibility testing. Both *S. aureus* strains were obtained from the National Collection of Agricultural and Industrial Microorganisms, Budapest, Hungary.

4.1.2 Sampling and bacterial isolation

Sampling sites were chosen due to their utilization as irrigation water. Sampling was also done in different regions, where irrigation water contamination (e.g., by manure) could have occurred or from irrigated crops (e.g., corn, lettuce, onion, sorrel, spinach, and tomato) where irrigation water might have transmitted the microbes on to them. One sample was taken from each sampling site. Water samples were collected from Danube River in different locations such as Csepel (Central Hungary) and Kalocsa (Southern Hungary). Samples from Tisza River were collected in different locations such as Szolnok (Eastern Hungary) and Tiszakécske (Eastern Hungary). Samples from Vajas River were collected in Bátya (Southern Hungary). Samples from still water were also taken from two different lakes such as Szelidi-tó (Southern Hungary) and Kavicsos-tó (Central Hungary). Samples from wells, used for irrigation, were taken from Central Hungary (Soroksár) and Eastern Hungary (Debrecen, Karcag, Kengyel, Nagykunsági-channel, Rákóczifalva, Szolnok). However, from Szolnok, two samples were taken from different sampling spots with the second one being artesian water. Experiments analyzing samples from Karcag, Kengyel, Rákóczifalva and Szolnok

were done in the Fall of 2021 at Institute of Food Science at University of Natural Resources and Life Sciences, Vienna (BOKU). The MALDI-TOF MS instrument (Bruker MALDI Biotyper) was provided by the EQ-BOKU VIBT GmbH and the BOKU Core Facility Food & Bio Processing. Experiments analyzing all the other samples were done in the period of 2019-2022 at Department of Food Microbiology, Hygiene and Safety at MATE Institute of Food Science and Technology.

Water samples were collected into sterile bottles to minimize microbial contamination while vegetables samples were collected into sterile plastic bags, and both were transported in cooled state.

Two groups were formed from the vegetables in accordance with their origins as Vegetables1 were collected from Soroksár while Vegetables2 were collected from Debrecen (Table 1). Group of Vegetables1 includes onion, corn, lettuce, spinach, while group of Vegetables2 contains tomato, spinach and sorrel. In addition to the crops manure samples were analyzed as well. Manure samples originated from different swine farms located in Bátya (Southern Hungary), Békéscsaba (Eastern Hungary) and Cegléd (Eastern Hungary). Sample of Manure1 was taken from Békéscsaba (Eastern Hungary) while Manure4 was taken from Cegléd (Eastern Hungary). Manure2 and Manure3 were from the same sample spot (Bátya, Southern Hungary), however Manure2 was liquid. One manure sample was taken from each sampling site. Three samples from wells in different towns are marked as Irrigation water samples. Sampling of each group of samples (water, manure, vegetables) was performed from the Summer of 2019 to the Spring of 2022. Bacterial isolates used in this study included 311 isolates from different water, vegetables, and manure samples. The isolates were chosen based on their different morphological characteristics. Altogether 353 isolates were cultivated, but the due to unpredictable identification results of 42 isolates, the remaining 311 isolates were included in this thesis. The correct genus identification of the aforementioned 311 isolates could be predicted from the MALDI-TOF MS identification results, however in some cases the 1.7 genus level identification score were not achieved. Details of the samples and names of them as referred later are shown in Table 1.

Table 1 Origin of samples and bacterial isolates. All analyzed samples were collected from the area of Hungary (HU)

Types	Sample name	Origin of samples	Location (city, region)	Date of collection
Still water	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	10.01.2022.
	Lake2	Szelidi-tó	Dunapataj (Central HU)	17.01.2022
Running water	River1	Tisza	Tiszakécske (Eastern HU)	15.08.2021.
	River2	Tisza	Szolnok (Eastern HU)	11.01.2022
	River3	Danube	Csepel (Central HU)	10.01.2022
	River4	Danube	Kalocsa (Southern HU)	17.01.2022.
	River5	Vajas	Bátya (Southern HU)	17.01.2022
Well	Irrigation water1	Soroksár	Soroksár (Central HU)	17.07.2019
	Irrigation water2	Debrecen	Debrecen (Eastern HU)	22.07.2019
	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	08.07.2019
	Irrigation water4	Karcag	Karcag (Eastern HU)	17.10.2021
	Irrigation water5	Kengyel	Kengyel (Eastern HU)	17.10.2021
	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	17.10.2021
	Irrigation water7	Szolnok1	Szolnok (Eastern HU)	17.10.2021
	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	17.10.2021
Liquid manure	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	08.07.2019
	Manure2	Bátya	Bátya (Southern HU)	17.01.2022
Manure	Manure3	Bátya	Bátya (Southern HU)	17.01.2022
	Manure4	Cegléd	Cegléd (Eastern HU)	24.01.2022
Vegetables	Vegetables1	Soroksár	Soroksár (Central HU)	15.07.2019
	Vegetables2	Debrecen	Debrecen (Eastern HU)	22.07.2019.

4.2 Cultivation of cultures

The bacterial cultures were derived from -80 °C storage frozen stocks. Form frozen stocks, the cultures were cultivated twice on Tryptic Soy Agar plates (Merck KGaA, Darmstadt, Germany), denoted them as *F1* and *F2* plates, with 24 hours of incubation at the appropriate growth temperature of the isolates. The *F1* plates were used for a month and the *F2* plates, prepared from the *F1* plates, were used for a maximum of 1.5 weeks. For the experiments, the overnight cultures were prepared from the *F2* plates, as 4-5 bacterial colonies were inoculated in 5 mL in Tryptic Soy Broth (Merck KGaA, Darmstadt, Germany) and were grown at their optimum temperature. Overnight cultures were always prepared in the same culture medium, that would be used in the experiments.

4.3 Chemicals

The measurements needed different chemicals and reagents. Therefore, the MALDI-TOF MS system was calibrated using IVD Bacterial Test Standard, consist of *E. coli* ribosomal protein standard (Bruker Daltonics GmbH & Co, Billerica, Massachusetts, USA). The identification of MALDI-TOF MS process involves 70% formic acid derived from Honeywell International Inc. (Charlotte, North Carolina, USA). The matrix, used to solve the ribosomal protein of isolates for identification, consists of α-cyano-4 hydroxycinnamic acid matrix solution (HCCA) (Bruker Daltonics GmbH & Co, Billerica, Massachusetts, USA).

DNA extraction of previously cultured isolates was performed by Chelex Method which needed reagents such as Chelex, a chelating material, derived from Bio-Rad Laboratories (Hercules, California, USA) and distilled water (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

4.4 Culture media and cultivation of isolates

4.4.1 Culture media used to isolate bacteria from environmental samples

Bacterial isolation was performed after preparing a ten-fold serial dilution in buffered peptone water (BPW) (Thermo Fisher Scientific Inc., Oxoid Ltd., Basingstoke, UK) up to dilution 10^{-3}. The dilutions of samples, except the irrigation water samples of Karcag, Kengyel, Rákóczifalva, Szolnok, were plated in duplicates on Trypticase Soy Agar (TSA) (Merck KGaA, Darmstadt, Germany), Reasoner's 2A agar (R2A agar)

(Merck KGaA, Darmstadt, Germany) and Yeast Extract Agar (Merck KGaA, Darmstadt, Germany) plates by spread plate method. Agar plates were incubated at 30 °C for 24-48 h.

4.4.2 Culture media used to test the effect of culture media on the mass spectra

Both *E. coli* strains (DSM 11250 and ATCC 13706) were plated on R2A (Biolab Inc., Budapest, Hungary), TSA (Merck KGaA, Darmstadt, Germany) and Yeast Extract Agar (Biolab Inc., Budapest, Hungary) agar plates. Agar plates were incubated at 37 °C to grow overnight cultures. Both *E. coli* strains (DSM 11250 and ATCC 13706) were cultivated on three different culture media (R2A, TSA, Yeast Extract Agar) in 10 replicates to obtain 60 spectra. Both *S. aureus* strains (ATCC 25923 and ATCC 43300) were plated on Baird-Parker Agar (Merck KGaA, Darmstadt, Germany), R2A (Biolab Inc., Budapest, Hungary), TSA (Merck KGaA, Darmstadt, Germany) and Yeast Extract Agar (Biolab Inc., Budapest, Hungary) agar plates. Agar plates were incubated at 37 °C to grow overnight cultures. Both *S. aureus* strains were cultivated on four different culture media (Baird-Parker, R2A, TSA, Yeast Extract agar) in 10 replicates to obtain 80 spectra.

4.4.3 Culture media used for the comparison of MALDI-TOF MS and 16S rRNA gene sequencing

The bacterial isolates from the irrigation water samples of Karcag, Kengyel, Rákóczifalva, Szolnok were plated on Trypticase Soy Agar (TSA) (Merck KGaA, Darmstadt, Germany), Violet Red Bile Dextrose agar (VRBD) (Merck KGaA, Darmstadt, Germany), Reasoner's 2A agar (R2A agar) (Merck KGaA, Darmstadt, Germany) and Yeast Extract Agar (Merck KGaA, Darmstadt, Germany) plates. VRBD agar was used with the purpose of isolating *E. coli* from the irrigation water samples. VRBD plates were incubated at 37 °C for 24-48 h while plates with the other culture media were incubated at 30 °C for 24-48 h.

4.5 Methods

4.5.1 MALDI-TOF MS to identify bacteria from environmental samples

To identify the isolates, extended direct transfer procedure was used, therefore each colony of isolates was placed onto the Bruker's ground steel target plate, overlaid with 1 µL of 70% formic acid after airdrying, overlaid with 1 µL of α-cyano-4 hydroxycinnamic acid matrix solution (HCCA). Each bacterial colony was measured

two times. The identification process was done using MALDI Biotyper (Bruker Daltonics GmbH & Co, Billerica, Massachusetts, USA). MALDI-TOF MS spectra of the isolates were collected using a Microflex LT/SH (Bruker Daltonics GmbH & Co, Bremen, Germany) mass spectrometer equipped with a nitrogen laser (lambda =337 nm) at a laser frequency of 60 Hz operating in linear positive ion detection mode under MALDI Biotyper 3.0 Realtime classification (RTC) (Bruker Daltonics GmbH & Co, Bremen, Germany) and FlexControl 3.4 (Bruker Daltonics GmbH & Co, Bremen, Germany). Mass spectra were acquired in the range of 2000-21000 Da for each sample analyzed for species level microbial identification which includes the measured ribosomal proteins, which forms up to 70% of a microbial cell in that range. MALDI-TOF MS spectra were generated from 240 single spectra that were created in 40-laser-shot steps from random positions of each isolate. The system was calibrated using *E. coli* ribosomal protein standard (Bruker IVD Bacterial Test Standard, Bruker Daltonics GmbH & Co, Bremen, Germany). FlexControl and FlexAnalysis (Bruker Daltonics GmbH & Co, Bremen, Germany) were used for data acquisition and data processing, respectively. FlexAnalysis was used to preprocess mass spectra which involves baseline subtraction, smoothing and peak picking.

MALDI-TOF MS identification results were accepted at genus or species level according to Bruker's instructions. High-confidence identification indicates a log score in the range of 2.00-3.00 which means reliable identification at species level. Low-confidence identification is accepted at genus level, with a log score of 1.7-1.99. Log scores below 1.7 are considered as not reliable identifications without reaching any level. Furthermore, besides MALDI-TOF MS identification, catalase and oxidase activities of the isolates were tested and were in concordance with the MALDI-TOF MS results. Moreover, Gram-staining was used to differentiate unidentifiable isolates that generated both Gram-positive and -negative mass spectra according to the results of Biotyper. Thus, the best hit in Biotyper which matched the result of Gram-staining was accepted to the analyzed isolates.

4.5.2 DNA Extraction and Sanger sequencing of waterborne isolates

DNA extraction of the previously cultured isolates was performed by Chelex Method. Chelex solution contained 2.5 g Chelex (Bio-Rad Laboratories, Hercules, California, USA), 2.5 mL 0.01 M Tris HCL and 95 mL distilled water (Thermo Fisher Scientific, Waltham, Massachusetts, USA). A colony of each isolate was put into 500 µL Chelex solution with a sterile inoculation loop. After mixing by vortexing, the samples were

placed into Eppendorf ThermoMixer® C (Eppendorf, Hamburg, Germany) and incubated for 10 minutes at 95 °C. Then samples were centrifuged at 15 000x *g* for 30 s and the supernatant was transferred into a fresh 2 mL Eppendorf tube. After extracting the DNA of isolates, 16S rRNA gene specific PCR was performed. The applied 16S rRNA gene primers were 27F, 5′-AGAGTTTGATCCTGGCTCAG-3′ and 1492R, 5′-GGTTACCTTGTTACGACTT-3′. The PCR thermal profile was set to 95 °C for 5 min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 minute, and concluded with a final elongation step at 72 °C for 10 minutes. PCR products were evaluated by 1% agarose gel electrophoresis. Samples were purified with the peqGOLD Cycle-Pure Kit (VWR International, Radnor, Pennsylvania, USA) following the manufacturer's instructions. 3 μL of 27F gene primer were added to 12 μL of purified DNA, then DNA Sanger sequencing was performed by Microsynth AG (Balgach, Switzerland). Sequences of the isolates were blasted against the NCBI RefSeq RNA sequence database to identify them.

As for 16S rRNA gene sequencing, in accordance with previous findings (Janda and Abbott 2007; Kim et al. 2014; Strejcek et al. 2018) 98.65% sequence similarity threshold was accepted to bacterial species demarcation and genus level identification was obtained at 95% sequence similarity (Schloss and Handelsman 2005; Johnson et al. 2019).

4.5.3 DNA-extraction and next-generation sequencing of irrigation water samples

For the isolation of microbial genomic DNA from irrigation water samples DNeasy PowerSoil Pro Kit (QIAGEN, Hilden, Germany) was used. The procedure was performed following the manufacturer's instructions. 250 μL of each sample was added to a dry bead tube with garnet beads from the DNA isolation kit. 800 μL of bead solution was added to the samples to disintegrate the cell walls. The samples were vortexed for 10 minutes. After centrifugation at 15000x *g* for 1 minute, the supernatant was transferred into a 2 mL collection tube. 200 μL of the respective solution of the isolation kit were added, and it was vortexed for 5 seconds to precipitate non-DNA organic and inorganic material. The tubes were centrifuged at 15000x *g* for 1 minute, and the supernatant was transferred into a clean 2 mL microcentrifuge tube. It was followed by adding 600 μL of a high-concentrate salt solution then 5 s of vortexing was performed, after that 650 μL of the lysate was loaded to an MB spin column and centrifuged for 1 minute at 15 000x *g*. Flow-through was discarded, and the column was centrifuged again. The MB spin column was transferred to a clean collection tube,

and 500 µL of a wash solution was added, followed by a centrifugation step. This step was repeated with 500 µL of an ethanol-based wash solution to further clean the DNA and allowing it to stay bound to the silica membrane. Flow-throughs were discarded after each centrifugation. Centrifugation for 2 minutes at 15000x *g* ensured the absence of remaining washing solutions. The column was placed into a new 1.5 mL elution tube. 50 µL of elution buffer were placed on the column, and DNA was eluted via centrifugation for 1 minute. Amplicon library generation, quality control and sequencing were performed at the Vienna Biocenter Core Facilities NGS Unit (www.vbcf.ac.at). The V3–V5 hypervariable region of the 16S rRNA gene was amplified and sequenced using a MiSeq Illumina platform with a 300 bp paired-end read protocol (Illumina, Inc., San Diego, California, USA). The PCR reactions were performed as described in Klindworth et al. (2013) using the forward primer 341f 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and the reverse primer 785r 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG. Raw sequence data is available in the European Nucleotide Archive under accession number PRJEB56665.

4.6 Data analysis

4.6.1 Analysis of data obtained from MALDI-TOF MS identification

The identification process was done using MALDI Biotyper (Bruker Daltonics GmbH & Co, Billerica, Massachusetts, USA). MALDI-TOF MS spectra of the samples were collected using a Microflex LT/SH (Bruker Daltonics GmbH & Co, Bremen, Germany) mass spectrometer equipped with a nitrogen laser (lambda =337 nm) at a laser frequency of 60 Hz operating in linear positive ion detection mode under MALDI Biotyper 3.0 Realtime Classification (RTC) (Bruker Daltonics GmbH & Co, Bremen, Germany) and FlexControl 3.4 (Bruker Daltonics GmbH & Co, Bremen, Germany). Mass spectra were acquired in the range of 2000-21000 Da for each sample analyzed for species level microbial identification which includes the measured ribosomal proteins, which forms up to 70% of a microbial cell in that range. MALDI-TOF MS spectra were generated from 240 single spectra that were created in 40-laser-shot steps from random positions of each isolate. The system was calibrated using *E. coli* ribosomal protein standard (Bruker IVD Bacterial Test Standard, Bruker Daltonics GmbH & Co, Bremen, Germany). FlexControl and FlexAnalysis (Bruker Daltonics GmbH & Co, Bremen, Germany) were used for data acquisition and data processing, respectively.

4.6.2 Analysis of data obtained from biomarker detection experiments

The identification process was done using MALDI Biotyper (Bruker Daltonics GmbH & Co, Billerica, Massachusetts, USA). MALDI-TOF MS spectra of the samples were collected using a Microflex LT/SH (Bruker Daltonics GmbH & Co, Bremen, Germany) mass spectrometer equipped with a nitrogen laser (lambda =337 nm) at a laser frequency of 60 Hz operating in linear positive ion detection mode under MALDI Biotyper 3.0 Realtime classification (RTC) (Bruker Daltonics GmbH & Co, Bremen, Germany) and FlexControl 3.4 (Bruker Daltonics GmbH & Co, Bremen, Germany). Mass spectra were acquired in the range of 2000-21000 Da for each sample analyzed for species level microbial identification which includes the measured ribosomal proteins, which forms up to 70% of a microbial cell in that range. MALDI-TOF MS spectra were generated from 240 single spectra that were created in 40-laser-shot steps from random positions of each isolate. The system was calibrated using *E. coli* ribosomal protein standard (Bruker IVD Bacterial Test Standard, Bruker Daltonics GmbH & Co, Bremen, Germany). FlexControl and FlexAnalysis (Bruker Daltonics GmbH & Co, Bremen, Germany) were used for data acquisition and data processing, respectively. Each raw spectrum was converted to a .csv file by the software and a list of intensities with the corresponding m/z data was created. To differentiate the two *E. coli* isolates and *S. aureus* isolates, Mass-Up was used to analyze further the bacterial mass spectra (López-Fernández et al. 2015). Peak matching, to generate consensus spectra for both *E. coli* (DSM 11250 and ATCC 13706) and *S. aureus* as *S. aureus* (ATCC 25923 and ATCC 43300) strains using forward intra- and inter-sample matching criteria, were performed by Mass-Up (tolerance value for assuming two peaks of the same set at 500 ppm). Applying this method, the previously measured 60 spectra of the two *E. coli* strains (DSM 11250 and ATCC 13706) and 80 spectra of *S. aureus* (ATCC 25923 and ATCC 43300) were reduced to six and eight consensus spectra to implement principal component analysis (PCA). PCA analysis performed by Mass-Up, was made on the peaks of consensus spectra. PCA measured the similarity between objects on a three-dimensional plot as the variance was set at 0.95.

Differentiation of mass spectra of the two *E. coli* strains (DSM 11250 and ATCC 13706) and *S. aureus* strains (ATCC 25923 and ATCC 43300) cultivated on different culture media was made by Discriminant analysis (DA) using IBM SPSS Statistics 27. The previously measured spectra (60 and 80, respectively) were used for DA cultivated on the different types of media. One-way ANOVA, using using IBM SPSS Statistics

27, was applied to compare the score values of identification of isolates cultivated on the different culture media (IBM Corp. 2020).

4.6.3 Analysis of data obtained from culture media experiments

The identification process was done using MALDI Biotyper (Bruker Daltonics GmbH & Co, Billerica, Massachusetts, USA). MALDI-TOF MS spectra of the samples were collected using a Microflex LT/SH (Bruker Daltonics GmbH & Co, Bremen, Germany) mass spectrometer equipped with a nitrogen laser (lambda =337 nm) at a laser frequency of 60 Hz operating in linear positive ion detection mode under MALDI Biotyper 3.0 Realtime classification (RTC) (Bruker Daltonics GmbH & Co, Bremen, Germany) and FlexControl 3.4 (Bruker Daltonics GmbH & Co, Bremen, Germany). Mass spectra were acquired in the range of 2000-21000 Da for each sample analyzed for species level microbial identification which includes the measured ribosomal proteins, which forms up to 70% of a microbial cell in that range. MALDI-TOF MS spectra were generated from 240 single spectra that were created in 40-laser-shot steps from random positions of each isolate. The system was calibrated using *E. coli* ribosomal protein standard (Bruker IVD Bacterial Test Standard, Bruker Daltonics GmbH & Co, Bremen, Germany). FlexControl and FlexAnalysis (Bruker Daltonics GmbH & Co, Bremen, Germany) were used for data acquisition and data processing, respectively.

One-way ANOVA was applied to compare the identification score values of *S. aureus* isolates cultivated on the four culture media (IBM SPSS Statistics 27, Armonk, New York, USA). Based on the values of Skewness and Kurtosis, model residuals had normal distribution. Based on Levene's, homogeneity of variance was violated ($p<0.001$). ANOVA was significant ($F=22.164$; $p<0.001$), therefore Games-Howell test (Post hoc) was used because of the error variances was violated (IBM Corp. 2020). One-way ANOVA was also applied on data obtained from *E. coli* culture media experiments and showed that there was no significant difference between the log score values of the identification ($p>0.05$) (IBM SPSS Statistics 27, Armonk, New York, USA). Kolmogorov–Smirnov test ($p>0.05$) proved that model residuals had normal distribution and homogeneity of variance was checked by Levene's test ($p>0.05$) (IBM Corp. 2020).

4.6.4 Data analysis of comparing MALDI-TOF MS and 16S rRNA gene sequencing

Paired t-test was used to compare the efficacy of identification of the MALDI-TOF MS and 16S rRNA gene sequencing (IBM SPSS Statistics 27, Armonk, New York, U.S.) (IBM Corp 2020).

4.6.5 Bioinformatics and sequence analysis of next-generation sequencing data

Primers were removed from the raw sequences using cutadapt v2.1 (Martin 2011). Raw sequences were further processed with the dada2 v1.14.1 pipeline in R v3.6.3 (Callahan et al. 2016; R Core Team 2021). Briefly, low quality sequences were filtered using 'filterAndTrim' with a maximum number of expected errors of 2 and trimming set at a length where the quality score dropped below 30. After learning the error rates with the 'learnErrors' command, samples were dereplicated using 'derepFastq' and the dada2 sample inference algorithm was run with default parameters. Then, forward and reverse reads were merged with the 'mergePairs' command, choosing a minOverlap = 10 and a maxMismatch = 1. ASV tables were constructed with the 'makeSequenceTable' command. Chimeric sequences were removed using the 'removeBimeraDenovo' command with the consensus method. Taxonomic assignment was performed via the SILVA rRNA database SSU 138 using the 'assignTaxonomy' command (Quast et al. 2013).

MicrobiomeAnalyst was used to analyze data derived from 16S rRNA amplicon sequencing (Dhariwal et al. 2017; Chong et al. 2020). A total of 33 low abundance ASVs were removed based on low prevalence (set at 20%) and low count (<4). After data filtering step, 730 ASVs were used for further analysis and included in the results. Data were normalized by total sum scaling (TSS), i.e. the number of reads from the same ASV were divided by the total number of reads in each sample. Hierarchical Clustering and Heatmap visualization were based on Euclidean distance with the application of Ward clustering algorithm.

5 RESULTS AND DISCUSSION

5.1 Identifying bacteria from environmental samples using MALDI-TOF MS

MALDI-TOF MS was used to identify bacteria from different types of environmental samples. Altogether, 21 samples were used for bacterial isolation. The samples contained a higher number of Gram-negative isolates of which only 22.3% were categorized as unidentified. Moreover, Biotyper could not identify 27.9% of Gram-positive bacteria, whereas only 24.11% of the total isolates were not identified (Table 2). However, in general, taking into consideration the results of both Gram-positive and -negative environmental isolates, better identification results were obtained both at genus (77.2%) and species level (36.6%). Detailed information about the results of the 311 isolates can be found in Appendix Table 1.

Table 2 Identification result of bacteria isolated from different environmental samples

Organisms	MALDI-TOF MS identification scores			
	Isolates	Species identification ≥2	Genus identification ≥1.7	Not reliable identification <1.7
Gram-positive bacteria	86	22 (25.6%)	62 (72.1%)	24 (27.9%)
Gram-negative bacteria	225	92 (40.9%)	178 (79.1%)	47 (20.9%)
Total	311	114 (36.6%)	240 (77.2%)	71 (22.8%)

These results are in concordance with Strejcek et al. (2018) who used MALDI-TOF MS to identify microbes found in soils and sediments and obtained concordant genus level identification (92%) while at species level 35% of the isolates identified coincided with those identified by 16S rRNA gene sequencing analysis. Kopcakova et al. (2014) used MALDI-TOF MS to identify the microflora from waste disposal sites with an overall identification rate lower than 20% at species level which result is lower compared to my identifications. It has been reported by Avanzi et al. (2017) that MALDI-TOF MS was able to identify specific copper resistant microorganisms from soil and water with an identification result of 97% at genus level. Furthermore, El-

Nemr et al. (2019) used MALDI-TOF MS to identify bacteria isolated from a market area (e.g., vegetables, soil, air and hand palms of fresh produce handlers) at species level (41%). This result is close to the results presented in this thesis as 36.6% of the isolates were identified at species level. In another study, Pandey et. al (2019) identified psychrotolerant bacteria isolated from high altitude soil with only 4.92% and 59% of the isolates identified similarly by MALDI-TOF MS and 16S rRNA gene sequencing at species and genus level.

Suzuki et al. (2018) identified waterborne coliform bacteria from sewage, river water and groundwater, obtaining identical results at genus level in 96%, 74%, and 62% of the isolates respectively applying MALDI-TOF MS. These results are also similar to the ones presented in this thesis as 77.2% of the isolates were identified at genus level.

Topić Popović et al. (2021) identified 184 of 321 (57%) Gram-negative bacteria isolated from water and fish samples at species level. Those result are higher than the ones presented in this thesis regarding species identifications of Gram-negative isolates (40.9%), however it should be noted that this thesis includes not only waterborne isolates but isolates from the food chain as well from different matrices (water, vegetables, manure) representing higher biodiversity. Moreover, comparing the Gram-negative bacterial isolates from running water to the aforementioned study of Topić Popović et al. (2022), only 25.3% of the isolates were identified at species level, however better results were obtained at genus level (85.3%). Besides, in the aforementioned study of Pandey et. al (2019) 19.67% of the isolates were not identified at any level by MALDI-TOF MS which value is concordant with the result obtained in this thesis as 22.8% of the isolates remained unidentified. One fact which could have contributed to this phenomenon in the case of the aforementioned study is that at the time of its conduction some of the unidentified isolates (*Bacillus wiedmannii*, *Bacillus velezensis*, *Bacillus paramycoides*) were not included in the database. Moreover, genus *Bacillus* is also one of the most abundantly identified genera in this thesis.

It should be noted that the above-mentioned studies in this paragraph include significantly lower number of isolates (Avanzi et al. (2017) 88, El-nemr et al. (2019) 105, Kopcakova et al. (2014) 51, Pandey et al (2019) 61; Strejcek et al. (2018) 49, Suzuki et al. (2018) 100) which can contribute to the fact that in some cases higher values in species level identification were achieved. It should be also taken into consideration that this thesis includes 311 bacterial isolates from 37 different bacterial

genera therefore it gives a more comprehensive picture about the identification performance of Bruker's Biotyper in environmental microbiology.

The application of MALDI-TOF mass spectrometry is gaining space for identification of food- and waterborne pathogens in the complex food production chain due to its faster and inexpensive identification process compared to traditional or molecular methods. However, as it can be seen in this subchapter, the identification score of environmental isolates is lower compared to studies involving clinical isolates (Ferriera et al. 2010; Van Veen et al. 2010; Ponderand et al. 2020; Chung et al. 2021). Low identification rate can be explained by several factors. As Bruker's database is mostly made for clinically relevant microbes, environmental isolates regarding food safety and quality are underrepresented in it. This finding is similar to previously reported by De Koster and Brul (2016), Strejcek et al. (2018) and El-Nemr et al. (2019). Moreover, several species of genus *Bacillus*, one of the most abundantly occurring Gram-positive genus in my thesis, such as *Bacillus drentensi*, *Bacillus pumilus*, *Bacillus thuringiensis* have either been reported as missing from database, misidentified or identified with low confidence by other authors (Ashfaq et al. 2022). Another factor which could contribute to the lower identification scores is the quality of the peaks. It has been suggested by Rahi et al. (2016) that a good quality spectrum should comprise at least 70-80 peaks for bacteria as those liberate proteins easier than fungi, for that reason spectra for fungi should have around 30-40 peaks for appropriate identification. Moreover, lack of reference spectra or the inability to differentiate closely related species can also lead to misidentifications or not reliable identifications (Bizzini et al. 2011; Seng et al. 2013).

One way to tackle this obstacle is to create in-house databases. It has been widely reported in the literature that wide range of microorganisms that underrepresented in the database were added to it. To name a few, Alispahic et al. (2009) added several *Campylobacter*, *Arcobacter* and *Helicobacter* strains to the database. 56 species from the Rhizobiaceae family, 30 strains from the genus *Prevotella*, 57 species from the genus *Staphylococcus*, 18 strains from the genus *Bradyrhizobium* and 30 strains from the genus *Brachyspira* were also added to each respective database to test the efficacy of the system (Ferreira et al. 2011; Wybo et al. 2012; Sánchez-Juanes et al. 2013; Murugaiyan et al. 2014; Warneke et al. 2014; Shih et al. 2015). However, creating such improved databases is a huge effort which should be taken into consideration.

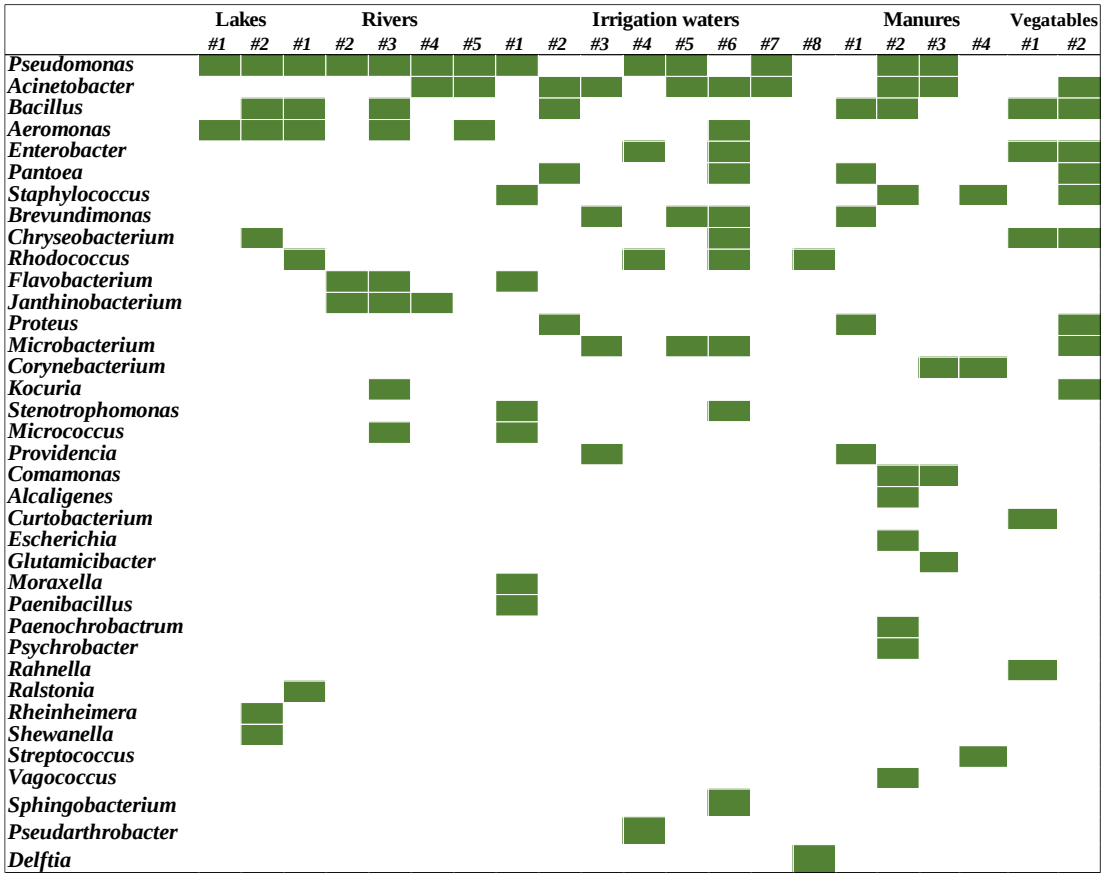
Nonetheless, the data obtained from the measurements presented here and the previously mentioned different environmental studies also suggest that MALDI-TOF MS Biotyper can be a reliable and fast tool to identify bacteria from the environment at genus level. However, as it can be seen in the data presented in this thesis and also in the aforementioned studies, to achieve the same feat at species level, the database needs more improvement by adding more species originated from environment.

Water samples from lakes comprise six bacterial genera as five of them were Gram-negative and one was Gram-positive. Regarding cultivated genera from water samples from lakes, huge differences can be observed as only two genera, *Pseudomonas* and *Aeromonas*, were cultivated from Lake1. However, the cultivated genera from Lake2 showed a bigger variety as not only genera of *Pseudomonas* and *Aeromonas* but *Bacillus*, *Chryseobacterium*, *Rheinheimera*, *Shewanella* were cultivated (Table 3).

River water samples include isolates belonging to 10 different bacterial genera of which four were Gram-positive while six were Gram-negative. This group of samples is the only one in which the ratio of Gram-positive and -negative bacteria are closer to 1:1. Cultivated genera of river water samples showed differences, however genus *Pseudomonas*, the most abundantly cultivated genus in this study, were the only genus which could be detected in every river water sample. The bacterial composition of River1 was similar to River3 as isolates of *Pseudomonas*, *Bacillus* and *Aeromonas* were cultivated from both samples.

Interestingly, the samples of River1 and River2 are from the same river, but 38 km away and from each other, the detected bacterial genera of those are not so similar as only *Pseudomonas* were found in both samples. However, the bacterial composition of River3 showed similarities to both River1 and River2 as *Pseudomonas* could be found in all three samples. Moreover, isolates of *Bacillus* and *Aeromonas* were found in both River1 and River3 while isolates of *Flavobacterium* and *Janthinobacterium* were detected in River2 and River3.

Table 3 Bacterial isolates identified by MALDI-TOF MS from environmental samples. Green coloring indicates bacterial genera, within more species were identified from the samples



River4 is similar to River5 as isolates of both *Pseudomonas* and *Acinetobacter* were cultivated, the only difference is that River4 contained *Janthinobacterium* isolates whereas in River5 isolates of *Aeromonas* could be detected. River5 is flown from River4 which might explain the similar bacterial genera. River3 contained the most different bacterial genera as isolates of seven bacterial genera (*Pseudomonas*, *Bacillus*, *Aeromonas*, *Flavobacterium*, *Janthinobacterium*, *Kocuria*, *Micrococcus*) were cultivated from it as isolates of *Kocuria* and *Micrococcus* only detected in that sample.

Among the samples the most diverse group, regarding cultivated bacterial genera, was the irrigation water samples. Those samples contained 21 bacterial genera of the 37 cultivated genera. The samples comprise more Gram-negative (14) than Gram-positive (7) bacterial genera. In this category, the most abundantly cultivated genus was *Acinetobacter* as isolates of this genus were detected in 5 of 8 samples (Irrigationwater2, Irrigationwater3, Irrigationwater5-7). However, the samples had diverse microbial composition as 14 of the 21 bacterial genera identified from

irrigation water samples were only detected in one sample (e.g., *Flavobacterium, Micrococcus, Moraxella*, *Paenibacillus, Staphylococcus* – Irrigationwater1; *Bacillus, Pantoea, Proteus* – Irrigationwater2; *Providencia* – Irrigationwater3; *Pseudarthrobacter* – Irrigationwater4; *Aeromonas, Chryseobacterium, Sphingobacterium* – Irrigationwater6; *Delftia* -Irrigationwater8). The most diverse sample among irrigation waters was Irrigationwater6 as it contained 10 different bacterial genera (*Acinetobacter, Aeromonas*, *Brevundimonas*, *Chryseobacterium, Enterobacter*, *Microbacterium*, *Pantoea*, *Rhodococcus, Sphingobacterium, Stenotrophomonas*).

Manure samples include isolates of 17 bacterial genera of which only six were Gram-positive while 11 were Gram-negative. The composition of Manure2 was the most diverse as it contained 10 bacterial genera (*Acinetobacter, Alcaligenes*, *Bacillus*, *Comamonas*, *Escherichia*, *Paenochrobactrum*, *Pseudomonas*, *Psychrobacter*, *Staphylococcus*, *Vagococcus*). Nine of 17 genera were only detected in one sample (*Brevundimonas, Pantoea*, *Proteus*, *Providencia* – Manure1; *Escherichia*, *Paenochrobactrum*, *Psychrobacter, Vagococcus* – Manure2; *Streptococcus* – Manure4).

Bacteria isolated from vegetables also showed a wide range of variety as only three of 11 (*Bacillus, Chryseobacterium* and *Enterobacter*) identified genera were detected in both groups of vegetables. Five of 11 (45.45%) identified genera were Gram-positive which is a higher proportion than in any other samples analyzed (lakes, 14.3%; irrigation waters, 33.3%; manures, 35.3%; rivers, 40%). Group of Vegetables2 were more diverse compared to Vegetables1 as the former contained nine bacterial genera while latter comprised only five.

Species of the genus *Pseudomonas* were cultivated from different environments such as lakes (Lake1,2) and rivers (River1-5), irrigation waters (Irrigation water4,5,7) and manures (Manure2,3) as well. Moreover, 70 (22.5%) of the 311 bacteria isolated from environment belonged to genus *Pseudomonas* (Table 4). Overall, the system was able to identify 54 (77.1%) of 70 *Pseudomonas* isolates at genus level and 22 (31.4%) at species level, while only 16 (22.8%) isolates were not identified at any level (Table 4).

Table 4 Identification performance of MALDI-TOF MS regarding Gram-negative bacterial isolates

MALDI-TOF MS identification results of Gram-negative bacterial isolates

Bacterial genus	# of isolate	Species level ≥2	Genus level ≥1.7	No ID <1.7
Pseudomonas	70	22 (31.4%)	54 (77.1%)	16 (22.8%)
Acinetobacter	38	22 (57.9%)	34 (89.5%)	4 (10.5%)
Aeromonas	26	9 (34.6%)	22 (84.6%)	4 (15.4%)
Comamonas	13	2 (15.4%)	6 (46.2%)	7 (53.8%)
Escherichia	11	9 (81.8%)	9 (81.8%)	2 (18.2%)
Brevundimonas	8	5 (62.5%)	7 (87.5%)	1 (12.5%)
Flavobacterium	8	3 (37.5%)	6 (75%)	2 (25%)
Enterobacter	8	4 (50%)	8 (100%)	0 (0%)
Stenotrophomonas	6	1 (16.7%)	2 (33.3%)	4 (66.7%)
Pantoea	5	3 (60%)	5 (100%)	0 (0%)
Chryseobacterium	5	3 (60%)	4 (80%)	1 (20%)
Janthinobacterium	7	0 (0%)	4 (60%)	3 (40%)
Proteus	3	3 (100%)	3 (100%)	0 (100%)
Providencia	2	2 (100%)	2 (100%)	0 (0%)
Alcaligenes	2	1 (50%)	2 (100%)	0 (0%)
Rahnella	2	1 (50%)	2 (100%)	0 (0%)
Shewanella	2	1 (50%)	1 (50%)	1 (50%)
Sphingobacterium	2	0 (0%)	0 (0%)	2 (100%)
Delftia	2	1 (50%)	2 (100%)	0 (0%)
Moraxella	1	0 (0%)	1 (100%)	0 (0%)
Paenochrobactrum	1	0 (0%)	1 (100%)	0 (0%)
Psychrobacter	1	0 (0%)	1 (100%)	0 (0%)
Ralstonia	1	0 (0%)	1 (100%)	0 (0%)
Rheinheimera	1	0 (0%)	1 (100%)	0 (0%)
Total	225	92 (40.9%)	178 (79.1%)	47 (20.9%)

The most abundant *Pseudomonas* species was *Pseudomonas extremorientalis*, first isolated from a drinking water reservoir near Vladivostok City, Russia (Ivanova et al. 2002), as four isolates of 22 high-confidence *Pseudomonas* isolates were identified as *P. extremorientalis*. *Pseudomonas veronii*, one of the nearest phylogenetic relatives of *P. extremorientalis*, was also isolated as two isolates were identified as such with high confidence. Two isolates were identified as members of the *Pseudomonas gessardii* Subgroup, *Pseudomonas brenneri* and *Pseudomonas proteolytica*. Members of the

Pseudomonas mandelli Subgroup were also identified as one isolate was identified as *Pseudomonas mandelli* and one as *Pseudomonas frederiksbergensis*. Other *Pseudomonas* species isolated and identified at species level include *Pseudomonas stutzeri*, an opportunistic human pathogen, *Pseudomonas alcaligenes* and *Pseudomonas antarctica*. The dominance of the *P. fluorescens* Subgroup can be observed as species belonging to that, *P. fluorescens*, *P. antarctica*, *P. extremorientalis*, *P. marginalis* and *P. veronii*, were also identified.

Genus *Acinetobacter* was the second most abundant one as species of it were identified from 10 of 21 samples (Table 4). Species of the genus *Acinetobacter* were cultivated from all types of samples such as rivers, irrigation waters, manures and vegetables (River4,5; Irrigation water2,3, 5-7; Manure2,3; Vegetables2) except the ones from the lakes. 38 isolates were identified of which 22 (57.9%) were identified at species level. As a species of genus *Acinetobacter* with the most abundant one being *Acinetobacter junii* with 15 isolates. This species has been previously reported to be found in aquatic environments such as wastewater (Weidmann-Al-Ahmad et al. 1994), sewage (Guardabassi et al. 1999), surface water (Goswami et al. 2015) and marine sediments (Roseline and Sachindra 2016). Other species such as *Acinetobacter pittii* was identified three times, while both *Acinetobacter ursingii* and *Acinetobacter johnsonii* were identified two times.

The third most abundant genus was found to be genus *Bacillus* as at least one species of that were isolated and identified from 8 of 21 samples (Table 5). This genus is also widely distributed in nature as species of it were cultivated from lakes (Lake2), rivers (River1, 3), irrigation waters (Irrigationwater2), manures (Manure1, 2) and vegetables (Vegetables1, 2) (Table 3). This highlights the fact the genus *Bacillus* is a ubiquitous one as this was the only genus among cultivated genera that was found in all of the environmental samples. Seven of 36 (19.4%) *Bacillus* isolates were identified at species level while also seven isolates were not identified reliably (Table 5). Majority (80.6%) of the *Bacillus* isolates were identified at genus level. Four isolates were identified as *B. cereus*, a common toxin producer pathogen responsible for severe foodborne illnesses. The distribution of this species was quite different as two of isolates were found in River3, Vegetables1 and 2. *Bacillus pumilus* was also isolated and identified from two different types of samples such as River3 and Manure2. Another species, *Bacillus megaterium*, was isolated from Vegetables1. These species are recently found to be useful for plant growth as the former improves tomato growth

and nitrogen uptake under nitrogen fertilization (Masood et al. 2020), while the latter besides having improved the bioavailability of soil phosphorus and potassium, also increased the yields of cucumber (Zhao et al. 2021).

Table 5 Identification performance of MALDI-TOF MS regarding Gram-positive bacterial isolates

MALDI-TOF MS identification results of Gram-positive bacterial isolates

Bacterial genus	# of isolate	Species level ≥2	Genus level ≥1.7	No ID <1.7
Bacillus	36	7 (18.9%)	29 (79.4%)	7 (21.6%)
Staphylococcus	9	3 (33.3%)	6 (66.7%)	3 (22.2%)
Pseudarthrobacter	6	0 (0%)	1 (16.7%)	5 (83.3%)
Glutamicibacter	5	2 (40%)	5 (100%)	0 (0%)
Rhodococcus	5	1 (20%)	4 (80%)	1 (20%)
Corynebacterium	4	1 (25%)	3 (75%)	1 (25%)
Microbacterium	4	2 (50%)	3 (75%)	1 (25%)
Micrococcus	4	0 (0%)	3 (75%)	1 (25%)
Arthrobacter	2	0 (0%)	0 (0%)	2 (100%)
Curtobacterium	2	1 (50%)	2 (100%)	0 (0%)
Kocuria	2	2 (100%)	2 (100%)	0 (0%)
Lactobacillus	2	0 (0%)	0 (0%)	2 (100%)
Paenibacillus	1	1 (50%)	1 (100%)	0 (0%)
Streptococcus	1	1 (100%)	1 (100%)	0 (0%)
Total	86	22 (25.6%)	62 (72.1%)	24 (27.9%)

The genus *Aeromonas* is usually found in aquatic environments with being isolated from rivers, lakes, ponds, seawater (estuaries), drinking water, groundwater, wastewater, and sewage in various stages of treatment. Therefore, it was found in every type of water samples e.g., lakes (Lake1, 2), rivers (River1, 3, 5) and irrigation waters (Irrigation water6) (Table 3). This genus was the fourth most abundant genus as it was isolated from 6 of 21 samples. Nine isolates of 26 (34.6%) were identified at species level, while 22 (84.6%) isolates were identified at genus level and only four (15.4%) isolates were categorized as not reliable identifications (Table 4). Four of the nine isolates, that were identified at species level, were identified as *Aeromonas salmonicida,* an important fish pathogen infecting salmonid populations and other species. Three isolates were identified as *Aeromonas veronii*, a potentially human pathogen that first isolated from victims of drowning or near drowning in fresh water,

from infected wounds of patients previously exposed to fresh water and stools of patients with diarrhea (Hickman-Brenner et al. 1987). The most abundantly isolated and identified species of this genus, *A. salmonicida*, is commonly found in fresh waters whereas *A. veronii*, the second most abundantly cultivated species of this genus in this study, are reportedly rarely found in fresh waters (Janda and Abbott 2010).

The fifth most abundantly cultivated genus was *Comamonas*. However, only two isolates of the 13 were identified at species level. Both of them were identified as *Comamonas jiangduensis*, an aerobe, mesophilic, Gram-negative bacterium, that was first isolated from soil from rice field. Moreover, these species have been isolated from several environments such as termite gut, wetland, activated sludge, soil, hay-infusion filtrate, humans, fresh water, subterranean forest sediment and sulfur spring sediment (Sun et al. 2013). In this study, species of genus *Comamonas* were isolated from solid and liquid manures.

The sixth most abundantly cultivated genus was *Escherichia* with 11 isolates. Nine of 11 isolates were identified as *Escherichia coli*, however as of now Biotyper cannot differentiate this species from other closely related species such *Shigella/Escherichia fergusonii*. Although this genus was the sixth abundant one, species of it have been cultivated from only one sample (Manure2).

Brevundimonas, *Enterobacter*, *Flavobacterium* and *Staphyloccocus* were the seventh most abundantly cultivated ones as eight isolates of each of those genera were cultivated. Genera of *Brevundimonas* and *Flavobacterium* were detected in several samples as species of the former were isolated from irrigation waters (Irrigationwater3, 5, 6) and manures (Manure1) while species of the latter were isolated from rivers (River2, 3) and irrigation waters (Irrigationwater1).

Better results were obtained regarding species of genus *Brevundimonas* as five of eight isolates were identified at species level with four of them being *Brevundimonas vesicularis* and one being *Brevunidmonas diminuta*, the type species of the genus. Both of these species are considered as emerging global opportunistic pathogens. Species of genus *Brevundimonas* have been isolated from various matrices including soils (Kang et al. 2009), different aquatic environments (Abraham et al. 2010) and wide range of clinical specimens (blood, urine, wound culture) (Ryan and Pembroke 2018).

Regarding species of *Flavobacterium* only three isolates were identified at species level with two of them being *Flavobacterium aquatile* and one being *Flavobacterium*

pectinovorum. Eight isolates were identified as a species of the genus *Enterobacter*, four of them were identified at species with three being *Enterobacter cloacae* and one being *Enterobacter hormaechei*. Regarding *Staphylococcus* isolates, three different species were identified such as *Staphyloccocus equorom*, *Staphylococcus saprophyticus* and *Staphylococcus warneri*.

Among the 37 identified bacterial genera, 13 Gram-positive and 26 Gram-negative genera included. Regarding species level identification, although genus *Kocuria* contained only two isolates, the best result by Biotyper was achieved in that case as 100% of those isolates were identified at species level. Taken into consideration Gram-positive genera of which at least five isolates were cultivated and analyzed (*Bacillus*, *Staphylococcus*, *Glutamicibacter*, *Rhodococcus*), the best result was achieved with genus *Glutamicibacter* as two (40%) of five isolates were identified at species level with one as *Glutamicibacter arilaitensis* while the other as *Glutamicibacter protophormiae*. That is followed by *Staphylococcus* (28.6%), *Rhodococcus* (20%) and *Bacillus (*19.4%). However, it should be mentioned that species of genus *Bacillus* were the most dominant by far with 36 isolates, compared to seven isolates of *Staphylococcus* and five-five isolates of both *Glutamicibacter* and *Rhodococcus*. Regarding genus level identification a huge difference was observed with the best result achieved with genus *Glutamicibacter* (100%), followed by *Bacillus* (80.5%), *Rhodococcus* (80%) and *Staphylococcus* (71.4%).

The best results by Biotyper regarding Gram-negative bacteria were achieved in the case of the genus *Escherichia* as 81.8% of those isolates were identified at species level. That is followed by the genus *Brevundimonas* (62.5%), *Pantoea* (60%), *Chryseobacterium* (60%), *Acinetobacter* (58.9%), *Enterobacter* (50%), *Flavobacterium* (37.5%), *Aeromonas* (34.6%), *Pseudomonas* (30.3%), *Stenotrophomonas* (16.7%) and *Comamonas* (12.5%). However, at genus level identification, like in the case of Gram-positive bacteria, a difference was recognized as the best result was obtained with *Enterobacter* (100%) and *Pantoea* (100%) with the former having 8 isolates and the latter having 5. The rest of the result of genus level identifications as follows, *Acinetobacter* (89.5%), *Brevundimonas* (87.5%), *Aeromonas* (84.6%), *Escherichia* (81.8%), *Pseudomonas* (78.8%), *Flavobacterium* (75%), *Comamonas* (37.5%), *Stenotrophomonas* (33.3%).

5.2 Comparing sample preparation techniques of MALDI-TOF MS

62 isolates from different environmental matrices were chosen to test whether an additional formic acid extraction step could enhance the log scores of MALDI-TOF MS identification (Table 6). Therefore, two sample preparation methods, extended direct transfer procedure (on-target extraction) and direct transfer procedure was compared to determine which is the best suited one for environmental isolates. The former involves an additional step of formic acid extraction before adding the matrix while the latter uses only the matrix.

Table 6 Identification results of MALDI-TOF MS applying formic acid extraction

	Results of MALDI-TOF MS identification					
	Extended direct transfer procedure			**Direct transfer procedure**		
Total number of Isolates	Species level id. ≥2	Genus level id. ≥1.7	Not reliable id. 1.7>	Species level id. ≥2	Genus level id. ≥1.7	Not reliable id. 1.7>
62	31 (50%)	60 (96.8%)	2 (3.2%)	21 (33.9%)	46 (73%)	16 (25.4%)

The results of the formic acid extraction are clearly visible as only 2 of the 62 isolates were not identified at any level whereas the unidentified number of isolates lacking formic acid extraction were 16. However, interestingly the aforementioned two isolates (#6; #23) were identified without adding formic acid as a *Pseudomonas* and a *Rheinheimera* isolate (Table 7). 16 isolates were unidentified without adding formic acid of which two were identified at species level and the remaining 14 at genus level with applying formic acid. The two isolates (#30; #33) were identified as a *Comamonas jiangduensis* and a *Pseudomonas antarctica* isolates with applying formic acid extraction. The usefulness of formic acid extraction is clearly visible as 60 (96.8%) of the 62 isolates were identified at genus level whereas without applying formic acid only 46 (73%) were identified. It is also remarkable that 31 (50%) of the 62 isolates were identified at species level with formic acid while without it only 22 (34.9%).

The identification results of two isolates (#31; #48) were not accepted at species level identification because despite reaching the threshold of log score 2, the best match for

the former was a *Psychrobacter* spp. isolate while for the latter, without applying formic, an *Aeromonas* spp. isolate. Therefore, these results were only accepted as genus level identifications. 16 isolates were identified at species level with both sample preparation techniques of which 13 were identified identically. The identification result of three isolates (#47; #50; #62) were different as the first two isolates were identified as *Aeromonas salmonicida* with formic acid while without it those were identified as *Aeromonas bestiarum*. The identification result of #62 isolate was also differed as with applying formic acid it was identified as *Pseudomonas fluorescens* whereas without it, it was identified as *Pseudomonas marginalis*.

The benefit of applying the extended direct transfer procedure is well-marked from the results. It should also be noted the average log scores of the extended direct transfer procedure were above 2 therefore reaching the species level threshold, while the average log scores of the direct transfer procedure was only 1.85. Besides, it was proved by Paired t-test ($t=16.09$, $p<0.001$) that the difference of the average log scores of the two sample preparation methods was significant.

Table 7 MALDI-TOF MS identification results of each isolates comparing extended direct transfer procedure and direct transfer procedure

No.	Isolate	Extended direct transfer procedure (log score)	Direct transfer procedure (log score)
#1	SZE 0/7 R2A	*Bacillus* spp. (1.97)	No ID (1.42)
#2	SZT 2/7 R2A 22	*Janthinobacterium* spp. (1.98)	No ID (1.54)
#3	SZT 2/6 TSA 22	*Pseudomonas* spp. (1.82)	*P. extremorientalis* (2.05)
#4	SZT 2/5 TSA 22	*Flavobacterium* spp. (1.74)	*Flavobacterium* spp. (1.8)
#5	SZT 2/1 TSA 22	*P. brenneri* (2.22)	*P. brenneri* (2.21)
#6	SZT 2/4 TSA 22	No ID (1.68)	*Pseudomonas* spp. (1.72)
#7	BSZ 3/4	*Glutamicibacter* spp. (1.89)	No ID (1.52)
#8	BSZ 3/3	*Pseudomonas* spp. (1.84)	*Pseudomonas* spp. (1.77)
#9	BSZ 3/2	*Glutamicibacter* spp. (1.91)	No ID. (1.36)
#10	BSZ 3/1	*Pseudomonas* spp. (1.9)	*Pseudomonas* spp. (1.87)
#11	CS 2/1 R2A 22	*Pseudomonas* spp. (2)	*Pseudomonas* spp. (1.84)
#12	CS 1/1 TSA 22	*F. aquatile* (2.22)	*Flavobacterium* spp. (1.98)
#13	SZE 0/1 TSA	*Shewanella baltica* (2.03)	*Shewanella* spp. (1.96)
#14	KD 2/4 TSA	*A. johnsonii* (2.16)	*Acinetobacter* spp. (1.98)
#15	KD 2/5 TSA	*A. johnsonii* (2.05)	*A. johnsonii* (2.11)
#16	KD 2/2 TSA	*P. extremorientalis* (2.11)	*P. extremorientalis* (2.12)
#17	KD 2/1 TSA	*P. extremorientalis* (2.27)	*P. extremorientalis* (2.07)
#18	SZE 0/5 TSA A	*Bacillus* spp. (1.92)	*Bacillus* spp. (1.88)
#19	SZE 0/5 TSA B	*Bacillus* spp. (1.91)	No ID (1.63)
#20	BT 1/3 B	*E. coli* (2.17)	*E. coli* (2.19)
#21	BT 1/3 1	*Comamonas* spp. (1.74)	*Comamonas* spp. (1.78)
#22	BT 1/3 2	*E. coli* (2.46)	*E. coli* (2.24)
#23	SZE 0/9 TSA	No ID (1.63)	*Rheinheimera* spp. (1.73)
#24	CS 1/2 TSA 22 A	*Micrococcus* spp. (1.93)	*Micrococcus* spp. (1.89)

#25	CS 1/2 TSA 22 B	*Janthinobacterium* spp. (1.73)	No ID (1.4)
#26	BSZ 3/5/1 H	*P. extremorientalis* (2.18)	*Pseudomonas* spp. (1.86)
#27	BSZ 3/5/2 H	*P. extremorientalis* (2.08)	*Pseudomonas* spp. (1.77)
#28	BT 2/3 5/7 H	*Acinetobacter* spp. (1.75)	*Acinetobacter* spp. (1.72)
#29	BT 2/3 5/8 TSA	*Alcaligenes* spp. (1.94)	*Alcaligenes faecalis* (2.1)
#30	BT 2/3 5/3	*C. jiangduensis* (2.07)	No ID (1.62)
#31	BT 2/3 5/5	*Psychrobacter* spp. (2)	*Psychrobacter* spp. (1.8)
#32	BT 1/3 5/3 TSA	*E. coli* (2.4)	*E. coli* (1.97)
#33	BSZ 3/5/5 TSA	*P. antarctica* (2.05)	No ID (1.65)
#34	BSZ 3/5/6 TSA	*Pseudomonas* spp. (1.75)	No ID (1.64)
#35	BSZ 3/5/2 TSA	*P. extremorientalis* (2.11)	*Pseudomonas* spp. (1.82)
#36	CS 1/2 TSA B 22	*Janthinobacterium* spp. (1.73)	No ID (1.42)
#37	BT 2/3 5/9 TSA	*Comamonas* spp. (1.71)	No ID (1.27)
#38	BT 1/3 5/7	*E. coli* (2.37)	*E. coli* (2.18)
#39	BT 2/3 5/10	*Pseudomonas* spp. (1.82)	No ID (1.5)
#40	BT 2/3 5/9	*C. jiangduensis* (2.14)	*C. jiangduensis* (2.01)
#41	BT 2/3 5/7 TSA	*Paenochrobactrum* spp. (1.86)	*Paenochrobactrum* spp. (1.79)
#42	BT 2/3 5/8	*E. coli* (2.11)	*E. coli* (2.15)
#43	BT 2/3 5/12	*Acinetobacter* spp. (1.82)	No ID (1.32)
#44	BSZ 3/5/3	*G. protophormiae* (2.29)	*Glutamicibacter* spp. (1.96)
#45	BT 3/3 5/1	*Alcaligenes faecalis* (2.33)	*A. faecalis* (2.01)
#46	BSZ 3/5/4	*P. extremorientalis* (2.03)	*Pseudomonas* spp. (1.94)
#47	KAV 0/1 R2A	*A. salmonicida* (2.18)	*A. bestiarum* (2.02)
#48	KAV 0/2 R2A	*A. salmonicida* (2.12)	*Aeromonas* spp. (2.01)
#49	SZT 2/8 B TSA 22	*Pseudomonas fragi* (2.14)	*Pseudomonas* spp. (1.96)
#50	SZE 0/2 TSA	*A. salmonicida* (2.09)	*A, bestiarum* (2.13)
#51	CS 1/1 TSA	*P. extremorientalis* (2.3)	*P. extremorientalis* (2.18)
#52	BT 2/3 5/8	*E. coli* (2.11)	*E. coli* (2.23)
#53	BSZ 3/6 B TSA	*Acinetobacter* spp. (1.77)	No ID. (1.32)
#54	BV 2/2 TSA	*P. marginalis* (2.26)	*P. marginalis* (2.15)
#55	BT 1/3 5/3	*E. coli* (2.51)	*E. coli* (2.38)
#56	T 3/3 TSA A	*Staphylococcus* spp. (1.96)	No ID (1.39)
#57	BV 2/3 TSA	*Acinetobacter* spp. (1.95)	*A. johnsonii* (2.14)
#58	T 3/3 TSA B	*Staphylococcus* spp. (1.96)	No ID (1.58)
#59	SZE 0/2 R2A 22	*Pseudomonas* spp. (1.72)	*Pseudomonas* spp. (1.9)
#60	SZT 2/6 TSA 22	*Pseudomonas* spp. (1.82)	*P. extremorientalis* (2.05)
#61	SZE 1/6 A R2A 22	*Bacillus* spp. (1.78)	*Bacillus* spp. (1.88)
#62	KD 2/4 R2A 22	*P. fluorescens* (2.19)	*P. marginalis* (2.05)
	Total	Average log score (2.01)	Average log score (1.85)

The different sample preparation methods have been tested since the introduction of MALDI-TOF MS systems. Alatoom et al. (2011) tested 305 clinical isolates of staphylococci, streptococci and related genera and found that MALDI-TOF MS Biotyper correctly identified 95% and 69% isolates at genus and species levels using the full extraction method. However, direct colony testing (direct transfer procedure) identified only 56% and 20% of the isolates at genus and species levels, respectively. In another study, Barcelos et al. (2019) generated 87.1% and 89.8% correct species and genus-level identification using mastitis causing bacteria, such as *Staphylococcus*

spp., *Streptococcus* spp., *Enterococcus* spp., *Aerococcus* spp. and *Lactococcus* spp. by comparing the on-plate method (extended direct transfer procedure) to the standard, full extraction protocol. These results are congruent to the results presented in this subchapter, highlighting the necessity of an additional formic acid extraction step before applying matrix.

In addition, Anderson et al. (2012) found that direct spotting or direct transfer procedure gives lower identification scores compared to methods involving extraction. However, direct transfer procedure is able to identify Gram-negative rod-shaped bacteria and also isolates regularly occurring in clinical laboratories but excluding Gram-positive bacteria which require additional extraction to obtain proper results (Alatoom et al. 2011; Tsuchida et al. 2020; Wang et al. 2021).

Nevertheless, the library of Biotyper was developed based on results of the standard/full extraction protocol (Alatoom et al. 2011, Barcelos et al. 2019) which involves approximately 13 steps with dissolving bacterial colonies in 70% ethanol, pelleting, drying the colonies followed by 70% formic acid extraction and 2 steps of centrifugation before spotting onto the plate for identification. Thus, identification by full extraction method can lasts around 2-3 hours for 40-48 samples therefore requiring considerably more time (approximately 45 minutes more) for the same number of isolates than using the direct or extended direct transfer protocol (Dhiman et al. 2011; Matsuda et al. 2012). Therefore, the full extraction method is not suitable for rapid identification and its relatively labor-intensive nature also makes it more difficult to fit well into the workflow of the clinical laboratory (McElvania TeKippe et al. 2013). Moreover, its reagent demand per isolate is also considerably more significant (700 µL formic acid, 900 µL ethanol, 300 µL HPLC grade water) compared to extended direct transfer procedure which only needs 1-1 µL of matrix and formic acid per isolate, respectively.

However, the application of on-target extraction (extended direct transfer procedure) yielded better identification scores compared to direct transfer procedure and full extraction method regarding environmental isolates of *Photobacterium damselae* subsp. *piscicida* (Kazazic et al. 2019). Moreover, McElvania TeKippe et al. (2013) could achieve an improved genus- and species level identification by 20% with the application of on-target extraction compared to direct transfer procedure using aerobic Gram-positive bacteria isolated from clinical specimens. In another study of Haigh et al. (2011), 93% of the previously unidentified clinical isolates by direct transfer

procedure, mostly Gram-positive bacilli, coagulase-negative staphylococci, yeasts, and anaerobes, were identified at least at genus level by formic acid extraction. These findings are similar to those presented in this subchapter of this thesis. However, the aforementioned studies involved mostly clinical isolates, but the improvement of identification results arising from applying formic acid to environmental isolates is also demonstrated here.

5.3 The effect of culture media on MALDI-TOF MS identification regarding waterborne isolates

The effect of different culture media was tested on the identification performance of MALDI-TOF MS; therefore, the results of the validated culture medium (TSA) were compared to the results of R2A and Yeast Extract Agar (YEA). The 23 isolates chosen for these measurements have been isolated previously from different aqueous environment (lakes, rivers). The isolates contain mostly Gram-negative waterborne bacteria from the genera of *Aeromonas, Acinetobacter, Pseudomonas, Shewanella* and Gram-positive genera such as *Bacillus* and *Micrococcus* (Table 8). Results from TSA agar, the validated culture medium for MALDI-TOF MS identifications, shows that 11 of 23 waterborne isolates were identified at species level. In addition, eight isolates were identified only at genus level, whereas only four isolates were categorized as unsuccessful identifications. One of three isolates that were unable to be identified on TSA agar, was identified as a *Pseudomonas* (#17) isolate on YEA. However, the remaining two isolates (#82; #86) were not identified at any level on R2A and YEA, respectively. Two Gram-positive isolates belonging to the genera of *Bacillus* (#99) and *Micrococcus* (#94) were only identified at genus level on TSA agars, while the latter was identified as *Micrococcus luteus* on both R2A and YEA. However, another Gram-positive isolate (#64), identified as *M. luteus* on TSA, was also identified similarly on YEA while on R2A agar only genus level identification was achieved. Out of the 12 *Pseudomonas* isolates, seven were identified at species level on TSA of which only three isolates were identified similarly on all three culture media as two *P. marginalis* isolates (#10; #112) and one *P. extremorientalis* (#32) isolate. Another *Pseudomonas* isolate (#15) was identified as *P. fragi* on TSA, however only low-confidence log scores (genus level) were achieved on both R2A and Yeast Extract Agar while having the same species result. The isolate of #18 was identified only at genus level on both TSA and YEA, while on R2A agar it was identified as a *P. marginalis* isolate. The isolate of #22 was identified as a *P. fragi* isolate on both TSA and R2A while it was

identified only at genus level on YEA. Another *Pseudomonas* isolate was identified as *Pseudomonas chlororaphis* on TSA whereas it was only identified at genus level on YEA, however the best match was also *P. chlororaphis*, but it was not identified on R2A. The isolate of #101 was identified as *Pseudomonas tolaasi* on TSA agar while it was identified only at genus level on YEA. In addition, this isolate was generated a log score of above 2 on R2A agar, therefore achieved the species level identification threshold, however the identification process gave the same result for both *Pseudomonas extremorientalis* and *Pseudomonas cedrina* meaning that this result was only accepted as a *Pseudomonas* isolate without species determination. The isolate of #108 was identified only at genus level as a *Pseudomonas* isolate on all three culture media, however in each case identified as a *Pseudomonas fragi* isolate without achieving the species level threshold. However, the most interesting isolate was #113, because while achieved the score of 2 on each culture medium, the results were contradictory to each other as it was identified as a *P. fragi* on TSA, *P. extremorientalis* on R2A and *P. cedrina* on YEA.

Among three *Aeromonas* isolates, only one (#20) was identified at species level on TSA, as an *Aeromonas eucrenophila* while this isolate was identified as *Aeromonas bestiarum* on both R2A and YEA. The remaining two *Aeromonas* isolates (#23; #24) were identified only at genus level on TSA while the former was identified as *Aeromonas bestiarum* on both R2A and YEA. The isolate of #24 was identified as an *Aeromonas* isolate on TSA, however it was identified as *Aeromonas popoffii* on R2A whereas it was identified as *Aeromonas eucrenophila* on YEA. In total, 12 of 23 isolates were identified at species level on R2A agar which is a higher value than that of identified on TSA agar. However, the number of unidentified isolates is also higher than that of on TSA agars as five of 23 isolates were not identified at any level compared to the four unidentified isolates on TSA. However, two of these unidentified isolates were identified on TSA agars as *Pseudomonas* isolates (#21; #100) while the rest of those remained unidentified. Nevertheless, five isolates which were identified only at genus level (#18; #23; #24; #94) or even not identified (#102) on TSA were identified at species level R2A. Isolate of #18 was identified as *P. marginalis* on R2A whereas it was only identified at genus level as a *Pseudomonas* isolate on both TSA and YEA.

The most interesting isolate was found to be isolate #102 because it was not identified on TSA whereas, it was identified as an *Acinetobacter johnsonii* isolate on both R2A

and YEA. In summary, better identification scores were achieved on R2A agars than on TSA agars in the case of #18; #23; #24; #94; #102, while worse identifications were obtained in the case of #21; #64; #100; #101.

Results on Yeast Extract Agar, the third applied culture medium, were similar to those obtained on TSA and R2A because 11 of the 23 isolates were identified at species level. However, this culture medium was found to be the best due to its lowest number of unidentified isolates, only three isolates (#21; #82; #87) were not identified at any level. These isolates were not identified at any level on three culture media. Nine of the 23 isolates were identified at genus level on YEA. However, four isolates (#23; #24; #94; #102) were identified at species level on YEA which were only identified at genus level on TSA. Nevertheless, five isolates (#15; #21; #22; #100; #101) were identified at species level on TSA agar, whereas these isolates were only identified at genus level on Yeast Extract Agar. All of the aforementioned isolates were different *Pseudomonas* species.

In general, it was proved by ANOVA (F=0.10, p= 0.90, F_{crit}= 3.13) that all three culture media are suitable to identify waterborne bacteria as no significant difference was observed regarding the averages of log scores. However, the average log scores were below 2, the species level identification threshold, in each case (TSA, 1.97; R2A, 1.95; YEA, 1.97). This also highlights the limitation of MALDI-TOF MS Biotyper to identify environmental bacteria at species level therefore this measurement also necessitates the enrichment of the database by a wide range of environmental isolates. It can also be recommended to parallel identify waterborne isolates on TSA and Yeast Extract Agar to achieve better identification at species level. However, even higher identification scores could have been achieved by applying lower cutoff scores from 2 to ≥1.9 as it has been proposed by Seng et al. (2009) and Risch et al. (2010). Moreover, a further decrease to ≥ 1.8 was achieved an 86% increase in species-level identification of clinical isolates (Fedorko et al. 2012). However, more measurements are needed to thoroughly assess at what extent a cutoff score reduction is acceptable for waterborne isolates.

Table 8 Identification results of MALDI-TOF MS obtained on three different culture media

No. of isolate in the collection	Result on TSA (log score)	Result on R2A (log score)	Result on YEA (log score)
#10	*P. marginalis* (2.2)	*P. marginalis* (2.26)	*P. marginalis* (2.15)
#15	*Pseudomonas fragi* (2.07)	*P. fragi* (1.88)	*P. fragi* (1.92)
#17	No id. (1.58)	No id. (1.3)	*P. antarctica* (1.76)
#18	*Pseudomonas grimonti* (1.92)	*P. marginalis* (2.02)	*P. grimonti/marginalis* (1.99)
#20	*Aeromonas eucrenophila* (2.01)	*Aeromonas bestiarum* (2.15)	*A. bestiarum* (2.07)
#21	*P. fragi* (1.86)	No id. (1.65)	No id. (1.57)
#22	*P. fragi* (2.04)	*P. fragi* (2)	*P. fragi* (1.9)
#23	*A. bestiarum* (1.83)	*A. bestiarum* (2.08)	*A. bestiarum* (2.03)
#24	*Aeromonas* sp. (2.09)	*Aeromonas popoffii* (2.17)	*A. eucrenophila* (2.02)
#32	*P. extremorientalis* (2.12)	*P. extremorientalis* (2.27)	*P. extremorientalis* (2.02)
#64	*Micrococcus luteus* (2.19)	*M. luteus* (1.96)	*M. luteus* (2.05)
#82	No id. (1.59)	No id. (1.44)	No id. (1.49)
#86	*A. johnsonii* (2.16)	*A. johnsonii* (2.17)	*A. johnsonii* (2.29)
#87	*Pseudomonas carica* (1.52)	*P. anguilliseptica* (1.61)	*P. anguilliseptica* (1.53)
#91	*Shewanella baltica* (2.03)	*S. baltica* (2.27)	*S. baltica* (2.16)
#94	*M. luteus* (1.93)	*M. luteus* (2.02)	*M. luteus* (2.02)
#99	*B. pumilus* (1.9)	*B. pumilus* (1.72)	*B. pumilus* (1.95)
#100	*Pseudomonas chlororaphis* (2.12)	No id. (1.69)	*P. chlororaphis* (1.95)
#101	*Pseudomonas tolaasii* (2.13)	*P. extremorientalis/cedrina* (2.06)	*P. extremorientalis* (1.99)
#102	No id. (1.65)	*A. johnsonii* (2.12)	*A. johnsonii* (2.13)
#108	*P. fragi* (1.91)	*P. fragi* (1.88)	*P. fragi* (1.91)
#112	*P. marginalis* (2.26)	*P. marginalis* (2.04)	*P. marginalis* (2.08)
#113	*P. fragi* (2.22)	*P. extremorientalis* (2.08)	*P. cedrina* (2.08)
Average log score	**1.97**	**1.95**	**1.97**

5.4 Discriminating bacterial strains while simultaneously testing the effect of culture media on PMF of strains using MALDI-TOF MS

Since its introduction to microbiology, besides basic microbial identification, MALDI-TOF MS has also been used to discriminate antibiotic resistant bacteria (Yoon and

Jeong 2021) such as methicillin-resistant *S. aureus* (Croxatto et al. 2012; Tang et al. 2019), vancomycin-resistant enterococci (Griffin et al. 2012; Nakano et al. 2014) or even determining the efflux activity of *E. coli* strains (Lu et al. 2020).

5.4.1 Discriminating bacterial strains via multivariate statistical methods and proteomics

The ability of MALDI-TOF MS to discriminate bacterial strains was tested and simultaneously the effect of culture media on the PMF of strains was also analyzed. Two *S. aureus* strains were tested as *S. aureus* ATCC 25923, a quality control strain while *S. aureus* ATCC 43300, a methicillin-resistant one. A specific peak was observed at m/z 5868 in the mass spectrum of *S. aureus* ATCC 43300 on all four culture media which peak can be considered as a specific biomarker for this strain as it was present only in the mass spectra of the aforementioned strain. Different specific biomarkers for MSSA and MRSA have also been reported at m/z 2302, m/z 3048, m/z 3086, m/z 3124 and m/z 3871 (Jackson et al. 2005; Drake et al. 2011). However, the mass spectra included in this study did not comprise such peaks in a way that it could help differentiate MRSA from non-resistant *S. aureus* strains. The aforementioned peaks have either been missing from the mass spectra or were observed in both *S. aureus* strains therefore those peaks were not appropriate for the differentiation of the strains.

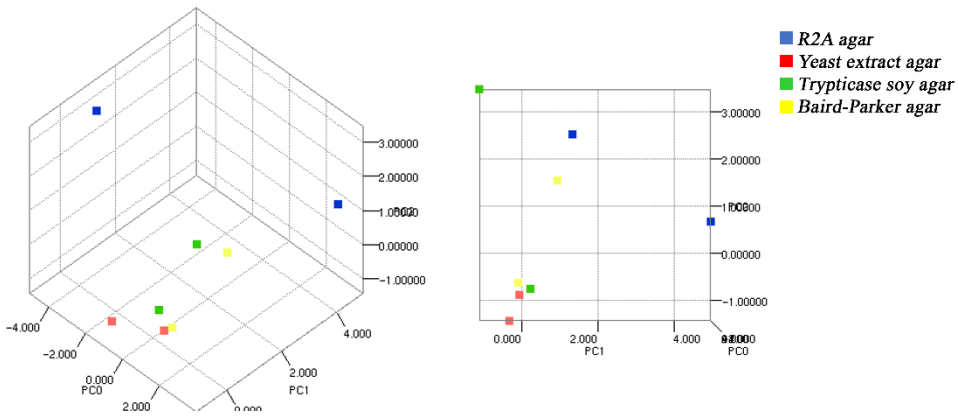


Figure 1 PCA on the level of culture media's type with the mass spectra of *S. aureus* ATCC 25923 and *S. aureus* ATCC 43300. Colors are to distinguish between culture media

Principal component analysis (PCA) was used to differentiate the strains of *S. aureus* ATCC 25923 and *S. aureus* ATCC 43300. The separated eight groups were clearly visible as each group contained mass spectra obtained on them separately per strains (Figure 1). Culture media applied in this study coupled with PCA were all suitable to differentiate *S. aureus* ATCC 43300 from *S. aureus* ATCC 25923 as each group were well separated from each other. However, there is a difference regarding differentiation by culture media as the groups of TSA and YEA for *S. aureus* ATCC 25923 were closer to each other therefore the spectra obtained on them were similar. The effect of those culture media on the mass spectra seemed negligible which was also reflected by the log scores of identifications (Table 9). Yeast Extract agar seemed the least useful for the differentiation of the two strains as those were also close to each other. However, it can be observed that groups of culture media containing the mass spectra of *S. aureus* ATCC 43300 and *S. aureus* ATCC 25923 were separated but R2A agar seemed to be the best one. Because the groups of those were separated regarding each *S. aureus* strain.

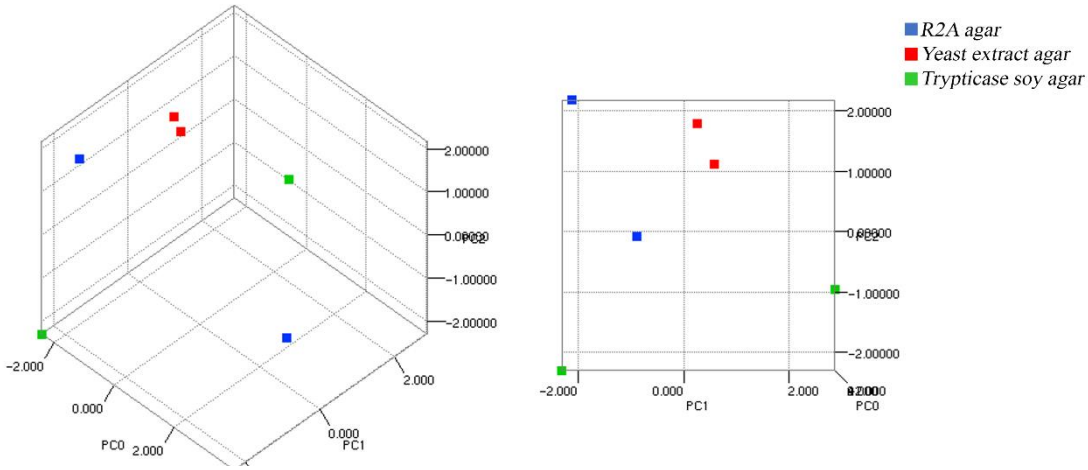


Figure 2 PCA on the level of culture media's type with the mass spectra of *E. coli* ATCC 13706 and *E. coli* DSM 11250. Colors are to distinguish between culture media

PCA was used to differentiate the strains of *E. coli* ATCC 13706 and *E. coli* DSM 11250. The six groups are clearly separated as each group contain the mass spectra obtained on them separately per strains such as *E. coli* ATCC 13706 on R2A agar etc. (Figure 2). Culture media applied in this study coupled with PCA were all suitable to differentiate the two aforementioned strains as each group were well separated from

each other. However, there is a difference regarding differentiation by culture media as the groups of YEA are closer to each other therefore the spectra obtained on them are similar so the effect of that culture media on the mass spectra seemed to be weak. Thus, YEA seemed the least useful for the differentiation of the two strains. However, analyzing the strains it can be noticed that each group of culture media containing the mass spectra of the analyzed strains were separated, especially TSA and R2A. For that reason, the groups of those were well separated regarding each *E. coli* strain.

Santos et al. (2015) found specific peaks which can be considered as *E. coli* specific peaks at m/z 5379 ± 3 and 6253 ± 3, respectively. Feng et al. (2020) also detected these *E. coli* specific peaks at m/z 5381 and 6255. The results presented here are similar to those studies because those *E. coli* specific peaks were also found in the mass spectra obtained with these measurements at m/z 5376 and 6250. Moreover, two strain specific peaks at m/z 6640 and m/z 8912 were also detected in the mass spectra of *E. coli* ATCC 13706. These peaks made it possible to distinguish the two *E. coli* strains used in this study as these were only present in the mass spectra of *E. coli* ATCC 13706 on all three culture media. The occurrence of proteins in different culture conditions can be associated with housekeeping genes as those are expressed constitutively, despite varied culture conditions (Topić Popović et al. 2021). The presence of these peaks in the mass spectra of *E. coli* isolates identified by MALDI-TOF MS can lead to the rapid identification of *E. coli* ATCC 13706, so these can be considered as important biomarkers for this species. Thus, these results show the ability of MALDI-TOF MS to discriminate waterborne microbial strains by applying multivariate statistical methods on the mass spectra of different isolates.

5.4.2 Analyzing the effect of culture media on the PMF of bacterial strains

Discriminant analysis (DA) was used to analyze the effect of culture media on the protein mass fingerprint of *S. aureus* strains. Regarding *S. aureus* ATCC 25923 DA clearly separated Baird-Parker Agar from the rest of the groups which was also visible in the identification results as the lowest scores were achieved on that culture medium (Table 9). Moreover, the groups of all four culture media were distinguished from each other. The groups of TSA, R2A and YEA were closer to each other because those spectra were similar to each other and their identification scores were also closer as well.

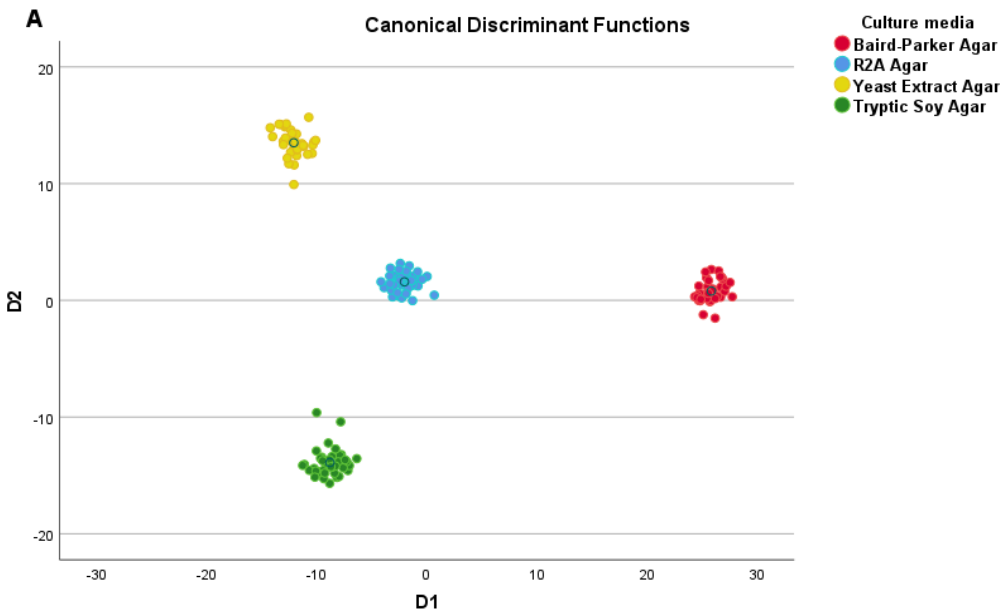


Figure 3 Discriminant analysis (DA) on the level of culture media's type with the mass spectra of *S. aureus* ATCC 25923. As independent variable D1 represents the first canonical discriminant function, the weighted linear function of variables that maximizes the differences between groups. D2 (dependent variable) is the second discriminant function that maximizes the remained differences between groups. Four separate groups are formed regarding the applied culture media. Colors are to distinguish between culture media

Regarding *S. aureus* ATCC 43300, DA showed that the spectra obtained on the two high-nutrient culture media, TSA and YEA, were similar to each other. However, not only the group of Baird-Parker Agar was further away from the two high nutrient culture media groups but the group of R2A, a low-nutrient culture medium, as well.

In the case of *S. aureus* ATCC 25923, 100% successful grouping was obtained in each case except the group of YEA of which 97.1% was achieved. Regarding *S. aureus* ATCC 43300, 100% successful was obtained for each group. Thus, the grouping of mass spectra obtained from different culture media can be considered as successful. Based on the test of equality of group means, significance level of each variable is <0.05, meaning that the group centroids of the three groups were not equal, they were clearly separated for both *S. aureus* strains. Classification was based on Mahalanobis distances to the group centroids and the result was compared to the original classification.

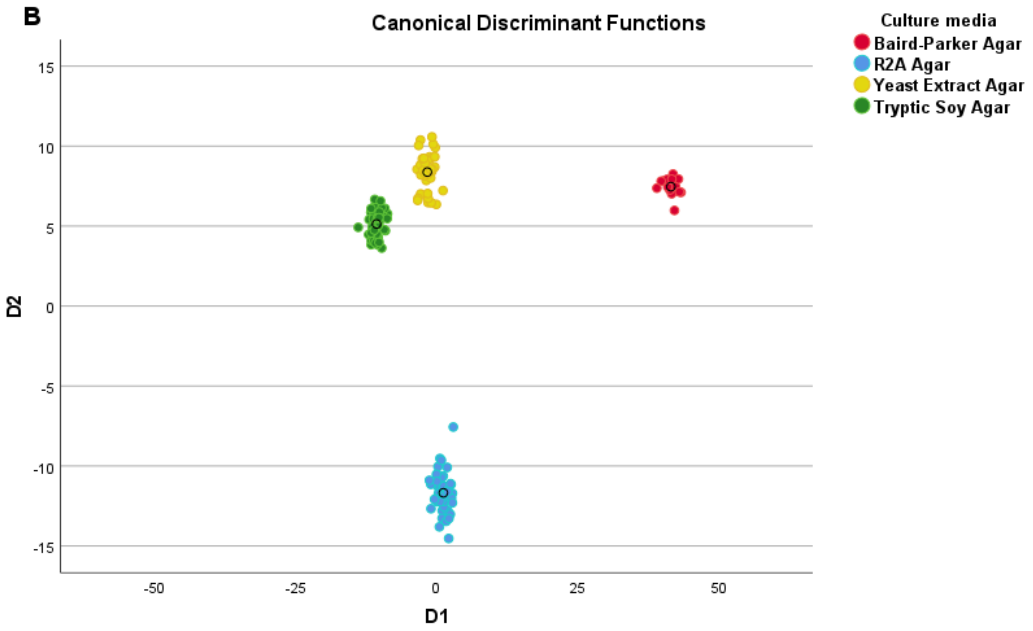


Figure 4 Discriminant analysis (DA) on the level of culture media's type with the mass spectra of *S. aureus* ATCC 43300 with the same settings as detailed in the caption of Figure 3

Interestingly, Baird-Parker Agar, the selective and differential culture medium for the isolation and enumeration of *S. aureus* in foods, environmental and clinical specimens, was the least effective of the tested culture media in terms of the confidence of identification (Table 9). The best culture medium for both *S. aureus* strains was found to be the TSA as the highest log scores were achieved on that medium. The log scores of *S. aureus* ATCC 25923 were higher on Baird-Parker Agar, YEA, but on R2A agar *S. aureus* ATCC 43300 achieved better results. However, it is also important to add that *S. aureus* ATCC 25923 can be found in the Bruker's database which fact could explain the higher log scores of identifications. Given the fact that the other *S. aureus* strain (ATCC 43300) is not part of the database as an added reference strain this also highlights that the database enhancement is inevitable thus better results could be achieved.

Table 9 Identification results of MALDI-TOF MS obtained on different culture media regarding *S. aureus* isolates

	Baird-Parker Agar		R2A Agar		Yeast Extract Agar		Tryptic Soy Agar	
Bacterial strains	*S. aureus* ATCC 25923	*S. aureus* ATCC 43300	*S. aureus* ATCC 25923	*S. aureus* ATCC 43300	*S. aureus* ATCC 25923	*S. aureus* ATCC 43300	*S. aureus* ATCC 25923	*S. aureus* ATCC 43300
Identification score values	1.92	1.81	2.11	2.14	2.15	2.07	2.16	2.16

One-way ANOVA was used to compare the identification score values of *S. aureus* isolates cultivated on the four culture media. ANOVA was significant ($F=22.164$; $p<0.001$), therefore Games-Howell test (Post hoc) was used as the error variances were violated. Games-Howell post hoc test showed that the averages of spectra obtained on Baird-Parker Agar were significantly different from the TSA agar ($p<0.001$). Thus, demonstrating that TSA agar can also be a good fit to identify *S. aureus* isolates as the highest identification score values were obtained on that culture medium. The averages of spectra obtained on R2A and Yeast Extract Agar were not differed significantly from the average of TSA agar. Therefore, these measurements showed that R2A and YEA agars are also suitable to identify *S. aureus* by MALDI-TOF MS as the application of both culture media generated species level identifications.

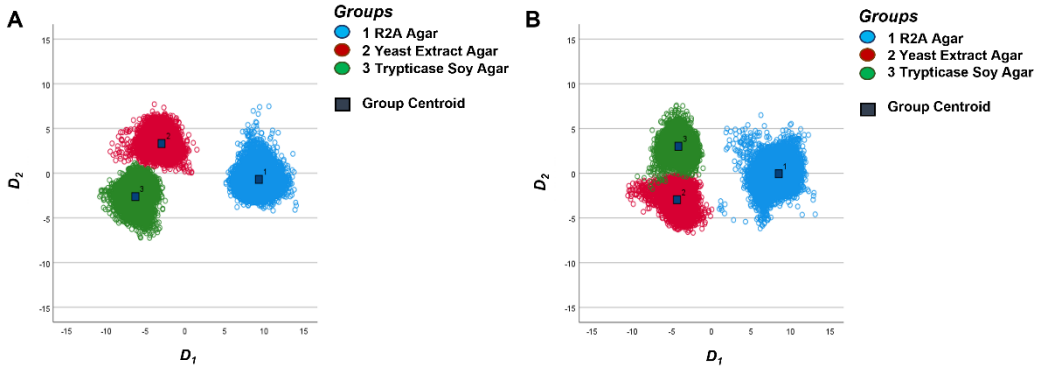


Figure 5 DA on the level of culture media's type with the mass spectra of *E. coli* ATCC 13706 (A) and *E. coli* DSM 11250 (B) with the same settings as detailed in the caption of Figure 2

The results of DA considering the type of culture media used to cultivate two *E. coli* strains (ATCC 13706, DSM 11250) are shown in Figure 5. As a result of DA, three

groups were formed considering the type of *E. coli* strains, Group1 (R2A agar), Group2 (Yeast Extract Agar) and Group3 (TSA) could be distinguished for each bacterial strain. DA produced a clear separation of bacterial mass spectra generated on the different culture media for both strains. The three media differ in terms of composition. R2A agar is a low-nutrient medium used for microbial monitoring of treated potable water, whereas YEA is a nutrient rich culture medium used for the plate count of organisms in water. TSA is a nonselective culture medium providing enough nutrients to cultivate a wide variety of microbes. The differences of generated spectra by the applied culture media were displayed by DA as groups of two high-nutrient media (YEA, TSA) are closer to each other, while the group of low-nutrient R2A agar is distant. In the case of *E. coli* ATCC 13706, 100% successful grouping for Group1 and Group3 were achieved, while for Group2 the grouping was 99.9% successful. Regarding *E. coli* DSM 11250, grouping for Group2 was 100% successful while grouping for Group1 and Group3 were 99.9% successful, respectively. Based on the test of equality of group means, significance level of each variable is <0.05, meaning that the group centroids of the three groups were not equal, they were clearly separated for both *E. coli* strains. Classification was on Mahalanobis distances to the group centroids and the result was compared to the original classification.

Table 10 Identification results of MALDI-TOF MS obtained on different culture media regarding *E. coli* isolates

	Yeast Extract Agar		R2A Agar		TSA	
Bacterial strains	*E. coli* DSM 11250	*E. coli* ATCC 13706	*E. coli* DSM 11250	*E. coli* ATCC 13706	*E. coli* DSM 11250	*E. coli* ATCC 13706
Identification score values	2.5	2.43	2.44	2.42	2.42	2.36

Although, there is a minor difference regarding the confidence of identifications, MALDI-TOF MS was able to identify both *E. coli* strains at species level on all three culture media used in this study. The results show that secure species identification can be achieved on the three examined culture media (YEA, TSA, R2A). Thus, the best culture media proved to be the Yeast Extract Agar because the application of it generated the highest identification scores. The score values of identifications are shown in Table 10. One-way ANOVA showed that there was no significant difference between the score values of identification, therefore the confidence of identification of

E. coli isolates was not differed ($p>0.05$). Kolmogorov–Smirnov test ($p>0.05$) proved that model residuals had normal distribution and homogeneity of variance was checked by Levene's test ($p>0.05$).

The present results bring a new point of view on identifying and discriminating bacteria by the MALDI-TOF MS technique. Regarding growth conditions before microbial identification, by applying discriminant analysis and principal component analysis on mass spectral data, the effect of different culture media on the confidence of identification could be successfully shown as well as differentiating bacterial strains.

5.5 Comparing MALDI-TOF MS and 16S rRNA gene sequencing to identify waterborne bacteria

In recent years, several authors evaluated the performance of MALDI-TOF MS in environmental microbiology by comparing it to 16S rRNA gene sequencing, the golden standard of bacterial identification. However, data on its efficacy in identifying waterborne microorganisms isolated directly from the environment, particularly from irrigation water, is limited. In an extensive environmental microbiology study, Uchida-Fuji et. al (2020) showed the potential of MALDI-TOF MS (Bruker Biotyper) in environmental microbiology as the authors were able to identify 86.2% of 3724 isolates at species level. In addition, MALDI-TOF MS (Bruker Biotyper) and 16S rRNA gene sequencing identification techniques have been compared in various environments such as environmental mining samples, high-altitude soil samples, soil samples and fresh vegetables (Avanzi et al. 2017; Strejcek et al. 2018; El-Nemr et al. 2019; Pandey et al. 2019;). Therefore, with the experiments presented below a comparison of MALDI-TOF MS to 16S rRNA gene sequencing regarding waterborne isolates can be made. Water samples in these experiments include samples from Kengyel, Karcag, Rákóczifalva and two samples from Szolnok from different sampling spots. Water samples are marked as Sample1 (Kengyel), Sample2 (Karcag), Sample3 (Rákóczifalva), Sample4 (Szolnok1), Sample5 (Szolnok2).

The applied methods resulted similar identification outcomes as both Sanger sequencing and MALDI-TOF MS identified more than 60% of the 42 waterborne isolates similarly at species level (Table 11). However, the application of MALDI-TOF MS made it possible to identify more isolates at species level. At genus level identification, a minor disparity was noticed as MALDI-TOF MS could identify more

isolates properly. However, a paired t-test showed that the identification results of the two methods did not differ significantly t(41)=2.02; p=0.57).

Most isolates identified by MALDI-TOF MS were classified as Gram-negative bacteria. The most frequently cultivated isolates belonged to genus *Acinetobacter*, *Enterobacter*, *Pseudomonas* and *Brevundimonas*. MALDI-TOF MS were not able to identify two isolates at any level, but those were categorized as a *Pseudomonas stutzeri* and a *Sphingobacterium kitahiroshimense* isolates with high-confidence by Sanger sequencing. Four isolates from the genera of *Acinetobacter*, *Pseudarthrobacter* and *Stenotrophomonas*, were not identified even at genus level by Sanger sequencing.

Table 11 Identification results of Sanger sequencing and MALDI-TOF MS regarding each waterborne isolate shown at genus level

		16S rRNA gene sequencing identification			MALDI-TOF MS identification		
Bacterial genus	# of isolates	Species level ≥98.5%	Genus level ≥95%	No ID <95%	Species level ≥2	Genus level ≥1.7	No ID <1.7
Acinetobacter	20	16	18	2	16	20	
Aeromonas	1	1	1		1	1	
Brevundimonas	3	3	3		3	3	
Chryseobacterium	1	1	1		1	1	
Enterobacter	5	1	5		2	5	
Microbacterium	1	1	1		1	1	
Pantoea	1		1		1	1	
Pseudarthrobacter	1			1		1	
Pseudomonas	5	2	5		3	4	1
Rhodococcus	2	1	2			2	
Sphingobacterium	1	1	1				1
Stenotrophomonas	1			1	1	1	
Total isolates	42	27 (64.3%)	38 (90.5%)	4 (9.5%)	28 (66.7%)	40 (95.2%)	2 (4.8%)

The first isolate was identified as *Acinetobacter ursingii* with high-confidence by MALDI-TOF MS, however with Sanger sequencing only 90.41% similarity was achieved. The next isolate was also identified as an *Acinetobacter ursingii* isolate with low-confidence by MALDI-TOF MS while with Sanger sequencing 91.67% similarity was obtained. The third isolate was identified as a member of *Pseudarthrobacter* at genus level by MALDI-TOF MS and it was related to a *Pseudarthrobacter siccitolerans* isolate with 89.91% similarity by Sanger sequencing. The fourth isolate was identified as *Stenotrophomonas maltophilia* by MALDI-TOF MS with high-confidence, but only 94.14% similarity was achieved by Sanger sequencing.

The most dominant genus was *Acinetobacter* with 20 isolates among the cultivated genera. Four of 20 *Acinetobacter* isolates were only identified at genus level by MALDI-TOF MS while 16 were identified at species level. Similarly, 16 *Acinetobacter* isolates were identified at species level by 16S rRNA gene sequencing, while 2 isolates were only identified at genus level and 2 more not at all. One isolate was identified with low-confidence as *Acinetobacter schindleri* by MALDI-TOF MS, while it was identified at species level similarly by Sanger sequencing.

The application of both methods resulted in similar outcomes in terms of identifying *Enterobacter* isolates. MALDI-TOF MS identified two *Enterobacter* isolates, one as *Enterobacter hormaechei* while the other being *Enterobacter cloacae*. Three isolates could not be identified at species level, because they had identical species identification scores for multiple species in both methods.

Isolates from the genus *Pseudomonas* were also frequent as five isolates were categorized into it. One isolate could not be identified by MALDI-TOF MS, but it was identified by Sanger sequencing as its sequence had 99.64% similarity score with sequences of *Pseudomonas stutzeri*. Sanger sequencing could not differentiate three *Pseudomonas* isolates correctly, however two of those isolates were identified as *Pseudomonas veronii* with high-confidence by MALDI-TOF MS. One isolate was identified correctly at species level as *P. stutzeri* by both methods.

Eleven of the 42 isolates were identified differently by MALDI-TOF MS and 16S rRNA gene sequencing (Table 12). Only two isolates were identified differently of the genus *Acinetobacter*, the most commonly found genus. In both cases MALDI-TOF MS identification resulted in *Acinetobacter junii*, whereas those isolates were identified as *Acinetobacter schindleri* by Sanger sequencing. An isolate, identified as

Rhodococcus erythropolis with low confidence by MALDI-TOF MS, was identified as *Rhodococcus qinsengii* by Sanger sequencing. Interestingly, neither of the techniques were able to identify this isolate at species level with high confidence.

Table 12 Differently identified isolates by MALDI-TOF MS and 16S rRNA gene sequencing

No.	Sample name	Isolate	MALDI-TOF MS identification (log score, consistency category)	16S rRNA identification (% similarity score)
#1	Irrigation water8	Sample5/9	*Acinetobacter junii* (2.34; A)	*Acinetobacter schindleri* (99.24%)
#2	Irrigation water8	Sample5/12	*Acinetobacter junii* (2.1; A)	*Acinetobacter schindleri* (98.78%)
#3	Irrigation water6	Sample3/1	*Rhodococcus* spp. (1.71; B)	*Rhodococcus qinsenghii* (96.2%)
#4	Irrigation water6	Sample3/3	No id (1.51; C)	*Sphingobacterium kitahiroshimense* (99.72%)
#5	Irrigation water6	Sample3/4	*Chryseobacterium indologenes* (2.01; A)	*Chryseobacterium lactis* (98.8%)
#6	Irrigation water4	Sample2/4	*E. hormaechei* (2.25; A)	*E. cloacae/E. hormaechei* (99.9%)
#7	Irrigation water4	Sample2/5	*Pseudarthrobacter scleromae/oxydans* (2.24; B)	*Pseudarthrobacter siccitolerans* (89.91%)
#8	Irrigation water4	Sample2/6	*Rhodococcus* spp. (1.99; B)	*Rhodococcus cerastii* (99.46%)
#9	Irrigation water4	Sample2/7	*E. cloacae* (2.27; A)	*E. hormaechei* (99.48%)
#10	Irrigation water4	Sample2/8	*P. veronii* (2.26; A)	*P. veronii/ P. extremaustralis* (100%)
#11	Irrigation water4	Sample2/9	*P. veronii* (2.2; A)	*Pseudomonas* spp. (99.34%)

Discrepancies were also found among *Enterobacter* and *Pseudomonas* isolates. Isolates marked as #6 and #9 were identified as *E. hormaechei* and *E. cloacae* with high confidence by MALDI-TOF MS, however Sanger sequencing could not differentiate the former as sequences of both *E. cloacae* and *E. hormaechei* showed 99.9% similarity. Although isolate #9 was identified as *E. cloacae* with high confidence, it was identified as *E. hormaechei* by Sanger sequencing with 99.48% similarity. Isolates marked as #10 and #11 were identified as *Pseudomonas veronii* with high confidence by MALDI-TOF MS, while the former was identified as *P. veronii*/*P. extremaustralis* showing 100% similarity scores for both species by Sanger sequencing. The best matched hit for the latter was an uncharacterized *Pseudomonas* species.

The applicability of MALDI-TOF MS and its databases (Bruker Biotyper, VITEK MS) have been studied and validated for clinical microbiology laboratories in recent years (Christner et al. 2010, Ferreira et al. 2010, Martiny et al. 2012, Scott et al. 2016, Fan et al. 2017). However, its application regarding isolates derived from the food production chain and its environment is challenging due to the microbial diversity in soil and water matrices which contain thousands of different bacterial species (Mauchline and Malone 2017; Zancarini et al. 2017). Strejcek et al. (2018) used both MALDI-TOF MS (Bruker Biotyper) and 16S rRNA gene sequencing to identify microorganisms from soils and sediments and achieved concordant genus level identification (92%) while at species level only 35% of the isolates identified coincided with those identified by 16S rRNA gene sequencing analysis. Kopcakova et al. (2014) utilized MALDI-TOF MS (Bruker Biotyper) to identify the microflora from waste disposal sites with an overall identification rate lower than 20% at species level. However, Suzuki et al. (2018) applied MALDI-TOF MS (Bruker Biotyper) to identify coliform bacteria from sewage, river water and groundwater, obtaining identical results at genus level in 96%, 74%, and 62% of the isolates respectively compared to 16S rRNA gene sequence analysis.

In congruence with the study of Suzuki et al. (2018), the results in this subchapter suggest that MALDI-TOF MS can be used to identify waterborne bacterial isolates, as more isolates were identified at species level than with 16S rRNA gene sequencing. MALDI-TOF MS generated 95.2% correct genus level identification of the cultivated 42 isolates which were higher than the results of 16S rRNA gene sequencing (90.5%). Moreover, 73.8% of the isolates were identified identically with both methods. Four of the isolates (9.5%) were not identified at any level by 16S rRNA gene sequencing while only two (4.7%) isolates were unidentified by MALDI-TOF MS.

Moreover, Böhme et al. (2013) also compared the efficacy of MALDI-TOF MS (Voyager STR-DE, Applied Biosystems) and 16S rRNA gene sequencing and pointed out that MALDI-TOF MS identified 76% of 50 seafood-borne bacterial strains isolated from commercial seafood products at species level while 16S rRNA gene sequencing only identified the species of 50% of the strains. The results in this chapter are also in congruence with the study of El-Nemr et al. (2019) in which MALDI-TOF MS identified more bacteria isolated from a market area (e.g., vegetables, soil, air and hand palms of fresh produce handlers) at species level (41%) than 16S rRNA gene sequencing (28%) did. In another environmental study, Pandey et. al (2019) identified

psychrotolerant bacteria isolated from high altitude soil with only 4.92% of the isolates identified similarly by MALDI-TOF MS and 16S rRNA gene sequencing, whereas in our case 73.8% of the isolates were identified identical. Findings in this subchapter are close to the study of Avanzi et al. (2017) in which 82% of the copper resistant bacteria, isolated from environmental mining samples (soil and water), were identified identically with both methods. Besides, in the afore-cited study of Pandey et al. (2019), 19.67% of the isolates were not identified at any level by MALDI-TOF MS which value is higher than the result (4.7%) achieved in this subchapter. The lower identification scores of the study of Pandey et al. (2019) can be explained by the fact that at the time of its conduction, the database not contained some of the unidentified isolates (*Bacillus wiedmannii*, *Bacillus velezensis*, *Bacillus paramycoides*).

In this study, most of the cultivable isolates belonged to the genus *Acinetobacter*. Regarding identifying species of this genus, both methods could achieve almost similar results as 16 isolates were identified at species level. MALDI-TOF MS (Bruker Biotyper) outperformed 16S rRNA gene sequencing as the former identified four isolates at genus level, while the latter besides identifying two isolates at genus level, was not able to identify two isolates at any level. Species of this genus have been isolated from agricultural and hydrocarbon-polluted soils, water, sediment, industrial wastewater and sewage (Adewoyin and Okoh 2018). Most of the *Acinetobacter* isolates were identified as *A. junii* which has been previously reported to be found in aquatic environments such as wastewater (Weidmann-Al-Ahmad et al. 1994), sewage (Guardabassi et al. 1999), surface water (Goswami et al. 2015), and marine sediments (Roseline and Sachindra 2016). Other isolates were identified as *A. schindleri* which species has been isolated from different sources before for example from soil samples (Choi et al. 2012), live-stock animals and pets (Rafei et al. 2015), head lice from primary school pupils (Sunantaraporn et al. 2015). *Acinetobacter ursingii*, also isolated from human fecal samples previously (Dijkshoorn et al. 2005) and from raw meat (Carvalheira et al. 2017), was also found in the irrigation water samples.

In these measurements it was successfully shown that MALDI-TOF MS (Bruker Biotyper) can act as an alternative to 16S rRNA gene sequencing of isolates to identify waterborne bacteria due its rapid and accurate nature. Currently MALDI-TOF MS is a basic instrument in several clinical microbiology laboratories (Christner et al. 2010; Ferreira et al. 2010; Martiny et al. 2012; Scott et al. 2016; Fan et al. 2017). Moreover, as the reliance on MALDI-TOF MS is increasing due to its routine application in

laboratories therefore sequencing is more likely used as a gold standard reference test for verification of MALDI-TOF MS identification of unusual or rare organisms. (Church et al. 2020). Nonetheless, as its identification process relies on matching the PMF of the measured isolate to the database, lack of entries will lead to misidentifications or not reliable identifications (Kopcakova et al. 2014; Strejcek et al. 2018). This could be seen in the previous subchapter of the results of this thesis (5.1. MALDI-TOF MS to identify bacteria from environmental samples). However, as more species' mass spectra are being generated and implemented into the commercially available mass spectral databases, the environmental applicability of MALDI-TOF MS will be even further improved. Thus, it will be thoroughly exploited in the whole spectrum of microbial identification.

5.6 16S rRNA amplicon sequencing to characterize the community of irrigation waters from the county of Jász-Nagykun-Szolnok

A common issue in environmental monitoring is that natural environments comprise a wide variety of microbial species, but 99% of bacteria are not culturable (Locey and Lennon 2016). Thus, amplicon sequence analysis of marker genes such as the bacterial 16S rRNA gene is applied to define the relative abundance of different bacterial genera and the entire bacterial community in environmental samples. Therefore, the five irrigation water samples (Kengyel as Sample1, Karcag as Sample2, Rákóczifalva as Sample3, Szolnok1 as Sample4, Szolnok2 as Sample5) which were used for the comparison of MALDI-TOF MS and 16S rRNA gene sequencing were applied in these experiments also to monitor the uncultivable part of the irrigation water regarding microbial water quality.

Altogether 730 amplicon sequence variants (ASV) were found in the five samples which comprised 82613 total read counts. After stringent quality filtering on average, 229 high quality 16S rRNA gene sequences per sample remained. Sample1 comprised the most read counts with 20785, followed by Sample3 with 19637 read counts. Sample4 contained 18368 while Sample5 consisted of 13086 read counts (Table 13).

Table 13 The number of different species and biodiversity represented in the irrigation water samples

Sample	Shannon index	Species richness	Total read counts
Sample1	4.02	148	20785
Sample2	4.08	106	10737
Sample3	4.40	276	19637
Sample4	4.99	392	18368
Sample5	4.36	224	13086

The lowest number of read counts were observed in Sample2 with only 10737 read counts. A rarefaction curve showed that all samples were sequenced deep enough to infer the full diversity of microorganisms in the samples (Figure 6). The species richness ranged from 106 ASVs in Sample2 to 392 ASVs in Sample4, whereas the Shannon index ranged from 4.02 to 4.99.

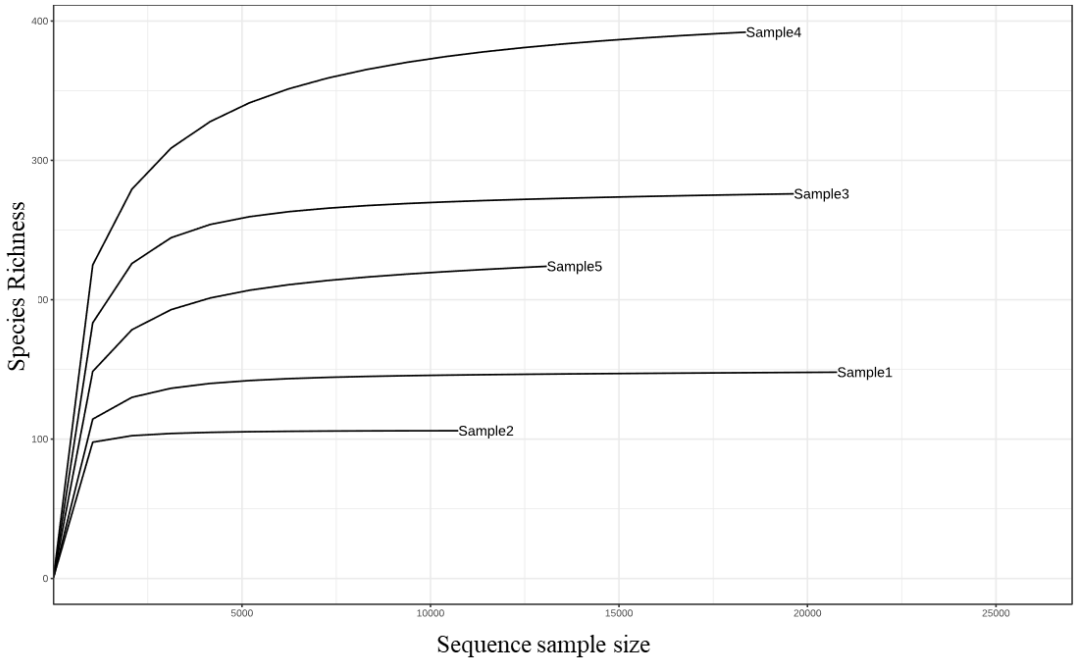


Figure 6 Species richness shown regarding the five analyzed irrigation water samples. On the figure the recorded Species Richness values are plotted against the Sequence Sample Size

The relative abundance of ASVs displays a wide range of variety at phylum level displaying 32 different phyla (Figure 7). The most abundant phylum was Proteobacteria followed by Actinobacteria, Bacteroidota, Patescibacteria and Verrucomicrobiota in the analyzed samples.

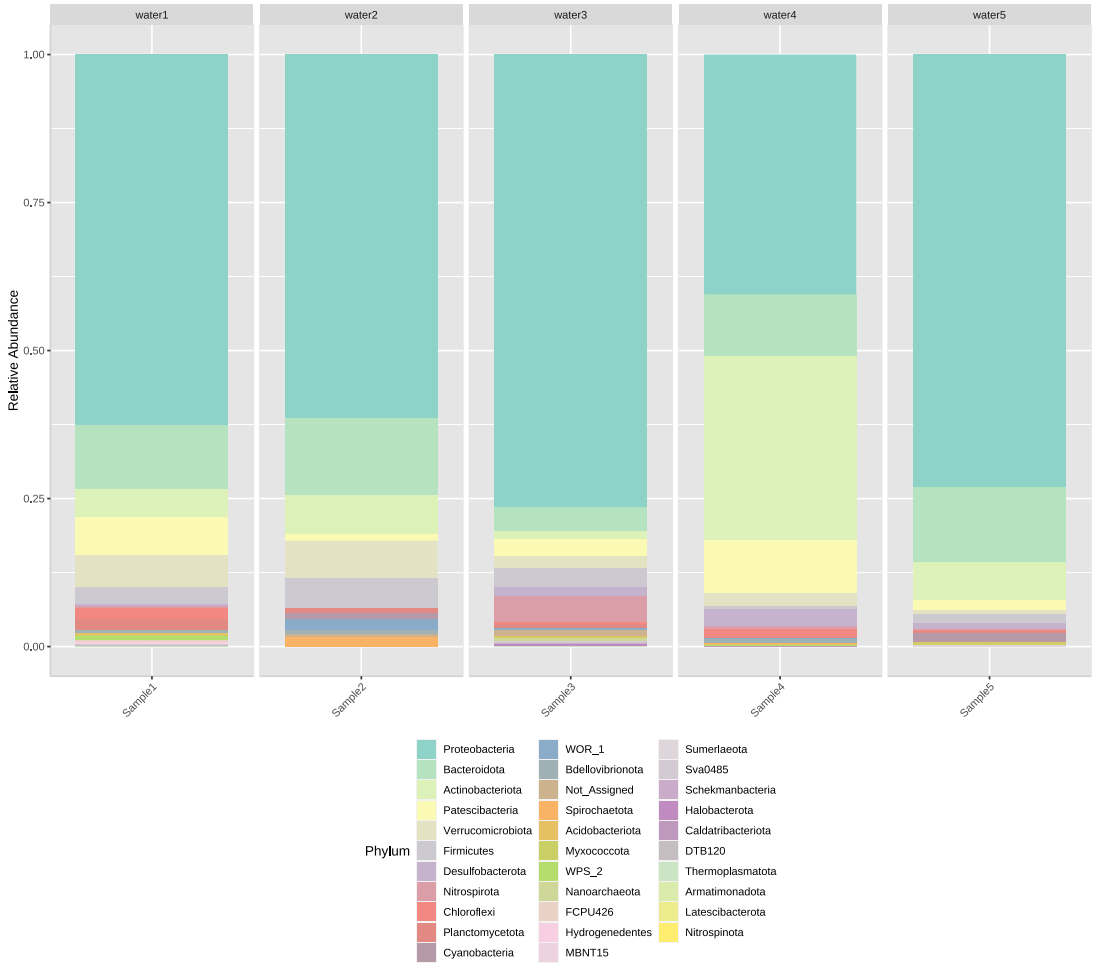


Figure 7 Relative abundance of ASVs regarding each irrigation water sample. ASVs are shown on phylum level

However, differences were noticed regarding the abundance of phyla in each sample. In Sample1, Proteobacteria (63%) were most abundant, while the other frequent phyla were Bacteriodota (11%), Patescibacteria (6%), Verrucomicrobiota (5%) and Actinobacteriota (5%). The abundance of phyla observed in Sample2 was similar to Sample5 as in both samples Proteobacteria (61%; 73%) was followed by Bacteroidota (13%; 13%), Actinobacteria (7%; 6%). Sample4 was dissimilar to the other samples, as most of the ASVs belonged to Proteobacteria (40%) while ASVs from Actinobacteria (31%) were also common, followed by ASVs belonging to Bacteroidota (10%), Patescibacteria (9%) and Desulfobacteria (3%). The dominance

of Proteobacteria could also be remarked in both Sample3 (76%) and Sample5 (73%). However, composition of these samples was varying as in Sample3 the next most abundant phyla were Nitrospirota (4%), Bacteriodota (4%), Firmicutes (3%) and Patescibacteria (3%), whereas in Sample5 the second most abundant phylum, Bacteriodota (13%), was followed by Actinobacteria (6%), Patescibacteria (2%) and Cyanobacteria (2%).

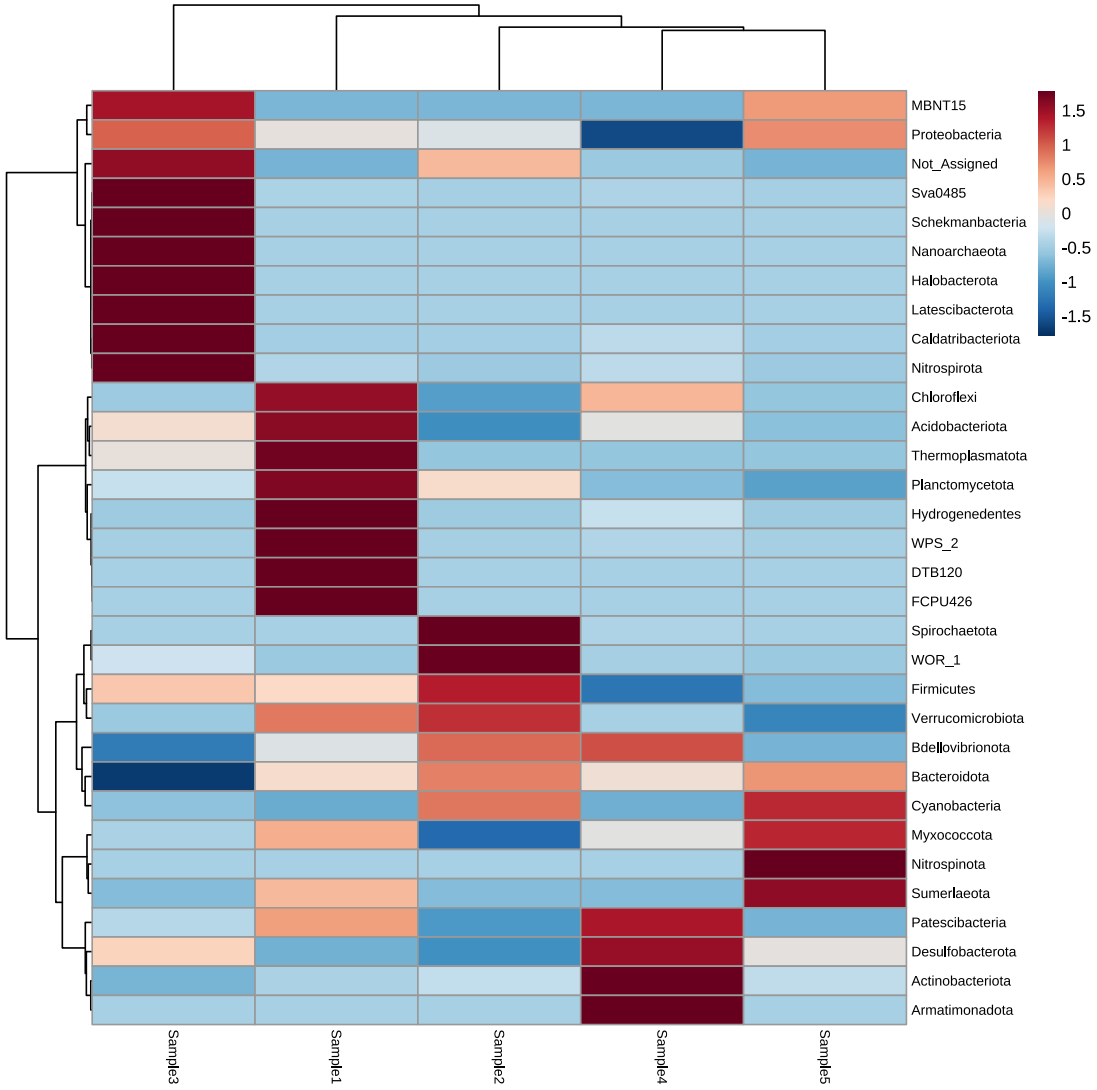


Figure 8 Taxonomic distribution of the phylogenetic groups at phylum level of irrigation water samples shown by combining Hierarchical Clustering and Heatmap visualization

The taxonomic distribution of phylogenetic groups of irrigation waters shows specific fingerprints regarding bacterial phyla in each sample (Figure 8). The use of Hierarchical Clustering classified the samples into clusters based on their microbial communities. Sample3 can be differentiated from other samples because phyla

including MBNT15, Sva0485, Schekmanbacteria, Nanoarchaeota, Halobacterota, Latescibacterota, Caldatribacteriota and Nitrospirota are mostly abundant in only that sample. In contrast, Sample4 and Sample5 are phylogenetically more related to each other. Cluster of Sample4 and Sample5 can be extended by adding Sample2, the cluster of these three samples can be further enlarged by adding Sample1. Thus, two large clusters can be differentiated with one comprising only Sample3 while the other contains the rest of the samples (Sample1, Sample2, Sample4 and Sample5).

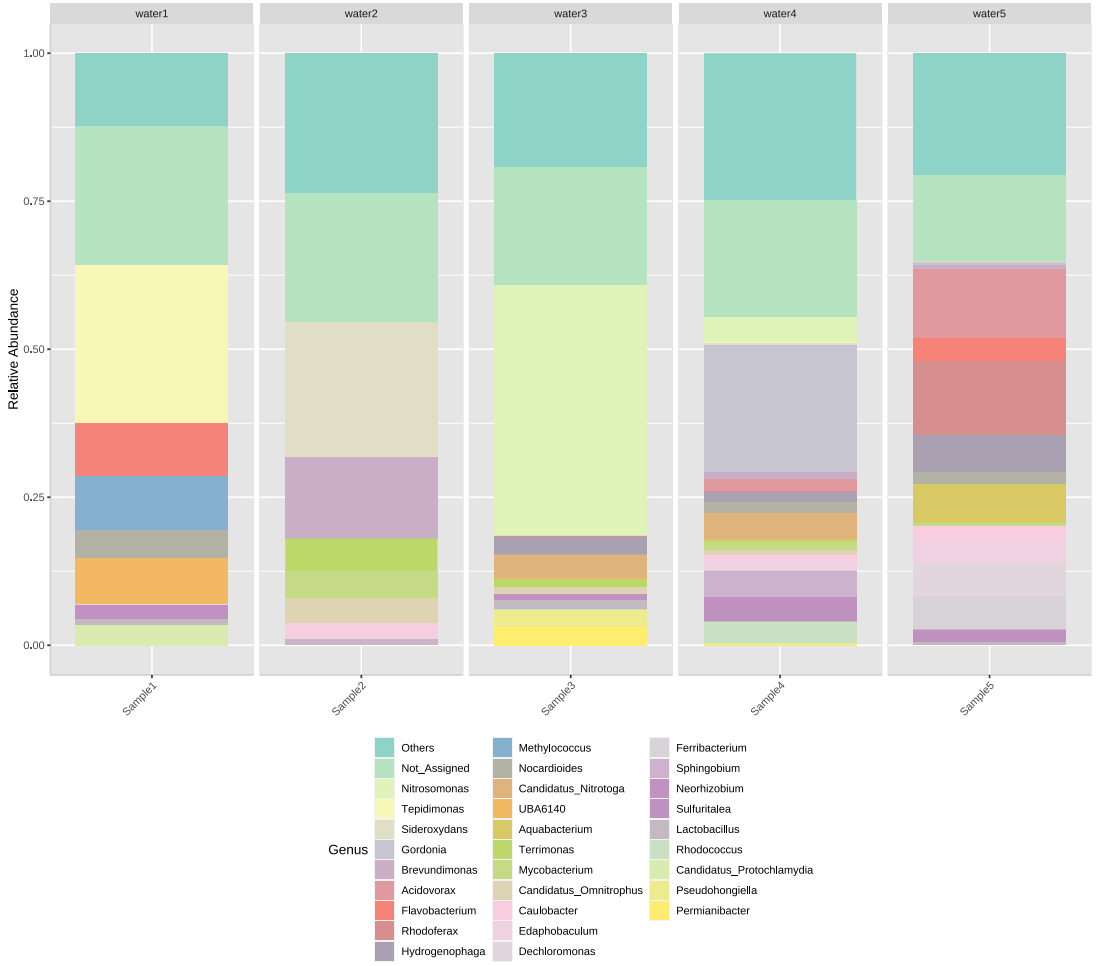


Figure 9 Relative abundance of the 30 most abundantly occurring bacterial genera in the irrigation water samples. The rest of the taxa merged into the group of Others

The microbial community of water samples was diversified as the five most abundant genera were different in each sample (Figure 9). In Sample1, the most abundant genus was *Tepidimonas* followed by *Flavobacterium*, *Methylococcus*, *Methylophilaceae* UBA6140 and *Nocardioides*. In Sample2, the most abundant genus was *Sideroxydans* which were followed by genus *Brevundimonas*, *Terrimonas*, *Mycobacterium*, and *Candidatus_Omnitrophus*. In Sample3, the most abundant genus was *Nitrosomonas*,

an ammonia-oxidizing genus, which was followed by *Candidatus Nitrotoga*, an uncultured nitrite-oxidizing and naturally occurring bacterial genus in aqueous ecosystems (Kitzinger et al. 2018) and *Permianibacter*. Genus *Hydrogenophaga*, a hydrogen oxidizing genus, and *Pseudohongiella*, of which species have been isolated from seawater (Xu et al. 2016), were also common. The abundance of genera monitored in Sample4 was comparable to Sample3 as nitrifying-bacterial genera such as *Nitrosomonas* and *Candidatus Nitrotoga* were the second and fourth most abundant genera. However, the most abundant genus was *Gordonia* and the third most abundant genus was *Sphingobium* while the fifth was genus *Rhodococcus*. In Sample5 the dominance of Comamonadaceae family could be noticed as four of the five most abundant genera belonged to that family. The most abundant genus was *Rhodoferax* followed by *Acidovorax, Hydrogenophaga, Aquabacterium* and *Dechloromonas*.

Although genus *Nitrosomonas* was the most dominant in terms of relative abundance (11.04%), zero isolates were cultivated from it. In contrast, *Acinetobacter*, the most dominant genus regarding cultivated isolates, was only the 31st in terms of relative abundance (0.64%) in the entire bacterial community (Table 14). Similarly, despite five isolates had been isolated from genera *Pseudomonas* and *Enterobacter*, their relative abundance was only 0.24% and 0.04%, respectively. Although genus *Brevundimonas* had the highest relative abundance (2.18%) among cultivated genera, only three isolates of it were cultivated. Furthermore, two isolates of the genus *Rhodococcus* were cultivated, which also had the second highest relative abundance (0.81%) value among cultivated genera. Although only one isolate was cultivated and identified as a member of the genus *Chryseobacterium* its relative abundance (0.35%) was the fourth highest among cultivated genera.

Table 14 Identified bacterial isolates and their relative abundance in the next-generation sequencing dataset

Bacterial genus	Number of isolates	Relative abundance of the genera
Brevundimonas	3	2.18%
Rhodococcus	2	0.81%
Acinetobacter	20	0.64%.
Chryseobacterium	1	0.35%
Pseudomonas	5	0.24%
Enterobacter	5	0.04%
Stenotrophomonas	1	0.03%
Sphingobacterium	1	0.02%
Aeromonas	1	<0.01%
Microbacterium	1	<0.01%
Pantoea	1	<0.01%
Pseudarthrobacter	1	<0.01%

Still, to identify and thoroughly characterize the bigger, uncultivable part of the microbial community of irrigation water, culture-independent methods such as amplicon sequence analysis of 16S rRNA genes are necessary. The results of the next generation sequencing approach showed a dominance of Proteobacteria followed by Actinobacteria, Bacteroidota, Patescibacteria, Verrucomicrobiota and Firmicutes. In the samples, Proteobacteria (62%) was by far the most abundant phylum while Actinobacteria (10%), Bacteriodota (10%), Patescibacteria (5%) Verrucomicrobiota (3%) and Firmicutes (2%) occurred less frequently. It is also notable that Actinobacteria were more abundant in Sample4 (32%) compared to the other samples (<10%). This sample was the only artesian water sample included in the study.

Jin et al. (2018) examined the microbial community characteristics of 16 surface water samples in the Beijing area applying 16S rRNA gene amplicon sequencing and found that Proteobacteria and Bacteroidetes were the most commonly identified phyla in all the samples, accounting for 21.9–78.5% and 19.1–74.7% of the sequences, respectively. Lehosmaa et al. (2021) analyzed the bacterial communities of groundwater-surface water ecotone of boreal springs and observed that the bacterial

communities were dominated by Proteobacteria (50%) based on relative abundance, followed by Bacteroidetes (18%), Patescibacteria and Acidobacteria (4% each). In the study of Jin et al. (2018) the most predominant genera among Proteobacterial sequences were *Hydrogenophaga* and *Rhodoferax* both of which were found to be dominant in the Hungarian irrigation water samples as well. The former was dominant in both Sample3 and Sample5 whereas the latter in Sample5. Moreover, genus *Hydrogenophaga* had both high relative abundance and prevalence in the samples analyzed in this study. Genus *Flavobacterium*, the predominant Bacteroidetes genus in the afore-cited study of Jin et al. (2018), had also high relative abundance in this study. Iliev et al. (2017) used 16S rRNA gene amplicon sequencing to characterize microbial freshwater communities in two Bulgarian reservoirs and found that Proteobacteria, Actinobacteria and Bacteroidetes contained more than 95% of the relative abundance, regardless of the reservoir's large hydrogeological differences. These findings are in congruence with the results obtained in this study, suggesting that Proteobacteria and Bacteroidetes are among the dominant phyla in different water bodies across the globe (Jin et al. 2018; Lehosmaa et al. 2021).

Nitrite-oxidizer bacteria were not cultivated but sequences of those were frequent in the samples as *Nitrosomonas* was the most abundant genus in Sample3 and the second most abundant in Sample4. Moreover, *Nitrotoga*, a main nitrite-oxidizer in activated sludge systems with nutrient removal (Saunders et al. 2016), was the second most abundant genus in Sample3 and the fourth most abundant in Sample4. The isolates were dominated by the genus *Acinetobacter*, although it only had the third highest relative abundance value among cultivated genera with the first one being genus *Brevundimonas*. Furthermore, while having cultivated several *Enterobacter* and *Pseudomonas* isolates in this study, two ubiquitous and potentially pathogenic genera, ASVs belonging to any of those genera had a low relative abundance in the samples. The majority of the cultivated genera (8 of 12) had a relative abundance of at least 0.01% in the amplicon dataset. Moreover, only 8 of 188 (4.25%) genera, which had a relative abundance above 0.01% were cultivated. Thus, highlighting the fact that most of the environmental bacteria are uncultivable.

6 CONCLUSIONS AND RECOMMENDATIONS

As the population of the world increases so does the demand for more food which can only be ensured by proper agricultural technologies involving irrigation. Therefore, high-quality irrigation water is crucial in order to provide sufficient yields. However, limited water resources and unpredictable precipitation infer the application of recycled and microbiologically non-characterized waters. The application of those is directly linked to a higher occurrence of foodborne diseases as those waters are perfect "culture media" for pathogens. The risk increases when irrigated produce such as fruits and vegetables are consumed raw. Thus, detecting and identifying bacteria or even analyzing the complete microbiome of irrigation water used for food production and its environment affecting the food production chain can prevent the increasing numbers of foodborne diseases related to fresh products.

In the first part of the thesis, to provide fast and trustworthy identification of food- and waterborne bacteria, Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) was applied to identify microorganisms from agricultural environment involving irrigation water, running water, manure and vegetables. A comprehensive picture about the identification performance of MALDI-TOF MS (Biotyper) was provided regarding environmental microbiology by analyzing 311 isolates. The results suggest that Gram-positive bacteria is more difficult to identify by MALDI-TOF MS (Biotyper) as lower (25.5%) species identification scores were achieved compared to Gram-negative isolates (40.9%). However, data obtained from the measurements indicates that MALDI-TOF MS (Biotyper) can be a reliable technique to identify bacteria from agricultural environment at genus level, but to realize similar results at species level, on which 36.6% identification scores were achieved, a database expansion with environmental isolates is inevitable.

Two types of MALDI-TOF MS sample preparation techniques, extended direct transfer procedure involving formic acid extraction and direct transfer procedure, were compared to determine which the best one is for environmental isolates. The average log scores of the former were above 2 therefore reached the species level threshold while the average log scores of the latter was only 1.85. Besides, paired t-test proved that extended direct transfer procedure was significantly superior to direct transfer procedure.

The effect of culture media on MALDI-TOF MS identification regarding waterborne bacteria was also tested using Tryptic Soy Agar, R2A and Yeast Extract Agar. It was proved by ANOVA that all three culture media are suitable to identify waterborne bacteria as no significant difference was observed regarding the averages of log scores. However, the species level identification threshold was not achieved using any of the culture media (TSA, 1.97; R2A, 1.95; YEA, 1.97). In these measurements, even higher identification scores could have been achieved by applying lower cutoff scores from 2 to ≥1.9, however that necessitates more measurements.

In the next part, it was shown that applying discriminant analysis on the mass spectra of isolates can help understanding the effect of culture media on the identification. Therefore, groups according to spectra obtained on specific culture media could be created which explained which culture medium the best is to use for the identification of *E. coli* and *S. aureus* isolates.

Moreover, it was possible to differ strains of *E. coli* from each other by applying principal component analysis on mass spectral data. Besides, using the same method it was also possible to distinguish MRSA from non-resistant *S. aureus*. Therefore, MALDI-TOF MS was able to strain type of waterborne and antibiotic-resistant bacteria. In the future, a study expanded with more strains with an emphasis on antibiotic-resistant bacteria could be performed.

In the following section, a comparison of MALDI-TOF MS to 16S rRNA gene sequencing regarding waterborne isolates were performed. Both methods resulted similar identification outcomes as more than 60% of the waterborne isolates were identified at species level by Sanger sequencing and MALDI-TOF MS as well. The application of MALDI-TOF MS made it possible to identify more isolates at both species and genus level, but a paired t-test showed that the identification results of the two methods did not differ significantly. Interestingly, 11 of 42 (26.2%) isolates were identified differently with the only discrepancies were observed at species level in the case of 10 isolates. Therefore, it was successfully demonstrated that MALDI-TOF MS (Biotyper) can act as an alternative to 16S rRNA gene sequencing of isolates to identify waterborne bacteria due its rapid and accurate nature.

In the last section, MALDI-TOF MS was coupled with 16S rRNA gene amplicon sequencing to monitor the quality of irrigation water. As 99% of bacteria are not culturable, a lower concentration of microbial pathogens might not be cultivated which

hinder the occurrence of those pathogens in irrigation waters. Interestingly, genera that were cultivated the most (*Acinetobacter*, *Pseudomonas*, *Enterobacter*, *Brevundimonas*) had a low relative abundance in the amplicon dataset. This highlights the fact that most of the environmental bacteria are uncultivable and the need for culture-independent methods to monitor natural waters. Based on this fact it would probably be worth to test more Hungarian irrigation waters by these methods.

7 THESES – NEW SCIENTIFIC RESULTS

1. For the first time, MALDI-TOF MS was utilized to identify bacteria from Hungarian agricultural environment involving irrigation water, running water, lakes, manure and vegetables. Thus, a broad picture about the identification performance of MALDI-TOF MS (Biotyper) was given regarding environmental microbiology by analyzing 311 bacteria with the genera of *Pseudomonas*, *Acinetobacter*, *Bacillus* and *Aeromonas* were the most frequently occurring ones. Moreover, the results of identifications also highlight the fact that environmental isolates of Gram-positive bacteria are more difficult to identify as lower (25.6%) species identification scores were obtained compared to Gram-negative isolates (40.9%).

2. I could demonstrate that extended direct transfer procedure was superior to identify environmental bacteria compared to direct transfer procedure as the averages of log scores of the former were 2.01, reaching the species level threshold while the latter only averaged 1.85. It was proven by paired t-test ($t=16.09$, $p<0.001$) that the difference was significant.

3. I could prove that by using discriminant analysis that the MALDI-TOF MS mass spectra can be separated based on the culture medium on which the given microbe was cultured. In addition, I found that Yeast Extract Agar and R2A agar can also be used to identify *E. coli* and *S. aureus* isolates from strain collections using MALDI-TOF MS.

4. For the first time it was demonstrated that R2A and Yeast extract agar were also appropriate to identify waterborne isolates as no significant disparity was observed regarding the averages of log scores. However, the species level identification threshold was not accomplished using any of the culture media (TSA, 1.97; R2A, 1.95; YEA, 1.97).

5. I could demonstrate that it is possible to differentiate bacterial strains by combining MALDI-TOF MS with principal component analysis (PCA). New biomarkers were found for *E. coli* ATCC 13706 at m/z 6640 and m/z 8912 therefore it was possible to distinguish from *E. coli* DSM 11250. Moreover, a specific peak was detected at m/z 5868 in the mass spectrum of the methicillin-resistant *S. aureus* ATCC 43300 therefore it was possible to differentiate the antibiotic-resistant strain from the non-resistant *S.*

aureus. Thus, present results bring a new point of view on identifying and discriminating bacteria by the MALDI-TOF MS technique utilizing PCA.

6. For the first time, the identification performance of MALDI-TOF MS using Hungarian waterborne bacterial isolates was tested against Sanger sequencing. Sanger sequencing and MALDI-TOF MS generated almost identical results as 64.3% and 66.7% of the isolates were identified at species level with the methods. However, using MALDI-TOF more isolates were identified at both species and genus level, but a paired t-test proved that the identification results of the two methods did not differ significantly ($t(41)=2.02$; $p=0.57$). By comparing MALDI-TOF MS to 16S rRNA gene sequencing regarding waterborne isolates it was successfully proven that MALDI-TOF MS (Biotyper) can be a great option to 16S rRNA gene sequencing of isolates to identify waterborne bacteria due its fast and accurate nature.

7. For the first time MALDI-TOF MS was coupled with next-generation sequencing to monitor the irrigation waters in Eastern-Hungary. Results suggested that genera that were cultivated the most (*Acinetobacter*, 0.64%; *Pseudomonas*, 0.24%; *Enterobacter*, 0.04%; *Brevundimonas*, 2.18%) had a low relative abundance in the amplicon dataset. Therefore, highlighting the need for the culture-dependent techniques to be supplemented with culture-independent methods to monitor natural waters as only 8 of 188 (4.25%) genera, which had a relative abundance above 0.01%, were cultivated.

8 SUMMARY

In order to provide an appropriate amount of food, agriculture must involve irrigation as a supplementation to cultivation of crops. However, due to the scarcity of water resources agriculture relies on microbiologically non-characterized or reused water, the application of which impose a risk as it is directly linked to a higher occurrence of foodborne diseases. The risk is even bigger considering raw eaten vegetables and fruits. As such almost 50 outbreaks connected with the consumption of tap and well water resulted 1969 cases in 2019 in the EU while 31 outbreaks related to vegetables and juices caused 626 cases in 2018 (European Food Safety Authority 2019; 2021b). Moreover, 22.8%-46% of foodborne illnesses were related to fresh produce in the period of 1998-2008 in the US and 2240 confirmed cases of illnesses were reported as the affected vehicles were romaine lettuce and spinach (Uyttendaele et al. 2015; Turner et al. 2019). Therefore, monitoring irrigation water regarding its microbial quality should be performed thoroughly. Thus, detecting and identifying microbes or even studying the complete microbiome of irrigation water used for food production and its environment affecting the food production chain can prevent the increasing numbers of foodborne diseases.

Thus, rapid and accurate microbiology identification methods are needed with which the water used for food production can be analyzed. Therefore, MALDI-TOF MS a rapid and accurate microbiology identification technique, already proven its usefulness in clinical microbiology, was used to identify bacteria from agricultural environment. The identification performance of MALDI-TOF MS (Biotyper) was demonstrated regarding environmental microbiology by analyzing 311 isolates. The most common genera were *Pseudomonas*, *Acinetobacter*, *Bacillus* and *Aeromonas*. The results indicate that identification of Gram-positive environmental isolates by MALDI-TOF MS is cumbersome as lower (25.6%) species identification scores were obtained compared to Gram-negative isolates (40.9%). However, data obtained from the measurements shows that MALDI-TOF MS (Biotyper) can be a useful method for bacterial identification from agricultural environment at genus level, but to achieve similar results at species level a database expansion with environmental isolates is needed.

The comparison of two MALDI-TOF MS sample preparation methods, extended direct transfer procedure involving formic acid extraction and direct transfer

procedure, could determine which is the best one for environmental isolates. Paired t-test proved that extended direct transfer procedure was significantly superior to direct transfer procedure as the average log score of the former was 2.01 therefore reached the species level threshold while latter generated only a log score of 1.85.

By applying discriminant analysis on the mass spectra of isolates using three culture media (Tryptic Soy Agar, R2A and Yeast Extract Agar) for *E. coli* isolates and four (the aforementioned three plus Baird-Parker) for *S. aureus* isolates, it was possible to determine which culture media is the best for the mentioned species. The best culture medium was found to be the Yeast Extract Agar for *E. coli* isolates with all three culture media were suitable to species level identification while for *S. aureus* the best one was Tryptic Soy Agar.

As it was demonstrated previously that R2A and Yeast Extract Agar are also appropriate culture media to identify isolates from strain collection, these two culture media were tested against TSA to determine which one is the best for waterborne isolates. ANOVA proved that all three culture media were applicable to identify waterborne isolates as no significant disparity was described regarding the averages of log scores. However, in each case the averages of the log scores were below 2, therefore the species level identification threshold was not accomplished by applying any of the culture media (TSA, 1.97; R2A, 1.95; YEA, 1.97).

The strains of *E. coli* (ATCC 13706, DSM 11250) and *S. aureus* (ATCC 25923, ATCC 43300) were used with a purpose to assess whether strain typing of bacterial isolates is possible with MALDI-TOF MS. Thus, principal component analysis was applied on mass spectral data. Regarding *E. coli* isolates it was possible to differentiate the two strains by finding two possible biomarkers for *E. coli* ATCC 13706 at the peaks of m/z 6640 and m/z 8912. Moreover, another strain specific peak was detected at m/z 5868 in the mass spectrum of the methicillin-resistant *S. aureus* ATCC 43300. Thus the results demonstrated here bring a new viewpoint on identifying and strain typing bacteria by the MALDI-TOF MS technique exploiting principal component analysis.

MALDI-TOF MS was compared to 16S rRNA gene sequencing, the golden standard of bacterial identification, regarding bacteria isolated from irrigation water. Both methods identified the analyzed isolates with more than 60% success at species level while a paired t-test proved that the identification results of the two methods did not differ significantly. Therefore, it was demonstrated that MALDI-TOF MS (Biotyper)

can be a valuable choice to 16S rRNA gene sequencing of isolates to identify waterborne bacteria due its fast, accurate and labor-saving nature.

By using MALDI-TOF MS supplemented with next-generation sequencing, irrigation waters of Eastern-Hungary were monitored by a culture-dependent and a culture-independent method. Interestingly, the most commonly cultivated bacterial genera (*Acinetobacter*, *Pseudomonas*, *Enterobacter*, *Brevundimonas*) showed low relative abundance in the next-generation sequencing dataset while bacterial genera with high relative abundance were not cultivated. Thus, these results clearly present the importance of supplementing culture-dependent methods with culture-independent ones in order to obtain a picture about the whole bacterial community composition of irrigation water therefore to assess whether irrigation water used in food production pose a potential risk in the food production chain.

It should be also mentioned that bacterial isolation and identification were performed using general purpose culture media without any preenrichment, selective enrichment or applying selective culture media. Thus, the MALDI-TOF MS identification process can save crucial time and reduce the number of additional cultivations.

The methods and its applications presented in this thesis could help the food industry by detecting and identifying food- and waterborne bacteria from their natural habitats, strain typing bacteria of public health concern and assessing the whole bacterial communities of water used for food production with which indirectly the number of foodborne infections could be reduced.

9 PUBLICATIONS

Botond Bendegúz Surányi, Benjamin Zwirzitz, Csilla Mohácsi-Farkas 1, Tekla Engelhardt 3 and Konrad J. Domig **(2023)** Comparing the efficacy of MALDI-TOF MS and sequencing-Based identification techniques (Sanger and NGS) to monitor the microbial community of irrigation water. *Microorganisms*, 11(2), 287. **Q2, IF: 4.926.** doi: https://doi.org/10.3390/microorganisms11020287

Anna Jánosity, József Baranyi, **Botond Bendegúz Surányi**, Sonja Smole Možina, Andrea Taczman-Brückner, Gabriella Kiskó, Anja Klančnik. **(2023)** Estimating the optimal efflux inhibitor concentration of carvacrol as a function of the bacterial physiological state. *Frontiers in Microbiology*, 14, 1073798. **Q1, IF: 6.064.** doi: https://doi.org/10.3389/fmicb.2023.1073798

Botond Bendegúz Surányi, Andrea Taczman-Brückner, Csilla Mohácsi-Farkas, Tekla Engelhardt. **(2023)** Rapid identification of bacteria from agricultural environment using MALDI-TOF MS. *Acta Alimentaria: An International Journal of Food Science*, 52, 1. **Q3, IF: 1.000. doi:** https://doi.org/10.1556/066.2022.00202

10 APPENDICES

10.1 References

Abadias, M., Usall, J., Anguera, M., Solsona, C., & Viñas, I. (2008): 'Microbiology quality of fresh, minimally-processed fruit and vegetables, and sprouts from retail establishments', *International Journal of Food Microbiology*, 123 (1-2), pp. 121-129. doi: https://doi.org/10.1016/j.ijfoodmicro.2007.12.013

Abraham, W-R., Estrela, A.B., Nikitin, D.I., Smit, J., & Vancanneyt, M. (2010): '*Brevundimonas halotolerans* sp. nov., *Brevundimonas poindexterae* sp. nov. and *Brevundimonas staleyi* sp. nov., prosthecate bacteria from aquatic habitats', *International Journal of Systematic and Evolutionary Microbiology*, 60(8), pp. 1837–43. doi: https://doi.org/10.1099/ijs.0.016832-0

Acheson, D., & Allos, B.M. (2001): '*Campylobacter jejuni* Infections: update on emerging issues and trends', *Clinical Infectious Diseases*, 32(8), pp. 1201-6. doi: https://doi.org/10.1086/319760

Adewoyin, M. A., & Okoh, A.I. (2018): 'The natural environment as a reservoir of pathogenic and non-pathogenic *Acinetobacter* species', *Reviews on Environmental Health*, 33(3), pp. 265-272. doi: https://doi.org/10.1515/reveh-2017-0034

Alatoom, A.A., Cunningham, S.A., Ihde, S.M., Mandrekar, J., & Patel, R. (2011): 'Comparison of direct colony method versus extraction method for identification of gram-positive cocci by use of Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry', *Journal of Clinical Microbiology*, 49(8), pp. 2868-73. doi: https://doi.org/10.1128/JCM.00506-11

Alispahic, M., Hummel, K., Jandreski-Cvetkovic, D., Nöbauer, K., Razzazi-Fazeli, E., Hess, M., & Hess, C. (2009): 'Species-specific identification and differentiation of *Arcobacter*, *Helicobacter* and *Campylobacter* by full-spectral matrix-associated laser desorption/ionization time of flight mass spectrometry analysis', *Journal of Medical Microbiology*, 59(3), pp. 295–301. doi: https://doi.org/10.1099/jmm.0.016576-0

Alnakip, M.E., Quintela-Baluja, M., Böhme, K., Caamaño-Antelo, S., Bayoumi, M.A., Kamal, R.M., Merwad, A.M., Calo-Mata, P., & Barros-Velázquez, J. (2019): 'Molecular characterisation and typing the methicillin resistance of *Staphylococcus* spp. isolated from raw milk and cheeses in Northwest Spain: a mini survey',

International Dairy Journal, 89, pp. 68–76. doi: https://doi.org/10.1016/j.idairyj.2018.09.006

Anagnostopoulos, D.A, Parlapani, F.F., Natoudi, S., Syropoulou, F., Kyritsi, M., Vergos, I., Hadjichristodoulou, C., Kagalou, I., & Boziaris, I.S. (2022): 'Bacterial Communities and Antibiotic Resistance of Potential Pathogens Involved in Food Safety and Public Health in Fish and Water of Lake Karla, Thessaly, Greece', *Pathogens*, 11(12), 1473. doi: https://doi.org/10.3390/pathogens11121473

Anderson, N.W., Buchan, B.W., Riebe, K.M., Parsons, L.N., Gnacinski, S., & Ledeboer, N.A. (2012): 'Effects of solid-medium type on routine identification of bacterial isolates by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry', *Journal of Clinical Microbiology*, 50(3), pp. 1008-13. doi: https://doi.org/10.1128/JCM.05209-11

API Reference Guide. bioMérieux, Inc. 2019. Marcy-l'Étoile, France. PPRN 053594 01.A https://www.biomerieux-usa.com/sites/subsidiary_us/files/18_api-ref-guide_v7.pdf accessed: 05.12.2022.

Applied Biosystems (ABI). 2009. DNA sequencing by capillary electrophoresis. Chemistry guide. ABI, Thermo Fisher Scientific, Foster City, CA. https://tools.thermofisher.com/content/sfs/manuals/cms_041003.pdf accessed: 05.12.2022.

Artimová, R., Játiová, M., Baumgartnerová, J., Lipková, N., Petrová, J., Maková, J., Javoreková, S., Hleba, L., Medová, J., & Medo, J. (2023): 'Microbial Communities on Samples of Commercially Available Fresh-Consumed Leafy Vegetables and Small Berries', *Horticulturae*, 9(2), 150. doi: https://doi.org/10.3390/horticulturae9020150

Ashfaq, M.Y., Al-Ghouti, M.A., Qiblawey, H., Rodrigues, D.F., Hu, Y., & Zouari, N. (2019): 'Isolation, identification and biodiversity of antiscalant degrading seawater bacteria using MALDI-TOF-MS and multivariate analysis', *Science of The Total Environment*, 656, pp. 910–920. doi: https://doi.org/10.1016/j.scitotenv.2018.11.477.

Ashfaq, M.Y., Da'na, D.A., & Al-Ghouti, M.A. (2022): 'Application of MALDI-TOF MS for identification of environmental bacteria: A review', *Journal of Environmental Management*, 305, 114359, doi: https://doi.org/10.1016/j.jenvman.2021.114359.

Avanzi, I.R., Gracioso, L.H., dos Passos Galluzzi Baltazar, M., Karolski, B., Perpetuo, E.A., & Nascimento, C.A.O. (2017): 'Rapid bacteria identification from

environmental mining samples using MALDI-TOF MS analysis', *Environmental Science and Pollution Research*, 24(4), pp. 3717–3726. doi: https://doi.org/10.1007/s11356-016-8125-8

Babich, O., Shevchenko, M., Ivanova, S., Pavsky, V., Zimina, M., Noskova, S., Anohova, V., Chupakhin, E., & Sukhikh, S. (2021):' Antimicrobial Potential of Microorganisms Isolated from the Bottom Sediments of Lake Baikal', *Antibiotics*, 10(8), 927. doi: https://doi.org/10.3390/antibiotics10080927

Bae, D., Macoy, D.M., Ahmad, W., Peseth, S., Kim, B., Chon, J-W., Ryu, G.R., Ban, G-H., Kim, S.A., Kang, H.J., Moon, J.S., & Kim, M.G. (2022): 'Distribution and Characterization of Antimicrobial Resistant Pathogens in a Pig Farm, Slaughterhouse, Meat Processing Plant, and in Retail Stores', *Microorganisms*, 10(11), 2252. doi: https://doi.org/10.3390/microorganisms10112252

Barcelos, M.M., Martins, L., Grenfell, R.C., Juliano, L., Anderson, K.L., Dos Santos, M.V., & Gonçalves, J.L. (2019): 'Comparison of standard and on-plate extraction protocols for identification of mastitis-causing bacteria by MALDI-TOF MS', *Brazilian Journal of Microbiology*, 50(3), pp. 849–857. doi: https://doi.org/10.1007/s42770-019-00110-5

Benagli, C., Rossi, V., Dolina, M., Tonolla, M., & Petrini, O. (2011): 'Matrix-assisted laser desorption ionization-time of flight mass spectrometry for the identification of clinically relevant bacteria', *PLoS One*, 6(1), e16424 doi: https://doi.org/10.1371/journal.pone.0016424

Berger, C.N., Sodha, S.V., Shaw, R.K., Griffin, P.M., Pink, D., Hand, P., & Frankel, G. (2010): 'Fresh fruit and vegetables as vehicles for the transmission of human pathogens', *Environmental Microbiology*, 12(9), pp. 2385-2397. doi: https://doi.org/10.1111/j.1462-2920.2010.02297.x

Betts, R. (2014): 'Microbial update: fruit and salad', *International Food Hygiene*, 25(3); 9–12. doi: https://assets.thermofisher.com/TFS-Assets/LSG/brochures/International-Food-Hygiene-Vol25-No3-Fruit-and-Fresh-Produce.pdf

Bhunia, K.A. (2018): Foodborne Microbial Pathogens. p. 193-196; 209-215; 229-234; 249-260; 271-280; 289-293. Food Science Text Series. Springer, New York, NY, USA: Springer. 377. p. doi: https://doi.org/10.1007/978-1-4939-7349-1

Bi, H., Tian, Y., Song, C., Li, J., Liu, T., Chen, Z., Chen, C., Huang, Y., Zhang, Y. (2019): 'Urinary microbiota – a potential biomarker and therapeutic target for bladder cancer', *Journal of Medical Microbiology*, 68(10), pp. 1471–1478. doi: https://doi.org/10.1099/jmm.0.001058

bioMérieux SA (2022): 'VITEK® MS Rapid Identification of Microorganisms. Improve Your Productivity & Increase Confidence', https://www.biomerieux-industry.com/products/vitek-ms-rapid-identification-microorganisms-massspectrometry accessed: 06.12.2022.

Bizzini, A., Durussel, C., Bille, J., Greub, G., & Prod'hom, G. (2010): 'Performance of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of bacterial strains routinely isolated in a clinical microbiology laboratory', *Journal of Clinical Microbiology*, 48(5), pp. 1549 –1554. doi: https://doi.org/10.1128/JCM.01794-09.

Bizzini, A., Jaton, K., Romo, D., Bille, J., Prodhom, G., & Greub, G. (2011): 'Matrix-assisted laser desorption ionization-time of flight mass spectrometry as an alternative to 16S rRNA gene sequencing for identification of difficult-to-identify bacterial strains', *Journal of Clinical Microbiol*, 49(2), pp. 693–696. doi: https://doi.org/10.1128/JCM.01463-10

Böhme, K., Fernández-No, I.C., Pazos, M., Gallardo, J.M., Barros-Velázquez, J., Canas, B., & Calo-Mata, P. (2013): 'Identification and classification of seafood-borne pathogenic and spoilage bacteria: 16S rRNA sequencing versus MALDI-TOF MS fingerprinting', *Electrophoresis*, 34(6), pp. 877–887. doi: https://doi.org/10.1002/elps.201200532

Borda-Molina, D., Seifert, J., Camarinha-Silva, A. (2018): 'Current Perspectives of the Chicken Gastrointestinal Tract and Its Microbiome', *Computational and Structural Biotechnology Journal*, 16, pp. 131–139. doi: https://doi.org/10.1016/j.csbj.2018.03.002

Bruker Daltonics GmbH & Co. KG (2022): 'MALDI Biotyper® Library. The heart of the MALDI Biotyper® - enabling fast, accurate and reliable microorganism identification', https://www.bruker.com/en/products-and-solutions/microbiology-and-diagnostics/microbial-identification/maldi-biotyper-library-ruo.html accessed: 06.12.2022.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., & Holmes, S.P. (2016): 'DADA2: high-resolution sample inference from Illumina amplicon data', *Nature Methods*, 13(7), pp. 581–583. doi: https://doi.org/10.1038/nmeth.3869

Carbonnelle, E., Mesquita, C., Bille, E., Day, N., Dauphin, B., Beretti, J.L., Ferroni, A., Gutmann, L., & Nassif, X. (2011): MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. *Clinical Biochemistry*. 44(1), pp. 104-9. doi: https://doi.org/10.1016/j.clinbiochem.2010.06.017

Carvalheira, A., Casquete, R., Silva, J., & Teixeira, P. (2017): 'Prevalence and antimicrobial susceptibility of *Acinetobacter* spp. isolated from meat', *International Journal of Food Microbiology*, 243, pp. 58–63. doi: https://doi.org/10.1016/j.ijfoodmicro.2016.12.001

Cayemitte, P.E., Raymond, P., & Aider, M. (2022): '*Bacillus cereus* as an Underestimated Foodborne Pathogen and New Perspectives on its Prevalence and Methods of Control: Critical and Practical Review', *ACS Food Science & Technology*, 2(8), pp. 1196–1212. doi: https://doi.org/10.1021/acsfoodscitech.2c00173

Centers for Disease Control and Prevention (CDC) (2018): Staphylococcal (Staph) Food Poisoning. URL: https://www.cdc.gov/foodsafety/diseases/staphylococcal.html. Accessed: 12/12/2022.

Cevallos-Cevallos, J.M., Gu, G., Richardson, S.M., Hu, J., & van Bruggen, A.H. (2014): 'Survival of *Salmonella enterica* Typhimurium in water amended with manure', *Journal of Food Protection*, 77(12), pp. 2035-42. doi: 10.4315/0362-028X.JFP-13-472

Choi, J.Y., Kim, Y., Ko, E.A., Park, Y.K., Jheong, W-H., Ko, G.P., & Ko, K.S. (2012): '*Acinetobacter* species isolates from a range of environments: species survey and observations of antimicrobial resistance', *Diagnostic Microbiology and Infectious Disease*, 74(2), pp. 177–80. doi: https://doi.org/10.1016/j.diagmicrobio.2012.06.023

Chong, J., Liu, P., Zhou, G., & Xia., J. (2020): 'Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data', *Nature Protocols*, 15(3), 799–821. doi: https://doi.org/10.1038/s41596-019-0264-1

Christner, M., Rohde, H., Wolters, M., Sobottka, I., Wegscheider, K., & Aepfelbacher, M. (2010): 'Rapid identification of bacteria from positive blood culture bottles by use of matrix-assisted laser desorption-ionization time of flight mass spectrometry

fingerprinting', *Journal of Clinical Microbiology*, 48(5), pp. 1584–1591. doi: https://doi.org/10.1128/JCM.01831-09

Chung, Y., Han, M., & Kim, J.S. (2021): 'Comparative Evaluation of Bruker Biotyper and ASTA MicroIDSys for Species Identification in a Clinical Microbiology Laboratory', *Diagnostics*, 11(9), 1683. doi: https://doi.org/10.3390/diagnostics11091683

Church, D.L., Cerutti, L., Gürtler, A., Griener, T., Zelazny, A., & Emler, S. (2020): 'Performance and application of 16S rRNA gene cycle sequencing for routine identification of bacteria in the clinical microbiology laboratory', Clinical Microbiology Reviews, 33(4), e00053-19, pp. 1–74. doi: https://doi.org/10.1128/CMR.00053-19

Clarridge, J.E. (2004): 'Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases', *Clinical Microbiology Reviews*, 17(4), pp. 840 – 862. doi: https://doi.org/10.1128/CMR.17.4.840-862.2004

Clinical Laboratory Standards Institute (CLSI). 2018. Interpretive criteria for identification of bacteria and fungi by DNA target sequencing; MM18-A2 approved guideline. CLSI, Wayne, PA.

Corry, J.E.L., Curtis, G.D.W., & Baird R.M. (2003): Skirrow campylobacter selective agar. 595-596. p. *Progress in Industrial Microbiology*, vol. 37. Amsterdam, Netherlands: Elsevier, 662. p. ISSN 0079-6352. ISBN 9780444510846. doi: https://doi.org/10.1016/S0079-6352(03)80092-9

Croxatto, A., Prod'hom, G., & Greub, G. (2012): 'Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology', *FEMS Microbiology Reviews*, 36(2), pp. 380–407. doi: https://doi.org/10.1111/j.1574-6976.2011.00298.x

Cunningham, S.A., & Patel, R. (2015): 'Standard matrix-assisted laser desorption ionization-time of flight mass spectrometry reagents may inactivate potentially hazardous bacteria', *Journal of Clinical Microbiology*, 53(8), pp. 2788-2789, doi: https://doi.org/10.1128/JCM.00957-15

de Jonge, N., Carlsen, B., Christensen, M.H., Pertoldi, C., Nielsen, J.L. (2022): 'The Gut Microbiome of 54 Mammalian Species', *Frontiers in Microbiology*, 13, 886252. doi: https://doi.org/10.3389/fmicb.2022.886252

de Koster, C.G., & Brul, S. (2016): 'MALDI-TOF MS identification and tracking of food spoilers and food-borne pathogens, *Current Opinion in Food Science*, 10 , pp. 76-84, doi: https://doi.org/10.1016/j.cofs.2016.11.004.

Delaleu, J., Battistella, M., Rathana, K., Vignon-Pennamen, M.-D., Laurent, C., Ram-Wolff, C., Fenaux, P., Jachiet, M., Zuelgaray, E., Bagot, M., Bouaziz, J.-D., & Masson, A. (2021): 'Identification of clonal skin myeloid cells by next-generation sequencing in myelodysplasia cutis', *British Journal of Dermatology*, 184(2), pp. 367–369. doi: https://doi.org/10.1111/bjd.19547

Dhariwal, A., Chong, J., Habib, S., King, I., Agellon, L.B., & Xia., J. (2017): 'MicrobiomeAnalyst – a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data', *Nucleic Acids Research*, 45(1), pp. 80-188. doi: https://doi.org/10.1093/nar/gkx295

Dhiman, N., Hall, L., Wohlfiel, S.L., Buckwalter, S.P., & Wengenack, N.L. (2011): 'Performance and cost analysis of matrix-assisted laser desorption ionization-time of flight mass spectrometry for routine identification of yeast', *Journal of Clinical Microbiology*, 49(4), pp. 1614-6. doi: https://doi.org/10.1128/JCM.02381-10

Didelot, X., Bowden, R., Wilson, D.J., Peto, T.E.A., & Crook, D.W. (2012): 'Transforming clinical microbiology with bacterial genome sequencing', *Nature Reviews Genetics*, 13, pp. 601– 612. doi: https://doi.org/10.1038/nrg3226

Dijkshoorn, L., van Aken, E., Shunburne, L., van der Reijden, T.J., Bernards, A.T., Nemec, A., & Towner, K.J. (2005): 'Prevalence of *Acinetobacter baumannii* and other *Acinetobacter* spp. in faecal samples from non-hospitalised individuals', *Clinical Microbiology and Infection*, 11(4), 329–32. doi: https://doi.org/10.1111/j.1469-0691.2005.01093.x

Drake, R.R., Boggs, S.R., & Drake, S.K. (2011): 'Pathogen identification using mass spectrometry in the clinical microbiology laboratory', *Journal of Mass Spectrometry*, 46(12), pp. 1223–32. doi: https://doi.org/10.1002/jms.2008

Drevinek, M., Dresler, J., Klimentova, J., Pisa, L., Hubalek, M. (2012): 'Evaluation of sample preparation methods for MALDI-TOF MS identification of highly dangerous bacteria', *Letters in Applied Microbiology*, 55 (1), pp. 40-46. doi: https://doi.org/10.1111/j.1472-765X.2012.03255.x

Dunne, W.M., Jr., Jaillard, M., Rochas, O., & Van Belkum, A. (2017): 'Microbial genomics and antimicrobial susceptibility testing', *Expert Review of Molecular Diagnostics*, 17(3), pp. 257–269. doi: https://doi.org/10.1080/14737159.2017.1283220.

Dunne, W.M., Jr., Westblade, L.F., & Ford, B. (2012): 'Next-generation and whole genome sequencing in the diagnostic clinical microbiology laboratory', *European Journal of Clinical Microbiology & Infectious Diseases*, 31(8), pp. 1719 –1726. doi: https://doi.org/10.1007/s10096-012-1641-7

Elbehiry, A., Marzouk, E., Abdeen, E., Al-Dubaib, M., Alsayeqh, A., Ibrahem, M., Hamada, M., Alenzi, A., Moussa, I., & Hemeg, H.A. (2019): 'Proteomic characterization and discrimination of *Aeromonas* species recovered from meat and water samples with a spotlight on the antimicrobial resistance of *Aeromonas hydrophila*', *Microbiology Open*, 8(11), e782. doi: https://doi.org/10.1002/mbo3.782.

El-Nemr, I.M., Mushtaha, M., Sundararaju, S., Fontejon, C., Suleiman, M., Tang, P., Goktepe, I., & Hasan, M.R. (2019): 'Application of MALDI Biotyper System for Rapid Identification of Bacteria Isolated from a Fresh Produce Market', *Current Microbiology*, 76, pp. 290–296. doi: https://doi.org/10.1007/s00284-018-01624-1

Escherich, T. (1885): 'Die darmbakterien des neugeborenen und säuglings'. Fortschritte der Medizin, 3, pp. 515–522; 547–554.

European Food Safety Authority, European Centre for Disease Prevention and Control (EFSA and ECDC). (2019): 'The European Union One Health 2018 Zoonoses Report', *EFSA Journal*, 17(12), e05926. doi: https://doi.org/10.2903/j.efsa.2019.5926

European Food Safety Authority, European Centre for Disease Prevention and Control (EFSA and ECDC). (2021a): 'The European Union One Health 2019 Zoonoses Report', *EFSA Journal*, 19(2), e06406. doi: https://doi.org/10.2903/j.efsa.2021.6406

European Food Safety Authority, European Centre for Disease Prevention and Control. (2021b): 'The European Union One Health 2020 Zoonoses Report', *EFSA Journal*, 19(12), e6971, pp. 324. doi: https://doi.org/10.2903/j.efsa.2021.6971

Falardeau, J., Johnson, R.P., Pagotto, F., & Wang, S. (2017): 'Occurrence, characterization, and potential predictors of verotoxigenic *Escherichia coli*, *Listeria monocytogenes*, and *Salmonella* in surface water used for produce irrigation in the

Lower Mainland of British Columbia, Canada', *PloS One*, 12(9), e0185437. doi: https://doi.org/10.1371/journal.pone.0185437

Fan, W.T., Qin, T.T., Bi, R.R., Kang, H.Q., Ma, P., & Gu, B. (2017): 'Performance of the matrix-assisted laser desorption ionization time-of flight mass spectrometry system for rapid identification of streptococci: a review', *European Journal of Clinical Microbiology & Infectious Diseases*, 36(6), pp. 1005–1012. doi: https://doi.org/10.1007/s10096-016-2879-2

Faron, M.L., Buchan, B.W., Hyke, J., Madisen, N., Lillie, J.L., Granato, P.A., Wilson, D.A., Procop, G.W., Novak-Weekley, S., Marlowe, E., Cumpio, J., Griego-Fullbright, C., Kindig, S., Timm, K., Young, S., & Ledeboer, N.A. (2015): 'Multicenter evaluation of the Bruker MALDI biotyper CA system for the identification of clinical aerobic gram-negative bacterial isolates', *PLoS One*, 10(11), e0141350. doi: https://doi.org/10.1371/journal.pone.0141350

Fedorko, D.P., Drake, S.K., Stock, F., & Murray, P.R. (2012): 'Identification of clinical isolates of anaerobic bacteria using matrix-assisted laser desorption ionization-time of flight mass spectrometry', *European Journal of Clinical Microbiology & Infectious Diseases*, 31(9), pp. 2257-62. doi: https://doi.org/10.1007/s10096-012-1563-4

Feng, B., Haimei, S., Fuxing, X., Hu, F., He, J., Yang, H., Ding, C., Chen, W., & Yu, S. (2020): 'FTIR-assisted MALDI-TOF MS for the identification and typing of bacteria', *Analytica Chimica Acta*, 1111, pp. 75-82. doi: https://doi.org/10.1016/j.aca.2020.03.037

Ferreira, L., Sánchez-Juanes, F., García-Fraile, P., Rivas, R., Mateos, P.F., Martínez-Molina, E., González-Buitrago, J.M., & Velázquez, E. (2011): 'MALDI-TOF mass spectrometry is a fast and reliable platform for identification and ecological studies of species from family *Rhizobiaceae*', *PLoS ONE* 6(5), e20223. doi: https://doi.org/10.1371/journal.pone.0020223

Ferreira, L., Sánchez-Juanes, F., González-Avila, M., Cembrero-Fuciños, D., Herrero-Hernández, A., González-Buitrago, J.M., & Muñoz-Bellido, J.L. (2010): 'Direct identification of urinary tract pathogens from urine samples by matrix-assisted laser desorption ionization-time of flight mass spectrometry', *Journal of Clinical Microbiology*, 48(6), pp. 2110-5. doi: https://doi.org/10.1128/JCM.02215-09

Fessler, A., Scott, C., Kadlec, K., Ehricht, R., Monecke, S., & Schwarz, S. (2010): 'Characterization of methicillin-resistant *Staphylococcus aureus* ST398 from cases of bovine mastitis', *Journal of Antimicrobial Chemotherapy*, 65(4), pp. 619-25. doi: https://doi.org/10.1093/jac/dkq021

Fox, G.E., Wisotzkey, J.D., & Jurtshuk, P. Jr. (1992): 'How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity', *International Journal of Systematic Bacteriology*, 42(1), 166 –170. doi: https://doi.org/10.1099/00207713-42-1-166.

Fox, K., Fox, A., Rose, J., & Walla, M. (2011): 'Speciation of coagulase negative staphylococci, isolated from indoor air, using SDS page gel bands of expressed proteins followed by MALDI TOF MS and MALDI TOF-TOF MS-MS analysis of tryptic peptides', *Journal of Microbiological Methods*, 84(2), 2011, pp. 243-250. ISSN 0167-7012. doi: https://doi.org/10.1016/j.mimet.2010.12.007.

Freiwald, A., & Sauer, S. (2009): 'Phylogenetic classification and identification of bacteria by mass spectrometry', *Nature Protocols*, 4(5), pp. 732–742. doi: https://doi.org/10.1038/nprot.2009.37

Funke, G., von Graevenitz, A., Clarridge, J.E., & Bernard, K.A. (1997): 'Clinical microbiology of coryneform bacteria', *Clinical Microbiology Reviews*, 10(1), pp. 125–59. doi: https://doi.org/10.1128/CMR.10.1.125

Garner, O., Mochon, A., Branda, J., Burnham, C.A., Bythrow, M., Ferraro, M., Ginocchio, C., Jennemann, R., Manji, R., Procop, G.W., Richer, S., Rychert, J., Sercia, L., Westblade, L., & Lewinski, M. (2014): 'Multi-centre evaluation of mass spectrometric identification of anaerobic bacteria using the VITEK(R) MS system', *Clinical Microbiology Infection*, 20(4), pp. 335-339. doi: https://doi.org/10.1111/1469-0691.12317

Gharsa, H., Ben Sallem, R., Ben Slama, K., Gómez-Sanz, E., Lozano, C., Jouini, A., Klibi, N., Zarazaga, M., Boudabous, A., & Torres, C. (2012): 'High diversity of genetic lineages and virulence genes in nasal *Staphylococcus aureus* isolates from donkeys destined to food consumption in Tunisia with predominance of the ruminant associated CC133 lineage', *BMC Veterinary Research*, 8, 203. doi: https://doi.org/10.1186/1746-6148-8-203

Glenn, G.M., Flyer, D.C., Ellingsworth, L.R., Frech, S.A., Frerichs, D.M., Seid, R.C., & Yu, J. (2007): 'Transcutaneous immunization with heat-labile enterotoxin: development of a needle-free vaccine patch', *Expert Review of Vaccines*, 6(5), pp. 809–19. doi: https://doi.org/10.1586/14760584.6.5.809

Gorliczay, E., Boczonádi, I., Kiss, N.É., Tóth, F.A., Pabar, S.A., Biró, B., Kovács, L.R., & Tamás, J. (2021): 'Microbiological Effectivity Evaluation of New Poultry Farming Organic Waste Recycling', *Agriculture*, 11(7), 683. doi: https://doi.org/10.3390/agriculture11070683

Goswami, R., Mukherjee, S., Rana, V.S., Saha, D.R., Raman, R., Padhy, P.K., & Mazumder, S. (2015): 'Isolation and characterization of arsenic-resistant bacteria from contaminated water-bodies in West Bengal, India', *Geomicrobiology Journal*, 32(1), pp. 17–26. doi: https://doi.org/10.1080/01490451.2014.920938

Grada, A., & Weinbrecht, K. (2013): 'Next-generation sequencing: methodology and application', *Journal of Investigative Dermatology*, 133(8), pp. 1-4. doi: https://doi.org/10.1038/jid.2013.248

Greenblum, S., Turnbaugh, P.J., & Borenstein, E. (2012): 'Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease', *Proceedings of the National Academy of Sciences*, 109(2), pp. 594–599 doi: https://doi.org/10.1073/pnas.1116053109

Griffin, P.M., Price, G.R., Schooneveldt, J.M., Schlebusch, S., Tilse, M.H., Urbanski, T., Hamilton, B., & Venter, D. (2012): 'Use of matrix-assisted laser desorption ionization-time of flight mass spectrometry to identify vancomycin-resistant enterococci and investigate the epidemiology of an outbreak', *Journal of Clinical Microbiology*, 50(9), pp. 2918–2931. doi: https://doi.org/10.1128/JCM.01000-12

Grutsch, A.A., Nimmer, P.S., Pittsley, R.H., Kornilow, K.G., & McKillip, J.L. (2018): 'Molecular pathogenesis of *Bacillus* spp., with emphasis on the dairy industry', *Fine Focus*, 4(2), pp. 203–222. doi: https://doi.org/10.33043/FF.4.2.203-222

Gu, G., Luo, Z., Cevallos-Cevallos, J.M., Adams, P., Vellidis, G., Wright, A., & van Bruggen, A.H. (2013): 'Factors affecting the occurrence of *Escherichia coli* O157 contamination in irrigation ponds on produce farms in the Suwannee River Watershed', *Canadian Journal of Microbiology*, 59(3), pp. 175-82. Epub 2012 Dec 3. PMID: 23540335. doi: https://doi.org/10.1139/cjm-2012-0599

Guardabassi, L., Dalsgaard, A., & Olsen, J.E. (1999): 'Phenotypic characterization and antibiotic resistance of *Acinetobacter* spp. isolated from aquatic sources', *Journal of Applied Microbiology*, 87(5), pp. 659–67. doi: https://doi.org/10.1046/j.1365-2672.1999.00905.x

Gupta, V., Chandran, S., Deep, A., Kumar, R., & Bisht, L. (2022): 'Environmental factors affecting the diversity of psychrophilic microbial community in the high altitude snow-fed lake Hemkund, India', *Current Research in Microbial Sciences*, 3, 100126. doi: https://doi.org/10.1016/j.crmicr.2022.100126

Haigh, J., Degun, A., Eydmann, M., Millar, M., & Wilks, M. (2011): 'Improved Performance of Bacterium and Yeast Identification by a Commercial Matrix-Assisted Laser Desorption Ionization – Time of Flight Mass Spectrometry System in the Clinical Microbiology Laboratory', *Journal of Clinical Microbiology*, 49(9), pp. 3441 – 3441. doi: https://doi.org/10.1128/JCM.00576-11

Hall, N. (2007): 'Advanced sequencing technologies and their wider impact in microbiology', *Journal of Experimental Biology*, 210(9), pp. 1518–25. doi: https://doi.org/10.1242/jeb.001370

Hamame, A., Davoust, B., Hasnaoui, B., Mwenebitu, D.L., Rolain, J-M., & Diene, S.M. (2022): 'Screening of colistin-resistant bacteria in livestock animals from France', *Veterinary Research*, 53, 96. doi: https://doi.org/10.1186/s13567-022-01113-1

Hickman-Brenner, F.W., MacDonald, K.L., Steigerwalt, A.G., Fanning, G.R., Brenner, D.J., Farmer, J.J. (1987): '*Aeromonas veronii*, a new ornithine decarboxylase-positive species that may cause diarrhea', *Journal of Clinical Microbiology*, 25(5), pp. 900-6. doi: https://doi.org/10.1128/jcm.25.5.900-906.1987

Hoelzer, K., Moreno Switt, A.I., & Wiedmann, M. (2011): 'Animal contact as a source of human non typhoidal salmonellosis', *Veterinary Research*, 42(1), pp. 34. doi: https://doi.org/10.1186/1297-9716-42-34

Hou, T-Y., Chuan, C-N., & Teng, S-H. (2019): 'Current status of MALDI-TOF mass spectrometry in clinical microbiology', *Journal of Food and Drug Analysis*, 27(2), 2019, pp. 404-414, doi: https://doi.org/10.1016/j.jfda.2019.01.001.

Hrenovic, J., Music, M.S., Durn, G., Dekic, S., Hunjak, B. & Kisic, I. (2019): 'Carbapenem-Resistant *Acinetobacter baumannii* Recovered from Swine Manure',

Microbial Drug Resistance, 25(5), pp. 725-730. doi: http://doi.org/10.1089/mdr.2018.0087

IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp. https://ibm.com

Iliev, I., Yahubyan, G., Marhova, M., Apostolova, E., Gozmanova, M., Gecheva, G., Kostadinova, S., Ivanova, A., & Baev, V. (2017): 'Metagenomic profiling of the microbial freshwater communities in two Bulgarian reservoirs', *Journal of Basic Microbiology*, 57(8), pp. 669-679. doi: https://doi.org/10.1002/jobm.201700137

Illikoud, N., Rossero, A., Chauvet, R., Courcoux, P., Pilet, M.-F., Charrier, T., Jaffres, E., & Zagorec, M. (2019): 'Genotypic and phenotypic characterization of the food spoilage bacterium *Brochothrix thermosphacta*', *Food Microbiology*, 81, pp. 22–31. PMID: 30910085. doi: https://doi.org/10.1016/j.fm.2018.01.015.

Ivanova, E.P., Gorshkova, N.M., Sawabe, T., Hayashi, K., Kalinovskaya, N.I., Lysenko, A.M., Zhukova, N.V., Nicolau, D.V., Kuznetsova, T.A., Mikhailov, V.V., & Christen, R. (2002): '*Pseudomonas extremorientalis* sp. nov., isolated from a drinking water reservoir', *International Journal of Systematic and Evolutionary Microbiology*, 52(6), pp. 2113-20. doi: https://doi.org/10.1099/00207713-52-6-2113

Jackson, K.A., Edwards-Jones, V., Sutton, C.W., & Fox, A.J. (2005): 'Optimisation of intact cell MALDI method for fingerprinting of methicillin-resistant *Staphylococcus aureus*. Journal of microbiological methods', 62(3), pp. 273–84. doi: https://doi.org/10.1016/j.mimet.2005.04.015

Jamal, W.Y., Shahin, M., & Rotimi, V.O. (2013): 'Comparison of two matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry methods and API 20AN for identification of clinically relevant anaerobic bacteria', *Journal of Medical Microbiology*, 62(4), pp. 540–544. doi: https://doi.org/10.1099/jmm.0.053256-0.

Jancova, P., Pachlova, V., Cechov, E., Cedidlova, K., Sera, J., Pistekova, H., Bunka, F., & Bunkova, L. (2020): 'Occurrence of biogenic amines producers in the wastewater of the dairy industry', *Molecules*, 25(21), 5143. doi: https://doi.org/10.3390/molecules25215143.

Janda, J.M., & Abbott, S.L. (2007): '16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls', *Journal of Clinical Microbiology*, 45(9), pp. 2761–2764. doi: 10.1128/JCM.01228-07.

Janda, J.M., & Abott, S.L. (2010): 'The Genus *Aeromonas*: Taxonomy, Pathogenicity, and Infection', *Clinical Microbiology Reviews* 23(1), pp. 35-53. doi: https://doi.org/10.1128/CMR.00039-09

Jesser, K.J., & Noble, R.T. (2018): '*Vibrio* Ecology in the Neuse River Estuary, North Carolina, Characterized by Next-Generation Amplicon Sequencing of the Gene Encoding Heat Shock Protein 60 (hsp60)', *Applied and Environmental Microbiology*, 84(13), e00333-18. doi: https://doi.org/10.1128/AEM.00333-18

Jin, D., Kong, X., Cui, B., Jin, S., Xie, Y., Wang, X., & Deng Y. (2018): 'Bacterial communities and potential waterborne pathogens within the typical urban surface waters', *Scientific Reports*, 8, 13368. doi: https://doi.org/10.1038/s41598-018-31706-w

Johnson, J.S., Spakowicz, D.J., Hong, B.Y., Petersen, L.M., Demkowicz, P., Chen, L., Leopold, S.R., Hanson, B.M., Agresta, H.O., Gerstein, M., Sodergren, E., & Weinstock, G.M. (2019): 'Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis', *Nature Communications*, 10(1), 5029. doi: https://doi.org/10.1038/s41467-019-13036-1

Jovanovic, J., Ornelis, V.F., Madder, A., & Rajkovic, A. (2021): '*Bacillus cereus* food intoxication and toxicoinfection', *Comprehensive Reviews in Food Science and Food Safety*, 20(4), pp. 3719-3761. doi: https://doi.org/10.1111/1541-4337.12785

Kang, S-J., Choi, N-S., Choi, J.H., Lee, J-S., Yoon, J-H., & Song, J.J. (2009): '*Brevundimonas naejangsanensis* sp. nov., a proteolytic bacterium isolated from soil, and reclassification of *Mycoplana bullata* into the genus *Brevundimonas* as *Brevundimonas bullata* comb. nov.', *International Journal of Systematic Evolutionary Microbiology*, 59(12), pp. 3155–60. doi: https://doi.org/10.1099/ijs.0.011700-0

Kazazić, S.P., Topić Popović, N., Strunjak-Perović, I., Florio, D., Fioravanti, M., Babić, S., & Čož-Rakovac, R. (2019): 'Fish photobacteriosis—the importance of rapid and accurate identification of *Photobacterium damselae* subsp. *piscicida*', *Journal of Fish Diseases*, 42(8), pp. 1201-1209. doi: https://doi.org/10.1111/jfd.13022.

Kelley, S.O. (2017): 'New Technologies for Rapid Bacterial Identification and Antibiotic Resistance Profiling', *SLAS Technology*, 22(2), pp. 113-121. doi: https://doi.org/10.1177/2211068216680207

Kim, M., Oh, H.S., Park, S.C., & Chun, J. (2014): 'Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes', *International Journal of Systematic Evolutionary Microbiology*, 64(2), pp. 346–351. doi: https://doi.org/10.1099/ijs.0.059774-0

Kitzinger, K., Koch, H., Lücker, S., Sedlacek, C.J., Herbold, C., Schwarz, J., Daebeler, A., Mueller, A.J., Lukumbuzya, M., Romano, S., Leisch, N., Karst, S.M., Kirkegaard, R., Albertsen, M., Nielsen, P.H., Wagner, M., & Daims, H. (2018): 'Characterization of the First "*Candidatus Nitrotoga*" Isolate Reveals Metabolic Versatility and Separate Evolution of Widespread Nitrite-Oxidizing Bacteria', *mBio*, 9(4), e01186-18. doi: https://doi.org/10.1128/mBio.01186-18

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F.O. (2013): 'Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies', *Nucleic Acids Research*, 41(1), pp. 1–11. doi: https://doi.org/10.1093/nar/gks808

Kopcakova, A., Stramova, Z., Kvasnova, S., Godany, A., Perhacova, Z., & Pristas, P. (2014): 'Need for database extension for reliable identification of bacteria from extreme environments using MALDI TOF mass spectrometry', *Chemical Papers*, 68(11), pp. 1435–1442. doi: https://doi.org/10.2478/s11696-014-0612-0

Laukova, A., Strompfova, V., Scerbova, J., Pogany, S., & Monika, P.S. (2019): 'Virulence factor genes incidence among enterococci from sewage sludge in eastern Slovakia following safety aspect', *BioMed Research International*, 2019, 5. doi: https://doi.org/10.1155/2019/2735895

Lazarevic, V., Whiteson, K., Hernandez, D., François, P., & Schrenzel, J. (2010): 'Study of inter- and intra-individual variations in the salivary microbiota', *BMC Genomics*, 11, 523. doi: https://doi.org/10.1186/1471-2164-11-523.

Lehosmaa, K., Muotka, T., Pirttilä A.M., Jaakola, I., Rossi, P.M., & Jyväsjärvi, J. (2021): 'Bacterial communities at a groundwater-surface water ecotone: gradual change or abrupt transition points along a contamination gradient?', *Environmental Microbiology*, 23(11), pp. 6694-6706. doi: https://doi.org/10.1111/1462-2920.15708

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., & Law, M. (2012): 'Comparison of Next-Generation Sequencing Systems', *Journal of Biomedicine and Biotechnology*, 2012, 251364. doi: https://doi.org/10.1155/2012/251364

Locey, K.J., & Lennon, J.T. (2016): 'Scaling laws predict global microbial diversity', *Proceedings of the National Academy of Sciences*, 113(21), pp. 5970–5975. doi: https://doi.org/10.1073/pnas.1521291113

López-Fernández, H., Santos, H.M., Capelo, J.L., Fdez-Riverola, F., Glez-Peña, D., & Reboiro-Jato, M. (2015): 'Mass-Up: an all-in-one open software application for MALDI-TOF mass spectrometry knowledge discovery', *BMC Bioinformatics*, 5(16), 318. doi: https://doi.org/10.1186/s12859-015-0752-4

Lu, W.J., Lin, H.J., Hsu, P.H., & Lin, H.V. (2020): 'Determination of Drug Efflux Pump Efficiency in Drug-Resistant Bacteria Using MALDI-TOF MS', *Antibiotics*, 9(10), 639. doi: https://doi.org/10.3390/antibiotics9100639

Luque-Sastre, L., Arroyo, C., Fox, E.M., McMahon, B.J., Bai, L., Li, F., Fanning, S. Antimicrobial Resistance in Bacteria from Livestock and Companion Animals. 2018. p. 237-239. Antimicrobial Resistance in *Listeria* Species. ASM Press. Washington, D.C. p. 237-259. doi: https://doi.org/10.1128/9781555819804.ch11

Martin, M. (2011): 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet. journal*, 17(1), 10 doi: https://doi.org/10.14806/ej.17.1.200

Martiny, D., Busson, L., Wybo, I., El Haj, R.A., Dediste, A., & Vandenberg, O. (2012): 'Comparison of the Microflex LT and Vitek MS systems for routine identification of bacteria by matrix-assisted laser desorption ionization-time of flight mass spectrometry', *Journal of Clinical Microbiology*, 50(4), pp. 1313–1325. doi: https://doi.org/10.1128/JCM.05971-11

Masood, S., Zhao, X.Q., & Shen, R.F. (2020): '*Bacillus pumilus* promotes the growth and nitrogen uptake of tomato plants under nitrogen fertilization', *Scientia Horticulturae*, 272, 109581. doi: https://doi.org/10.1016/j.scienta.2020.109581.

Matereke, L.T., & Okoh, A.I. (2020): '*Listeria monocytogenes* Virulence, Antimicrobial Resistance and Environmental Persistence: A Review', *Pathogens*, 9(7), 528, doi: https://doi.org/10.3390/pathogens9070528

Matsuda, N., Matsuda, M., Notake, S., Yokokawa, H., Kawamura, Y., Hiramatsu, K., & Kikuchi, K. (2012): 'Evaluation of a simple protein extraction method for species identification of clinically relevant staphylococci by matrix-assisted laser desorption ionization-time of flight mass spectrometry', *Journal of Clinical Microbiology*, 50(12), pp. 3862-6. doi: https://doi.org/10.1128/JCM.01512-12

Mauchline, T.H., & Malone, J.G. (2017): 'Life in earth - the root microbiome to the rescue?', *Current Opinion in Microbiology*, 37, pp. 23-28. doi: https://doi.org/10.1016/j.mib.2017.03.005

McDermott, P.F., Zhao, S., & Heather, T. Antimicrobial Resistance in Bacteria from Livestock and Companion Animals. 2018. p. 261-262. Antimicrobial Resistance in Nontyphoidal *Salmonella*. ASM Press. Washington, D.C. p. 261-287. doi: https://doi.org/10.1128/9781555819804.ch12

McElvania TeKippe, E., Shuey, S., Winkler, D.W., Butler, M.A., & Burnham, C.A.D. (2013): 'Optimizing identification of clinically relevant Gram-positive organisms by use of the Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry system', *Journal of Clinical Microbiology*, 51(5), pp. 1421-1427. doi: https://doi.org/10.1128/jcm.02680-12

Molina-Lopez, R.A., Valverdú, N., Martin, M., Mateu, E., Obon, E., Cerdà-Cuéllar, M., & Darwich, L. (2011): 'Wild raptors as carriers of antimicrobial-resistant *Salmonella* and *Campylobacter* strains', *Veterinary Record*, 168(21), pp. 565-565. doi: https://doi.org/10.1136/vr.c7123

Murugaiyan, J., Walther, B., Abou-Elnaga, Y., Brueggemann-Schwarze, S., Vincze, S., Wieler, L.H., & Lübke-Becker, A., Semmler, T., & Roesle, U. (2014): 'Species differentiation within the *Staphylococcus intermedius* group using a refined MALDI-TOF MS database', *Clinical Microbiology and Infection*, 20(10), pp. 1007–1015. doi: https://doi.org/10.1111/1469-0691.12662

Nakamura, S., Yang, C-S., Sakon, N., Ueda, M., Tougan, T., Yamashita, A., Goto, N., Takahashi, K., Yasunaga, T., Ikuta, K., Mizutani, T., Okamoto, Y., Tagami, M., Morita, R., Maeda, N., Kawai, J., Hayashizaki, Y., Nagai, Y., Horii, T., Iida, T., Nakaya, T. (2009): 'Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach', *PLoS One*, 4(1):e4219. doi: https://doi.org/10.1371/journal.pone.0004219

Nakano, S., Matsumura, Y., Kato, K., Yunoki, T., Hotta, G., Noguchi, T., Masaki Yamamoto, M., Nagao, M., Ito, Y., Takakura, S., & Ichiyama, S. (2014): 'Differentiation of *vanA*-positive *Enterococcus faecium* from *vanA*-negative *E. faecium* by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry', *International Journal of Antimicrobial Agents*, 44(3), pp. 256–259. doi: https://doi.org/10.1016/j.ijantimicag.2014.05.006

Nguyen-The, C., & Prunier, J.P. (1989): 'Involvement of pseudomonads in deterioration or 'ready-to-use' salads', *International Journal of Food Science & Technology*, 24(1), pp. 47-58. doi: https://doi.org/10.1111/j.1365-2621.1989.tb00618.x

Pál, T. (2013): Orvosi mikrobiológia. Budapest: Medicina Könyvkiadó Zrt. 303 p.

Pandey, A., Jain, R., Sharma, A., Dhakar, K., Kaira, G.S., Rahi, P., Dhyani, A., Pandey, N., Adhikari, P., & Shouche, Y.S. (2019): '16S rRNA gene sequencing and MALDI-TOF mass spectrometry based comparative assessment and bioprospection of psychrotolerant bacteria isolated from high altitudes under mountain ecosystem', *SN Applied Sciences*, 1, 278. doi: https://doi.org/10.1007/s42452-019-0273-2

Panek, M., Čipčić Paljetak, H., Barešić, A., Perić, M., Matijašić, M., Lojkić, I., Bender, D.V., Krznarić, Ž. & Verbanac, D. (2018): 'Methodology challenges in studying human gut microbiota – effects of collection, storage, DNA extraction and next generation sequencing technologies', *Scientific Reports*, 8, 5143. doi: https://doi.org/10.1038/s41598-018-23296-4

Pascale, M.R., Mazzotta, M., Salaris, S., Girolamini, L., Grottola, A., Simone, M.L., Cordovana, M., Bisognin, F., Dal Monte, P., Bucci Sabattini, M.A., Viggiani, M., & Cristino, S. (2020): 'Evaluation of MALDI-TOF mass spectrometry in diagnostic and environmental surveillance of *Legionella* species: a comparison with culture and Mip-gene sequencing technique', *Frontiers in Microbiology*, 11, 589369. doi: https://doi.org/10.3389/fmicb.2020.589369.

Pathogen Surveillance Tool: Centers for Disease Control and Prevention (CDC) (2021): FoodNet Fast: Pathogen Surveillance Tool. Atlanta, Georgia: U.S. Department of Health and Human Services. URL: http://wwwn.cdc.gov/foodnetfast. Accessed: 12/12/2022.

Patz, S., Witzel, K., Scherwinski, A-C., & Ruppel, S. (2019) 'Culture Dependent and Independent Analysis of Potential Probiotic Bacterial Genera and Species Present in the Phyllosphere of Raw Eaten Produce', *Int. J. Mol. Sci.*, 20(15), 3661. doi: https://doi.org/10.3390/ijms20153661

Pennanec, X., Dufour, A., Haras, D., & Réhel, K. (2010): 'A quick and easy method to identify bacteria by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry', *Rapid Communications in Mass Spectrometry*, 24(3), pp. 384-392. doi: https://doi.org/10.1002/rcm.4404

Pinar-Méndez, A., Galofré, B., Blanch, A.R., & García-Aljaro, C. (2022): 'Culture and molecular methods as complementary tools for water quality management', *Sci. Total Environ.*, 848, 157789, doi: https://doi.org/10.1016/j.scitotenv.2022.157789

Pinel, I.S.M., Moed, D.H., Vrouwenvelder, J.S., & van Loosdrecht, M.C.M. (2020): 'Bacterial community dynamics and disinfection impact in cooling water systems', *Water Research*, 172, 115505. doi: https://doi.org/10.1016/j.watres.2020.115505

Poirel, L., Madec, J-Y., Lupo, A., Schink, A., Kieffer, N., Nordmann, P., & Schwarz, S. Antimicrobial Resistance in Bacteria from Livestock and Companion Animals. 2018. p. 289-316. Antimicrobial Resistance in *Escherichia coli*. ASM Press. Washington, D.C. p. 289-291. doi: https://doi.org/10.1128/9781555819804.ch7

Pomastowski, P., Złoch, M., Rodzik, A., Ligor, M., Kostrzewa, M., & Buszewski, B. (2019): 'Analysis of bacteria associated with honeys of different geographical and botanical origin using two different identification approaches: MALDI-TOF MS and 16S rDNA PCR technique', *PLoS One*, 14(5), e0217078. doi: https://doi.org/10.1371/journal.pone.0217078.

Ponderand, L., Pavese, P., Maubon, D., Giraudon, E., Girard, T., Landelle, C., Maurin, M., & Caspar, Y. (2020): 'Evaluation of Rapid Sepsityper® protocol and specific MBT-Sepsityper module (Bruker Daltonics) for the rapid diagnosis of bacteremia and fungemia by MALDI-TOF-MS', *Annals of Clinical Microbiology and Antimicrobials*, 19(1), 60. doi: https://doi.org/10.1186/s12941-020-00403-w

Quail, M.A., Gu, Y., Swerdlow, H., & Mayho, M. (2012): 'Evaluation and optimisation of preparative semi-automated electrophoresis systems for Illumina library preparation', *Electrophoresis*, 33(23), pp. 3521–28. doi: https://doi.org/10.1002/elps.201200128

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F.O. (2013): 'The SILVA ribosomal RNA gene database project: improved data processing and web-based tools', *Nucleic Acids Research*, 41(D1), pp. 590-596. doi: https://doi.org/10.1093/nar/gks1219

R Core Team. (2021): R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. 2021. https://www.r-project.org.

Rafei, R., Hamze, M., Pailhoriès, H., Eveillard, M., Marsollier, L., Joly-Guillou M.L., Dabboussi, F., & Kempf, M. (2015): 'Extrahuman epidemiology of *Acinetobacter baumannii* in Lebanon', *Applied and Environmental Microbiology*, 81(7), pp. 2359–67. doi: https://doi.org/10.1128/AEM.03824-14

Rahi, P., Prakash, O., & Shouche, Y.S. (2016): 'Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass-Spectrometry (MALDI-TOF MS) Based Microbial Identifications: Challenges and Scopes for Microbial Ecologists', *Frontiers in Microbiology*, 7, 1359. https://doi.org/10.3389/fmicb.2016.01359

Reimer, L.C., Carbasse, J.S., Koblitz, J., Ebeling, C., Podstawka, A., & Overmann, J. (2022): 'BacDive in 2022: the knowledge base for standardized bacterial and archaeal data', *Nucleic Acids Research*, 50(D1), pp. D741-46. doi: https://doi.org/10.1093/nar/gkab961

Risch, M., Radjenovic, D., Han, J., Wydler, M., & Nydegger, U. (2010): 'Comparison of MALDI TOF with conventional identification of clinically relevant bacteria', *Swiss Medical Weekly*, 140(3738), p. w13095. doi: https://doi.org/10.4414/smw.2010.13095

Roseline, T.L. & Sachindra, N.M. (2016): 'Characterization of extracellular agarase production by *Acinetobacter junii* PS12B, isolated from marine sediments', *Biocatalysis and Agricultural Biotechnology*, 6(4), pp. 219–26. doi: https://doi.org/10.1016/j.bcab.2016.04.007

Ruan, X., Crupper, S.S., Schultz, B.D., Robertson, D.C., & Zhang, W. (2012): '*Escherichia coli* expressing EAST1 toxin did not cause an increase of cAMP or cGMP levels in cells, and no diarrhea in 5-day old gnotobiotic pigs', *PLoS One*, 7(8), e43203. doi: https://doi.org/10.1371/journal.pone.0043203

Ryan, M.P., & Pembroke, J.T. (2018): '*Brevundimonas* spp: Emerging global opportunistic pathogens', *Virulence*, 9(1), pp. 480-493, doi: https://doi.org/10.1080/21505594.2017.1419116

Rychert, J., Burnham, C.A., Bythrow, M., Garner, O.B., Ginocchio, C.C., Jennemann, R., Lewinski, M.A., Manji, R., Mochon, A.B., Procop, G.W., Richter, S.S., Sercia, L., Westblade, L.F., Ferraro, M.J., Branda, J.A. (2013): 'Multicenter evaluation of the Vitek MS matrix-assisted laser desorption ionization-time of flight mass spectrometry system for identification of Gram-positive aerobic bacteria', *Journal of Clinical Microbiology*, 51(7), pp. 2225-31. doi: https://doi.org/10.1128/JCM.00682-13

Sala-Comorera, L., Blanch, A.R., Vilaro, C., Galofre, B., & García-Aljaro, C. (2017): 'Heterotrophic monitoring at a drinking water treatment plant by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry after different drinking water treatments', *Journal of Water & Health*, 15(6), pp. 885–897. doi: https://doi.org/10.2166/wh.2017.090.

Sanad, Y.M., Closs Jr., G., Kumar, A., Jeffrey T. LeJeune, & Rajashekara, G. (2013): 'Molecular Epidemiology and Public Health Relevance of *Campylobacter* Isolated from Dairy Cattle and European Starlings in Ohio, USA', *Foodborne Pathogens and Disease*, 10(3), pp. 229-236. doi: https://doi.org/10.1089/fpd.2012.1293

Sánchez-Juanes, F., Ferreira, L., dela Vega, P.A., Valverde, A., Barrios, M.L., Rivas, R., Mateos, P.F., Martínez-Molina, E., González-Buitrago, J.M., Trujillo, M.E., & Velázquez, E. (2013): 'MALDI-TOF mass spectrometry as a tool for differentiation of *Bradyrhizobium* species: application to the identification of *Lupinus* nodulating strains', *Systematic and Applied Microbiology*, 36(8), pp. 565–571. doi: https://doi.org/10.1016/j.syapm.2013.09.003

Sanger, F., Nicklen, S., & Coulson, A.R. (1977): 'DNA sequencing with chain terminating inhibitors', *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp. 5463-7. doi: https://doi.org/10.1073/pnas.74.12.5463.

Santos, T., Capelo, J.L., Santos, H.M., Oliveira, I., Marinho, C., Gonçalves, A., Araújo, J.E., Poeta, P., & Igrejas, G. (2015): 'Use of MALDI-TOF mass spectrometry fingerprinting to characterize *Enterococcus* spp. and *Escherichia coli* isolates', *Journal of Proteomics*, 127(B), pp. 321-331. doi: https://doi.org/10.1016/j.jprot.2015.02.017

Santos, T.S., Campos, F.B., Padovani, N.F.A., Dias, M., Mendes, M.A., & Maffei, D.F. (2020): 'Assessment of the microbiological quality and safety of minimally processed vegetables sold in Piracicaba, SP, Brazil', *Lett. Appl. Microbiol*., 71(2), pp. 187–194, doi: https://doi.org/10.1111/lam.13305

Saunders, A.M., Albertsen, M., Vollertsen, J., & Nielsen, P.H. (2016): 'The activated sludge ecosystem contains a core community of abundant organisms', *The ISME journal*, 10(1), pp. 11-20. doi: https://doi.org/10.1038/ismej.2015.117

Schloss, P.D. & Handelsman, J. (2005): 'Metagenomics for studying unculturable microorganisms: cutting the Gordian knot', *Genome Biology*, 6(8), 229. doi: https://doi.org/10.1186/gb-2005-6-8-229

Schulthess, B., Bloemberg, G.V., Zbinden, A., Mouttet, F., Zbinden, R., Böttger, E.C., & Hombach, M. (2016): 'Evaluation of the Bruker MALDI Biotyper for Identification of Fastidious Gram-Negative Rods', *Journal of Clinical Microbiology*, 54(3), pp. 543-548. doi: https://doi.org/10.1128/JCM.03107-15

Schulthess, B., Bloemberg, G.V., Zbinden, R., Böttger, E.C., & Hombach, M. (2014): 'Evaluation of the Bruker MALDI Biotyper for Identification of Gram-Positive Rods: Development of a Diagnostic Algorithm for the Clinical Laboratory', *Journal of Clinical Microbiology*, 52(4), pp. 1089-1097. doi: https://doi.org/10.1128/JCM.02399-13

Schwarz, S., Feßler, A.T., Loncaric, I., Wu, C., Kadlec, K., Wang, Y., & Shen, J. Antimicrobial Resistance in Bacteria from Livestock and Companion Animals. 2018. p. 127-157. Antimicrobial Resistance among Staphylococci of Animal Origin. ASM Press. Washington, D.C. p. 127-128. doi: https://doi.org/10.1128/9781555819804.ch7

Scott, J.S., Sterling, S.A., To, H., Seals, S.R., & Jones, A.E. (2016): 'Diagnostic performance of matrix-assisted laser desorption ionization time-of-flight mass spectrometry in blood bacterial infections: a systematic review and meta-analysis', *Infectious Diseases*, 48(7), pp. 530–536. doi: https://doi.org/10.3109/23744235.2016.1165350

Seng, P., Abat, C., Rolain, J. M., Colson, P., Lagier, J. C., Gouriet, F., Fournier, P.E., Drancourt, M., La Scola, B., & Raoult, D. (2013): 'Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of matrix-assisted laser

desorption ionization-time of flight mass spectrometry. *Journal of Clinical Microbiology*, 51(7), pp. 2182–2194. doi: https://doi.org/10.1128/JCM.00492-13

Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P.E., Rolain, J.M., & Raoult, D. (2009): 'Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry', *Clinical Infectious Diseases*, 49(4), pp. 543-51. doi: https://doi.org/10.1086/600885

Shen, Z., Wang, Y., Zhang, Q., & Shen, J. Antimicrobial Resistance in Bacteria from Livestock and Companion Animals. 2018. p. 317-318. Antimicrobial Resistance in *Campylobacter* spp. ASM Press. Washington, D.C. pp. 317-330. doi: https://doi.org/10.1128/9781555819804.ch1

Shih, C.J., Chen, S.C., Weng, C.Y., Lai, M.C., & Yang, Y.L. (2015): 'Rapid identification of haloarchaea and methanoarchaea using the matrix assisted laser desorption/ionization time-of-flight mass spectrometry', *Scientific Reports*, 5:16326. doi: https://doi.org/10.1038/srep16326

Singhal, N., Kumar, M., Kanaujia, P.K., & Virdi, J.S. (2015): 'MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis', *Frontiers in Microbiology*, 6, 791. doi: https://doi.org/10.3389/fmicb.2015.00791

Sirichoat, A., Sankuntaw, N., Engchanil, C., Buppasiri, P., Faksri, K., Namwat, W., Chantratita, W., & Lulitanond, V. (2021): 'Comparison of different hypervariable regions of 16S rRNA for taxonomic profiling of vaginal microbiota using next-generation sequencing', *Archives of Microbiology*, 203, pp. 1159–1166. doi: https://doi.org/10.1007/s00203-020-02114-4

Skarp, C.P.A., Hänninen, M.-L., Rautelin, H.I.K. (2016): 'Campylobacteriosis: the role of poultry meat', *Clinical Microbiology and Infection*, 22(2), pp. 103-109. doi: https://doi.org/10.1016/j.cmi.2015.11.019.

Solomon, E.B., Yaron, S., & Matthews, K.R. (2002): 'Transmission of *Escherichia coli* O157: H7 from Contaminated Manure and Irrigation Water to Lettuce Plant Tissue and Its Subsequent Internalization', *Applied and Environmental Microbiology*, 68(1), pp. 397-400. doi: https://doi.org/10.1128/AEM.68.1.397-400.2002

Strejcek, M., Smrhova, T., Junkova P., & Uhlik, O. (2018): 'Whole-Cell MALDI-TOF MS Versus 16S rRNA Gene Analysis for Identification and Dereplication of Recurrent

Bacterial Isolates', *Frontiers in Microbiology*, 9, 1294. doi: https://doi.org/10.3389/fmicb.2018.01294

Sun, L.N, Zhang, J., Chen, Q., He, J., Li, Q.F, & Li, S.P. (2013): '*Comamonas jiangduensis* sp. nov., a biosurfactant-producing bacterium isolated from agricultural soil', *International Journal of Systematic and Evolutionary Microbiology*, 63(6), pp. 2168–73. doi: https://doi.org/10.1099/ijs.0.045716-0

Sunantaraporn, S., Sanprasert, V., Pengsakul, T., Phumee, A., Boonserm, R., Tawatsin, A., Thavara, U., & Siriyasatien, P. (2015): 'Molecular survey of the head louse *Pediculushumanus capitis* in Thailand and its potential role for transmitting Acinetobacter spp.', *Parasites & Vectors*, 8, 127. doi: https://doi.org/10.1186/s13071-015-0742-4

Suzuki, Y., Niina, K., Matsuwaki, T., Nukazawa, K., & Iguchi, A. (2018): 'Bacterial flora analysis of coliforms in sewage, river water, and ground water using MALDI-TOF mass spectrometry', *Journal of Environmental Science and Health*, 53(2), pp. 160-173. doi: https://doi.org/10.1080/10934529.2017.1383128

Tang, W., Ranganathan, N., Shahrezaei, V., & Larrouy-Maumus, G. (2019): 'MALDI-TOF mass spectrometry on intact bacteria combined with a refined analysis framework allows accurate classification of MSSA and MRSA', *PLoS ONE*, 14, e0218951. doi: https://doi.org/10.1371/journal.pone.0218951

Tekebayeva, Z., Temirbekova, A., Bazarkhankyzy, A., Bissenova, G., Abzhalelov, A., Tynybayeva, I., Temirkhanov, A., Askarova, N., Mkilima, T., & Sarmurzina, Z. (2023): 'Selection of Active Microorganism Strains Isolated from a Naturally Salty Lake for the Investigation of Different Microbial Potentials', *Sustainability*, 15(1), 51. doi: https://doi.org/10.3390/su15010051

Thompson, C., Garcia, G., Masi, B.P., Freitas, T., Paz, P.H.C., Leal, C.V., Otsuki, K., Tschoeke, D., Salazar, V., Soares, M., Lopes, G., Bacha, L., Cosenza, C., Vieira, V.V., Botelho, A.C.N., Vaz de Oliveira, B.C., de Rezende, C.E., Teixeira, L., & Thompson, F. (2023): 'Brumadinho dam collapse induces changes in the microbiome and the antibiotic resistance of the Paraopeba River (Minas Gerais, Brazil)', *Science of The Total Environment*, 865, 161278. doi: https://doi.org/10.1016/j.scitotenv.2022.161278.

Todar, K. (2012): 'Todar's Online Textbook of Bacteriology - The Normal Bacterial Flora of Humans', https://www.textbookofbacteriology.net/normalflora_3.html, accessed: 15.12.2022.

Topić Popović, N., Kazazić, S.P., Bojanić, K., Strunjak-Perović, I., & Čož-Rakovac, R. (2021): 'Sample preparation and culture condition effects on MALDI-TOF MS identification of bacteria: A review', *Mass Spectrometry Reviews*, pp. 1-15. doi: https://doi.org/10.1002/mas.21739

Topić Popović, N., Kepec, S., Kazazić, S.P., Strunjak-Perović, I., Bojanić, K., & Čož-Rakovac, R. (2022): 'Identification of environmental aquatic bacteria by mass spectrometry supported by biochemical differentiation', *PLoS One,* 17(6), e0269423. doi: https://doi.org/10.1371/journal.pone.0269423

Tsuchida, S., Umemura, H., & Nakayama, T. (2020): 'Current status of matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) in clinical diagnostic microbiology', *Molecules*, 25(20); 4775. doi: https://doi.org/10.3390/molecules25204775

Turner, K., Moua, C.N., Hajmeer, M., Barnes, A., & Needham, M. (2019): 'Overview of Leafy Greens–Related Food Safety Incidents with a California Link: 1996 to 2016', *Journal of Food Protection,* 82(3), pp. 405–414. doi: https://doi.org/10.4315/0362-028X.JFP-18-316

Turvey, M.E., Weiland, F., Meneses, J., Sterenberg, N., & Hoffmann, P. (2016): 'Identification of beer spoilage microorganisms using the MALDI Biotyper platform', *Applied Microbiology Biotechnology*, 100(6), pp. 2761–2773. doi: https://doi.org/10.1007/s00253-016-7344-8

Uchida-Fuji, E., Niwa, H., Kinoshita, Y., & Nukada, T. (2020): 'Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry (MALDI-TOF MS) for Identification of Bacterial Isolates from Horses', *Journal of Equine Veterinary Science*, 95, 103276. doi: https://doi.org/10.1016/j.jevs.2020.103276

Uyttendaele, M., Jaykus, L-A., Amoah, P., Chiodini, A., Cunliffe, D., Jacxsens, L., Holvoet, K., Korsten, L., Lau, M., McClure, P., Medema, G., Sampers, I., & Jasti, P.R. (2015): 'Microbial Hazards in Irrigation Water: Standards, Norms, and Testing to Manage Use of Water in Fresh Produce Primary Production', *Comprehensive Reviews in Food Science*, 14(4), pp. 336-356. doi: https://doi.org/10.1111/1541-4337.12133

van Veen, S.Q., Claas, E.C., Kuijper, E.J. (2010): 'High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories', *Journal of Clinical Microbiology*, 48(3), pp. 900-7. doi: https://doi.org/10.1128/JCM.02071-09

Veloo, A.C., Elgersma, P.E., Friedrich, A.W., Nagy, E., & van Winkelhoff, A.J. (2014): 'The influence of incubation time, sample preparation and exposure to oxygen on the quality of the MALDI-TOF MS spectrum of anaerobic bacteria', *Clinical Microbiology and Infection*, 20(12):O1091-7. doi: https://doi.org/10.1111/1469-0691.12644

Virtanen, S., Rantsi, T., Virtanen, A., K., Nieminen, P., Kalliala, I., & Salonen, A. (2019): 'Vaginal Microbiota Composition Correlates Between Pap Smear Microscopy and Next Generation Sequencing and Associates to Socioeconomic Status', *Scientific Reports*, 9, 7750. doi: https://doi.org/10.1038/s41598-019-44157-8

Wallis, Y., Morrell, N. (2011): Automated DNA Sequencing. 173-185. p. In: Theophilus, B., Rapley, R. (Eds.): PCR Mutation Detection Protocols. *Methods in Molecular Biology*, vol 688. Totowa, New Jersey, USA: Humana Press. doi: https://doi.org/10.1007/978-1-60761-947-5_12

Wang, J., Wang, H., Cai, K., Yu, P., Liu, Y., Zhao, G., Chen, R., Xu, R., & Yu, M. (2021): 'Evaluation of three sample preparation methods for the identification of clinical strains by using two MALDI-TOF MS systems', *Journal of Mass Spectrometry*, 56(2), e4696. doi: https://doi.org/10.1002/jms.4696

Warneke, H.L., Kinyon, J.M., Bower, L.P. ,Burrough, E.R., & Frana, T.S. (2014): 'Matrix-assisted laser desorption ionization time-of-flight mass spectrometry for rapid identification of *Brachyspira* species isolated from swine, including the newly described "*Brachyspirahampsonii*"', *Journal of Veterinary Diagnostic Investigation*, 26(5), 635–639. doi: https://doi.org/10.1177/1040638714541114

Weidmann-Al-Ahmad, M., Tichy, H.V., & Schön, G. (1994): 'Characterisation of *Acinetobacter* type strains and isolates obtained from wastewater treatment plants by PCR fingerprinting', *Applied and Environmental Microbiology*, 60(11), pp. 4066–71. doi: https://doi.org/10.1128/aem.60.11.4066-4071.1994

Welker, M., & Moore, E.R. (2011): 'Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology',

Systematic and Applied Microbiology, 34(1), pp. 2-11. doi: https://doi.org/10.1016/j.syapm.2010.11.013

Wendlandt, S., Feßler, A.T., Monecke, S., Ehricht, R., Schwarz, S., & Kadlec, K. (2013): 'The diversity of antimicrobial resistance genes among staphylococci of animal origin', *International Journal of Medical Microbiology*, 303(6–7), pp. 338-349. doi: https://doi.org/10.1016/j.ijmm.2013.02.006.

Wendlandt, S., Kadlec, K., Feßler, A.T., Monecke, S., Ehricht, R., van de Giessen, A.W., Hengeveld, P.D., Huijsdens, X., Schwarz, S., & van Duijkeren, E. (2019): 'Resistance phenotypes and genotypes of methicillin-resistant *Staphylococcus aureus* isolates from broiler chickens at slaughter and abattoir workers', *Journal of Antimicrobial Chemotherapy*, 68(11), pp. 2458-63. doi: https://doi.org/10.3382/ps/pez439

Wieczorek, K., & Osek, J. (2013): 'Antimicrobial resistance mechanisms among *Campylobacter*', *BioMed Research International*, 2013, 340605. doi: https://doi.org/10.1155/2013/340605

Willner, D., Furlan, M., Haynes, M., Schmieder, R., Angly, F.E., Silva, J., Tammadoni, S., Nosrat, B., Conrad, D., Rohwer, F. (2009): 'Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals', *PLoS One* 4(10), e7370, doi: https://doi.org/10.1371/journal.pone.0007370

Woese, C.R., Fox, G.E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B.J., Stahl, D. (1975): 'Conservation of primary structure in 16S ribosomal RNA', *Nature*, 254(5495), pp. 83– 86. doi: https://doi.org/10.1038/254083a0

Wu, H., Li, Y., Zhang, W., Wang, C., Wang, P., Niu, L., Du, J., & Gao, Y. (2019): 'Bacterial community composition and function shift with the aggravation of water quality in a heavily polluted river', *Journal of Environmental Management*, 237, pp. 433-441. doi: https://doi.org/10.1016/j.jenvman.2019.02.101.

Wybo, I., Soetens, O., DeBel, A., Echahidi, F., Vancutsem, E., Vandoorslaer, K., & Piérard, D. (2012): 'Species identification of clinical *Prevotella* isolates by matrix-assisted laser desorption ionization-time of flight mass spectrometry', *Journal of Clinical Microbiology*, 50(4), pp. 1415–1418. doi: https://doi.org/10.1128/JCM.06326-11

Xu, L., Wu, Y-H., Jian, S-L., Wang, C-H., Wu, M., Cheng, L., & Xu, X-W. (2016): '*Pseudohongiella nitratireducens* sp. nov., isolated from seawater, and emended description of the genus *Pseudohongiella*', *International Journal of Systematic and Evolutionary Microbiology*, 66(12), pp. 5155–516. doi: https://doi.org/10.1099/ijsem.0.001489

Yoon, E-J, & Jeong, S.H. (2021): 'MALDI-TOF Mass Spectrometry Technology as a Tool for the Rapid Diagnosis of Antimicrobial Resistance in Bacteria', *Antibiotics*, 10(8), 982. doi: https://doi.org/10.3390/antibiotics10080982

Yoon, S.H., Ha, S.M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017): 'Introducing EzBioCloud: A taxonomically united database of 16S rRNA and whole genome assemblies', *International Journal of Systematic and Evolutionary Microbiology*, 67(5), pp. 1613-1617. doi: https://doi.org/10.1099/ijsem.0.001755

Zancarini, A., Echenique-Subiabre, I., Debroas, D., Taïb, N., Quiblier, C., & Humbert, J.F. (2017): 'Deciphering biodiversity and interactions between bacteria and microeukaryotes within epilithic biofilms from the Loue River, France', *Scientific Reports*, 7(1), 4344. doi: https://doi.org/10.1038/s41598-017-04016-w

Zhao, Y., Mao, X., Zhang, M., Yang, W., Di, H.J., Ma, L., Liu, W., & Li, B. (2021): 'The application of *Bacillus Megaterium* alters soil microbial community composition, bioavailability of soil phosphorus and potassium, and cucumber growth in the plastic shed system of North China', *Agriculture, Ecosystems & Environment*, 7, 107236. doi: https://doi.org/10.1016/j.agee.2020.107236.

Żukowska, M.E. (2021): 'Advanced Methods of Bacteriological Identification in Clinical Microbiology Laboratory', *Journal of Pre-Clinical and Clinical Research*, 15(2), pp. 68–72. doi: https://doi.org/10.26444/jpccr/134646

10.2 Appendix of the supplementary tables

Appendix Table 1. List of the identified environmental isolates included in the thesis

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
1	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	*P. anguilliseptica* (1.87)	NA
2	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	*A. salmonicida* (2.12)	NA
3	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	*A. bestiarum*/*A. salmonicida* (2.1; 2.08)	NA
4	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	*A. salmonicida* (2.25)	NA
5	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	*A. salmonicida* (2.18)	NA
6	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	*A. salmonicida*/*A. bestiarum* (2.09; 2)	NA
7	Lake1	Kavicsos-tó	Szigetszentmiklós (Central HU)	No id. (1.52)	NA
8	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Chryseobacterium* spp. (1.85)	NA
9	Lake2	Szelidi-tó	Dunapataj (Central HU)	*P. mandelii* (2.08)	NA
10	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Shewanella baltica* (2.03)	NA
11	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Bacillus* spp. (1.97)	NA
12	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Bacillus* spp. (1.78)	NA
13	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Aeromonas* spp. (1.86)	NA
14	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Bacillus* spp. (1.85)	NA
15	Lake2	Szelidi-tó	Dunapataj (Central HU)	*A. salmonicida* (2.09)	NA
16	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Bacillus* spp. (1.99)	NA
17	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Pseudomonas* spp. (1.77)	NA
18	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Pseudomonas* spp. (1.76)	NA
19	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Pseudomonas* spp. (1.72)	NA
20	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Bacillus* spp. (1.74)	NA
21	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Rheinheimera* spp. (1.74)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
22	Lake2	Szelidi-tó	Dunapataj (Central HU)	No id. (1.58)	NA
23	Lake2	Szelidi-tó	Dunapataj (Central HU)	*Bacillus* spp. (1.92)	NA
24	River1	Tisza	Tiszakécske (Eastern HU)	*A. veronii* (2.3)	NA
25	River1	Tisza	Tiszakécske (Eastern HU)	*Rhodococcus* spp. (1.93)	NA
26	River1	Tisza	Tiszakécske (Eastern HU)	No id. (1.5)	NA
27	River1	Tisza	Tiszakécske (Eastern HU)	*Aeromonas* spp. (1.99)	NA
28	River1	Tisza	Tiszakécske (Eastern HU)	*Ralstonia* spp. (1.88)	NA
29	River1	Tisza	Tiszakécske (Eastern HU)	*B. cereus* (2.1)	NA
30	River1	Tisza	Tiszakécske (Eastern HU)	*Bacillus* spp. (1.93)	NA
31	River1	Tisza	Tiszakécske (Eastern HU)	*Aeromonas* spp. (1.97)	NA
32	River1	Tisza	Tiszakécske (Eastern HU)	*Bacillus* spp. (1.77)	NA
33	River1	Tisza	Tiszakécske (Eastern HU)	*Bacillus* spp. (1.92)	NA
34	River1	Tisza	Tiszakécske (Eastern HU)	*A. ichthiosmia*/ *A. hydrophila* (2.13; 2.03)	NA
35	River1	Tisza	Tiszakécske (Eastern HU)	*Bacillus cereus* (2.1)	NA
36	River1	Tisza	Tiszakécske (Eastern HU)	*Bacillus* spp. (1.7)	NA
37	River1	Tisza	Tiszakécske (Eastern HU)	No id. (1.65)	NA
38	River1	Tisza	Tiszakécske (Eastern HU)	*A. hydrophila* (2.1)	NA
39	River1	Tisza	Tiszakécske (Eastern HU)	No id. (1.3)	NA
40	River1	Tisza	Tiszakécske (Eastern HU)	*Aeromonas* spp. (1.7)	NA
41	River1	Tisza	Tiszakécske (Eastern HU)	No id. (1.36)	NA
42	River1	Tisza	Tiszakécske (Eastern HU)	No id. (1.39)	NA
43	River1	Tisza	Tiszakécske (Eastern HU)	*P. alcaligenes* (2.1)	NA
44	River2	Tisza	Szolnok (Eastern HU)	*Pseudomonas* spp. (1.91)	NA
45	River2	Tisza	Szolnok (Eastern HU)	*Pseudomonas* spp. (2)	NA
46	River2	Tisza	Szolnok (Eastern HU)	*Janthinobacterium* spp. (1.98)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
47	River2	Tisza	Szolnok (Eastern HU)	*Pseudomonas* spp. (1.75)	NA
48	River2	Tisza	Szolnok (Eastern HU)	*Flavobacterium pectinovorum* (2.13)	NA
49	River2	Tisza	Szolnok (Eastern HU)	*P. brenneri* (2.34)	NA
50	River2	Tisza	Szolnok (Eastern HU)	*Pseudomonas* spp. (1.72)	NA
51	River2	Tisza	Szolnok (Eastern HU)	*P. fragi* (2.14)	NA
52	River2	Tisza	Szolnok (Eastern HU)	*Pseudomonas* spp. (1.82)	NA
53	River2	Tisza	Szolnok (Eastern HU)	*Janthinobacterium* spp. (1.83)	
54	River2	Tisza	Szolnok (Eastern HU)	No id. (1.57)	NA
55	River2	Tisza	Szolnok (Eastern HU)	No id. (1.69)	NA
56	River2	Tisza	Szolnok (Eastern HU)	No id. (1.53)	NA
57	River3	Danube	Csepel (Central HU)	*Pseudomonas* spp. (1.79)	NA
58	River3	Danube	Csepel (Central HU)	*Pseudomonas* spp. (1.96)	NA
59	River3	Danube	Csepel (Central HU)	*P. extremorientalis* (2.3)	NA
60	River3	Danube	Csepel (Central HU)	*Bacillus* spp. (1.96)	NA
61	River3	Danube	Csepel (Central HU)	*P. flourescencs* (2.28)	NA
62	River3	Danube	Csepel (Central HU)	*Pseudomonas* spp. (1.91)	NA
63	River3	Danube	Csepel (Central HU)	No id. (1.41)	NA
64	River3	Danube	Csepel (Central HU)	No id. (1.39)	NA
65	River3	Danube	Csepel (Central HU)	*P. mandelii* (2.15)	NA
66	River3	Danube	Csepel (Central HU)	*Bacillus* spp. (1.91)	NA
67	River3	Danube	Csepel (Central HU)	*Aeromonas* spp. (1.93)	NA
68	River3	Danube	Csepel (Central HU)	*Aeromonas caviae* (2.2)	NA
69	River3	Danube	Csepel (Central HU)	*Bacillus* spp. (1.71)	NA
70	River3	Danube	Csepel (Central HU)	*Aeromonas* spp. (1.91)	NA
71	River3	Danube	Csepel (Central HU)	No id. (1.41)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
72	River3	Danube	Csepel (Central HU)	No id. (1.6)	NA
-73	River3	Danube	Csepel (Central HU)	*Pseudomonas* spp. (1.72)	NA
74	River3	Danube	Csepel (Central HU)	No id. (1.45)	NA
75	River3	Danube	Csepel (Central HU)	*P. fragi* (2.22)	NA
76	River3	Danube	Csepel (Central HU)	*Aeromonas* spp. (1.86)	NA
77	River3	Danube	Csepel (Central HU)	*Flavobacterium aquatile* (2.07)	NA
78	River3	Danube	Csepel (Central HU)	*Flavobacterium* spp. (1.95)	NA
79	River3	Danube	Csepel (Central HU)	*Flavobacterium aquatile* (2.22)	NA
80	River3	Danube	Csepel (Central HU)	*P. marginalis* (2)	NA
81	River3	Danube	Csepel (Central HU)	No id. (1.5)	NA
82	River3	Danube	Csepel (Central HU)	No id. (1.56)	NA
83	River3	Danube	Csepel (Central HU)	*Flavobacterium* spp. (1.81)	NA
84	River3	Danube	Csepel (Central HU)	*Micrococcus* spp. (1.93)	NA
85	River3	Danube	Csepel (Central HU)	No id. (1.5)	NA
86	River3	Danube	Csepel (Central HU)	No id. (1.55)	NA
87	River3	Danube	Csepel (Central HU)	*Janthinobacterium* spp. (1.73)	NA
88	River3	Danube	Csepel (Central HU)	*Kocuria rosea* (2.24)	NA
89	River4	Danube	Kalocsa (Southern HU)	No id. (1.25)	NA
90	River4	Danube	Kalocsa (Southern HU)	*P. flourescens* (2.19)	NA
91	River4	Danube	Kalocsa (Southern HU)	*P. frederiksbergensis* (2.01)	NA
92	River4	Danube	Kalocsa (Southern HU)	*Pseudomonas* spp. (1.95)	NA
93	River4	Danube	Kalocsa (Southern HU)	*P. antarctica*/*P. marginalis* (2.19/2.14)	NA
94	River4	Danube	Kalocsa (Southern HU)	*Janthinobacterium* spp. (1.93)	NA
95	River4	Danube	Kalocsa (Southern HU)	*Acinetobacter johnsonii* (2.16)	NA
96	River4	Danube	Kalocsa (Southern HU)	*Acinetobacter johnsonii* (2.11)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
97	River4	Danube	Kalocsa (Southern HU)	*P. extremorientalis* (2.27)	NA
98	River4	Danube	Kalocsa (Southern HU)	*P. extremorientalis* (2.12)	NA
99	River5	Vajas	Bátya (Southern HU)	*Aeromonas* spp. (1.7)	NA
100	River5	Vajas	Bátya (Southern HU)	*Aeromonas* spp. (1.98)	NA
101	River5	Vajas	Bátya (Southern HU)	*Acinetobacter* spp. (1.76)	NA
102	River5	Vajas	Bátya (Southern HU)	*Aeromonas veronii* (2.15)	NA
103	River5	Vajas	Bátya (Southern HU)	*Aeromonas* spp. (1.84)	NA
104	River5	Vajas	Bátya (Southern HU)	*P. extremorientalis* (2.25)	NA
105	River5	Vajas	Bátya (Southern HU)	*Aeromonas* spp. (1.96)	NA
106	River5	Vajas	Bátya (Southern HU)	*Acinetobacter* spp. (1.83)	NA
107	River5	Vajas	Bátya (Southern HU)	*P. marginalis* (2.26)	NA
108	River5	Vajas	Bátya (Southern HU)	*Acinetobacter* spp. (1.95)	NA
109	River5	Vajas	Bátya (Southern HU)	*Acinetobacter* spp. (1.77)	NA
110	River5	Vajas	Bátya (Southern HU)	*Pseudomonas* spp. (1.71)	NA
111	River5	Vajas	Bátya (Southern HU)	*Pseudomonas* spp. (1.84)	NA
112	River5	Vajas	Bátya (Southern HU)	No id. (1.54)	NA
113	River5	Vajas	Bátya (Southern HU)	No id. (1.35)	NA
114	Irrigation water1	Soroksár	Soroksár (Central HU)	No id. (1.46)	NA
115	Irrigation water1	Soroksár	Soroksár (Central HU)	No id. (1.49)	NA
116	Irrigation water1	Soroksár	Soroksár (Central HU)	No id. (1.36)	NA
117	Irrigation water1	Soroksár	Soroksár (Central HU)	*Pseudomonas* spp. (1.92)	NA
118	Irrigation water1	Soroksár	Soroksár (Central HU)	No id. (1.64)	NA
119	Irrigation water1	Soroksár	Soroksár (Central HU)	*Pseudomonas* spp. (1.79)	NA
120	Irrigation water1	Soroksár	Soroksár (Central HU)	*Flavobacterium* spp. (1.7)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
121	Irrigation water1	Soroksár	Soroksár (Central HU)	*Staphylococcus* spp. (1.99)	NA
122	Irrigation water1	Soroksár	Soroksár (Central HU)	No id. (1.45)	NA
123	Irrigation water1	Soroksár	Soroksár (Central HU)	No id. (1.43)	NA
124	Irrigation water1	Soroksár	Soroksár (Central HU)	*Micrococcus* spp. (1.86)	NA
125	Irrigation water1	Soroksár	Soroksár (Central HU)	No id. (1.57)	NA
126	Irrigation water1	Soroksár	Soroksár (Central HU)	*Micrococcus* spp. (1.98)	NA
127	Irrigation water1	Soroksár	Soroksár (Central HU)	*Moraxella* spp. (1.78)	NA
128	Irrigation water1	Soroksár	Soroksár (Central HU)	*Stenotrophomonas* spp. (1.96)	NA
129	Irrigation water1	Soroksár	Soroksár (Central HU)	*Paenibacillus amylolyticus* (2.1)	NA
130	Irrigation water2	Debrecen	Debrecen (Eastern HU)	*P. vulgaris* (2.1)	NA
131	Irrigation water2	Debrecen	Debrecen (Eastern HU)	No id. (1.6)	NA
132	Irrigation water2	Debrecen	Debrecen (Eastern HU)	*Pantoea* spp. (1.95)	NA
133	Irrigation water2	Debrecen	Debrecen (Eastern HU)	No id. (1.38)	NA
134	Irrigation water2	Debrecen	Debrecen (Eastern HU)	*Acinetobacter pittii* (2.32)	NA
135	Irrigation water2	Debrecen	Debrecen (Eastern HU)	No id. (1.6)	NA
136	Irrigation water2	Debrecen	Debrecen (Eastern HU)	No id. (1.65)	NA
137	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	*Acinetobacter pittii* (2.27)	NA
138	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	*Microbacterium testaceum* (2.1)	NA
139	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	*Acinetobacter pittii* (2.3)	NA
140	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	*Providencia rettgeri* (2.28)	NA
141	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	*Brevundimonas vesicularis* (2.27)	NA
142	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	No id. (1.65)	NA
143	Irrigation water3	Nagykunsági-főcsatorna	Abádszalók (Eastern HU)	No id. (1.68)	NA
144	Irrigation water4	Karcag	Karcag (Eastern HU)	*E. cloacae*/ *E. hormaechei* (2.25)	*E. cloacae*/ *E. hormaechei* (99.8%)
145	Irrigation water4	Karcag	Karcag (Eastern HU)		

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
				E. cloacae/ *E. hormaechei* (2.25)	*E. cloacae*/ *E. hormaechei* (99.23%)
146	Irrigation water4	Karcag	Karcag (Eastern HU)	*P. stutzeri* (2.26)	*P. stutzeri* (100%)
147	Irrigation water4	Karcag	Karcag (Eastern HU)	*E. hormaechei* (2.25)	*E. cloacae*/ *E. hormaechei* (99.9%)
148	Irrigation water4	Karcag	Karcag (Eastern HU)	*Pseudarthrobacter scleromae*/ *P. oxydans* (2.24)	*Pseudarthrobacter siccitolerans* (89.91%)
149	Irrigation water4	Karcag	Karcag (Eastern HU)	*Rhodococcus* spp. (1.99)	*R. cerastii* (99.46%)
150	Irrigation water4	Karcag	Karcag (Eastern HU)	*E. cloacae* (2.27)	*E. hormaechei* (99.48%)
151	Irrigation water4	Karcag	Karcag (Eastern HU)	*P. veronii* (2.26)	*P. veronii*/ *P. extremaustralis* (100%)
152	Irrigation water4	Karcag	Karcag (Eastern HU)	*P. veronii* (2.22)	*Pseudomonas* spp. (99.34%)
153	Irrigation water5	Kengyel	Kengyel (Eastern HU)	*A. ursingii* (2.17)	*A. ursingii* (90.41%)
154	Irrigation water5	Kengyel	Kengyel (Eastern HU)	*P. monteili*/*P. putida* (2.1)	*P. monteili*/ *P. putida* (99.31%)
155	Irrigation water5	Kengyel	Kengyel (Eastern HU)	*B. vesicularis* (2.02)	*B. vesicularis* (99.27%)
156	Irrigation water5	Kengyel	Kengyel (Eastern HU)	*Acinetobacter* spp. (1.81)	*A. junii* (98.5%)
157	Irrigation water5	Kengyel	Kengyel (Eastern HU)	*A. junii* (2.46)	*A. junii* (99.11%)
158	Irrigation water5	Kengyel	Kengyel (Eastern HU)	*A. ursingii* (2.11)	*A. ursingii* (91.67%)
159	Irrigation water5	Kengyel	Kengyel (Eastern HU)	*A. junii* (2.37)	*A. junii* (99.51%)
160	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*Rhodococcus* spp. (1.71)	*Rhodococcus qinsenghii* (96.2%)
161	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*B. vesicularis* (2.3)	*B. vesicularis* (99.17%)
162	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	No id. (1.51)	*Sphingobacterium kitahiroshimense* (99.72%)
163	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*C. indologenes* (2.01)	*C. lactis* (98.8%)
164	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*B. vesicularis* (2.15)	*B. vesicularis* (99.36%)
165	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*Microbacterium maritypicum* (2.29)	*M. marytipicum* (99.82%)

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
166	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*Stenotrophomonas maltophilia* (2.1)	*Stenotrophomonas maltophilia* (94.14%)
167	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*Pantoea* spp. (1.85)	*P. agglomerans* (97.38%)
168	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*E. hormaechei*/ *E. cloacae* (2.36; 2.29)	*E. hormaechei*/ *E. cloacae* (98.92%)
169	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*Aeromonas veronii* (2.17)	*A. veronii* (99.61%)
170	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*A. schindleri* (1.87)	*A. schindleri* (99.19%)
171	Irrigation water6	Rákóczifalva	Rákóczifalva (Eastern HU)	*A. schindleri* (1.85)	*A. schindleri* (99.61%)
172	Irrigation water7	Szolnok1	Szolnok (Eastern HU)	*Rhodococcus erythropolis* (2.04)	NA
173	Irrigation water7	Szolnok1	Szolnok (Eastern HU)	*Delftia acidovorans* (1.95)	NA
174	Irrigation water7	Szolnok1	Szolnok (Eastern HU)	*Delftia acidovorans* (2.03)	NA
175	Irrigation water7	Szolnok1	Szolnok (Eastern HU)	No id. (1.6)	NA
176	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.47)	*A. junii* (99.41%)
177	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.14)	*A. junii* (99.65%)
178	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.23)	*A. junii* (99.32%)
179	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.31)	*A. junii* (98.93%)
180	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.15)	*A. junii* (99.63%)
181	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.2)	*A. junii* (99.64%)
182	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.31)	*A. junii* (99.69%)
183	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	No id. (1.69)	*P. stutzeri* (99.64%)
184	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.34)	*A. schindleri* (99.24%)
185	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*Acinetobacter* spp. (1.98)	*A. junii* (98.09%)
186	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.42)	*A. junii* (99.18%)
187	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.1)	*A. schindleri* (98.78%)

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
188	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.25)	*A. junii* (99.45%)
189	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.33)	*A. junii* (99.82%)
190	Irrigation water8	Szolnok2	Szolnok (Eastern HU)	*A. junii* (2.34)	NA
191	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	*Pantoea agglomerans* (2.15)	NA
192	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	No id. (1.37)	NA
193	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	*Bacillus* spp. (1.91)	NA
194	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	*Providencia rettgeri* (2.32)	NA
195	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	*Brevundimonas diminuta* (2.18)	NA
196	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	No id. (1.61)	NA
197	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	No id. (1.37)	NA
198	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	*Proteus vulgaris* (2.6)	NA
199	Manure1	Békéscsaba	Békéscsaba (Eastern HU)	*Bacillus* spp. (1.9)	NA
200	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.11)	NA
201	Manure2	Bátya	Bátya (Southern HU)	No id. (1.66)	NA
202	Manure2	Bátya	Bátya (Southern HU)	*Comamonas jiangduensis* (2.07)	NA
203	Manure2	Bátya	Bátya (Southern HU)	*Psychrobacter* spp. (2)	NA
204	Manure2	Bátya	Bátya (Southern HU)	*Staphylococcus equorum* (2.01)	NA
205	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.26)	NA
206	Manure2	Bátya	Bátya (Southern HU)	*Alcaligenes* spp. (1.94)	NA
207	Manure2	Bátya	Bátya (Southern HU)	*Acinetobacter* spp. (1.75)	NA
208	Manure2	Bátya	Bátya (Southern HU)	*Vagococcus* spp. (1.88)	NA
209	Manure2	Bátya	Bátya (Southern HU)	*Vagococcus fluvialis* (2.19)	NA
210	Manure2	Bátya	Bátya (Southern HU)	*Comamonas jiangduensis* (2.14)	NA
211	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.49)	NA
212	Manure2	Bátya	Bátya (Southern HU)	No id. (1.67)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
213	Manure2	Bátya	Bátya (Southern HU)	*Pseudomonas* spp. (1.82)	NA
214	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.47)	NA
215	Manure2	Bátya	Bátya (Southern HU)	*Comamonas* spp. (1.71)	NA
216	Manure2	Bátya	Bátya (Southern HU)	*Acinetobacter* spp. (1.82)	NA
217	Manure2	Bátya	Bátya (Southern HU)	No id. (1.32)	NA
218	Manure2	Bátya	Bátya (Southern HU)	No id. (1.39)	NA
219	Manure2	Bátya	Bátya (Southern HU)	No id. (1.69)	NA
220	Manure2	Bátya	Bátya (Southern HU)	*Paenochrobactr um* spp. (1.86)	NA
221	Manure2	Bátya	Bátya (Southern HU)	No id. (1.52)	NA
222	Manure2	Bátya	Bátya (Southern HU)	No id. (1.39)	NA
223	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.19)	NA
224	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.46)	NA
225	Manure2	Bátya	Bátya (Southern HU)	*Comamonas* spp. (1.78)	NA
226	Manure2	Bátya	Bátya (Southern HU)	*Acinetobacter* spp. (1.77)	NA
227	Manure2	Bátya	Bátya (Southern HU)	No id. (1.32)	NA
228	Manure2	Bátya	Bátya (Southern HU)	No id. (1.35)	NA
229	Manure2	Bátya	Bátya (Southern HU)	No id. (1.42)	NA
230	Manure2	Bátya	Bátya (Southern HU)	No id. (1.35)	NA
231	Manure2	Bátya	Bátya (Southern HU)	No id. (1.35)	NA
232	Manure2	Bátya	Bátya (Southern HU)	*Bacillus* (1.85)	NA
233	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.37)	NA
234	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.51)	NA
235	Manure2	Bátya	Bátya (Southern HU)	*E. coli* (2.4)	NA
236	Manure2	Bátya	Bátya (Southern HU)	No id. (1.67)	NA
237	Manure2	Bátya	Bátya (Southern HU)	No id. (1.33)	NA
238	Manure2	Bátya	Bátya (Southern HU)	No id. (1.53)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
239	Manure2	Bátya	Bátya (Southern HU)	No id. (1.61)	NA
240	Manure2	Bátya	Bátya (Southern HU)	*Alcaligenes faecalis* (2.33)	NA
241	Manure2	Bátya	Bátya (Southern HU)	*Comamonas* spp. (1.71)	NA
242	Manure2	Bátya	Bátya (Southern HU)	No id. (1.63)	NA
243	Manure2	Bátya	Bátya (Southern HU)	No id. (1.69)	NA
244	Manure2	Bátya	Bátya (Southern HU)	*Comamonas* spp. (1.79)	NA
245	Manure2	Bátya	Bátya (Southern HU)	No id. (1.43)	NA
246	Manure2	Bátya	Bátya (Southern HU)	*Bacillus pumilus* (2.02)	NA
247	Manure2	Bátya	Bátya (Southern HU)	No id. (1.61)	NA
248	Manure3	Bátya	Bátya (Southern HU)	No id. (1.42)	NA
249	Manure3	Bátya	Bátya (Southern HU)	No id. (1.48)	NA
250	Manure3	Bátya	Bátya (Southern HU)	No id. (1.58)	NA
251	Manure3	Bátya	Bátya (Southern HU)	*Pseudomonas* spp. (1.76)	NA
252	Manure3	Bátya	Bátya (Southern HU)	*P. extremorientalis* (2.18)	NA
253	Manure3	Bátya	Bátya (Southern HU)	*P. extremorientalis* (2.08)	NA
254	Manure3	Bátya	Bátya (Southern HU)	*P. extremorientalis*/ *P. fluorescens* (2.03/2.01)	NA
255	Manure3	Bátya	Bátya (Southern HU)	*Pseudomonas* spp. (1.89)	NA
256	Manure3	Bátya	Bátya (Southern HU)	*Pseudomonas* spp. (1.75)	NA
257	Manure3	Bátya	Bátya (Southern HU)	*Pseudomonas* spp. (1.99)	NA
258	Manure3	Bátya	Bátya (Southern HU)	*P. extremorientalis*/ *P. fluorescens* (2.11/2.03)	NA
259	Manure3	Bátya	Bátya (Southern HU)	*Glutamicibacter protophormiae* (2.29)	NA
260	Manure3	Bátya	Bátya (Southern HU)	*Pseudomonas antarctica* (2.05)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
261	Manure3	Bátya	Bátya (Southern HU)	*Corynebacterium xerosis* (2.27)	NA
262	Manure3	Bátya	Bátya (Southern HU)	*Pseudomonas* spp. (1.84)	NA
263	Manure3	Bátya	Bátya (Southern HU)	*Glutamicibacter* spp. (1.89)	NA
264	Manure3	Bátya	Bátya (Southern HU)	*Glutamicibacter* spp. (1.91)	NA
265	Manure3	Bátya	Bátya (Southern HU)	*Pseudomonas* spp. (1.9)	NA
266	Manure3	Bátya	Bátya (Southern HU)	*Glutamicibacter* spp. (1.72)	NA
267	Manure3	Bátya	Bátya (Southern HU)	*Acinetobacter* spp. (1.77)	NA
268	Manure3	Bátya	Bátya (Southern HU)	*Glutamicibacter arilaitensis* (2.34)	NA
269	Manure4	Cegléd	Cegléd (Eastern HU)	*Staphylococcus* spp. (1.96)	NA
270	Manure4	Cegléd	Cegléd (Eastern HU)	*Staphylococcus* spp. (1.96)	NA
271	Manure4	Cegléd	Cegléd (Eastern HU)	No id. (1.63)	NA
272	Manure4	Cegléd	Cegléd (Eastern HU)	*Streptococcus alactolyticus* (2.08)	NA
273	Manure4	Cegléd	Cegléd (Eastern HU)	*Corynebacterium* spp. (1.91)	NA
274	Manure4	Cegléd	Cegléd (Eastern HU)	*Corynebacterium* spp. (1.73)	NA
275	Manure4	Cegléd	Cegléd (Eastern HU)	No id. (1.38)	NA
276	Manure4	Cegléd	Cegléd (Eastern HU)	No id. (1.37)	NA
277	Manure4	Cegléd	Cegléd (Eastern HU)	No id. (1.65)	NA
278	Manure4	Cegléd	Cegléd (Eastern HU)	No id. (1.4)	NA
279	Vegetables 1	Soroksár	Soroksár (Central HU)	*Chryseobacterium arthrosphaerae* (2.1)	NA
280	Vegetables 1	Soroksár	Soroksár (Central HU)	*Bacillus* spp. (1.84)	NA
281	Vegetables 1	Soroksár	Soroksár (Central HU)	No id. (1.56)	NA
282	Vegetables 1	Soroksár	Soroksár (Central HU)	*Enterobacter cloacae* (2.36)	NA
283	Vegetables 1	Soroksár	Soroksár (Central HU)	*Bacillus* spp. (1.72)	NA
284	Vegetables 1	Soroksár	Soroksár (Central HU)	*Chryseobacterium*	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
				arthrosphaerae (2.2)	
285	Vegetables 1	Soroksár	Soroksár (Central HU)	*Bacillus licheniformis* (2.01)	NA
286	Vegetables 1	Soroksár	Soroksár (Central HU)	No id. (1.51)	NA
287	Vegetables 1	Soroksár	Soroksár (Central HU)	*Bacillus cereus* (2.4)	NA
288	Vegetables 1	Soroksár	Soroksár (Central HU)	*Rahnella* spp. (1.9)	NA
289	Vegetables 1	Soroksár	Soroksár (Central HU)	*Curtobacterium flaccumfaciens* (2)	NA
290	Vegetables 1	Soroksár	Soroksár (Central HU)	No id. (1.62)	NA
291	Vegetables 1	Soroksár	Soroksár (Central HU)	*Rahnella aquatilis* (2)	NA
292	Vegetables 1	Soroksár	Soroksár (Central HU)	*Curtobacterium* spp. (1.82)	NA
293	Vegetables 2	Debrecen	Debrecen (Eastern HU)	No id. (1.47)	NA
294	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus* spp. (1.82)	NA
295	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Enterobacter cloacae* (2.1)	NA
296	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Chryseobacterium arthrosphaerae* (2)	NA
297	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Pantoea agglomerans* (2.26)	NA
298	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Proteus vulgaris* (2.2)	NA
299	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus megaterium* (2.24)	NA
300	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus subtilis* (1.89)	NA
301	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus* spp. (1.81)	NA
302	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus* spp. (1.96)	NA
303	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Acinetobacter calcoaceticus* (2.36)	NA
304	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Staphylococcus saprophyticus* (2.02)	NA
305	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Kocuria rosea* (2.35)	NA

# of isolate	Sample name	Origin of Samples	Location (city, region)	MALDI-TOF MS identification	16rRNA gene sequencing
306	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus* spp. (1.7)	NA
307	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus* spp. (1.86)	NA
308	Vegetables 2	Debrecen	Debrecen (Eastern HU)	No id. (1.65)	NA
309	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Pantoea agglomerans* (2.4)	NA
310	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Microbacterium* spp. (1.72)	NA
311	Vegetables 2	Debrecen	Debrecen (Eastern HU)	*Bacillus cereus* (2.24)	NA

11 ACKNOWLEDGEMENTS

First, I owe special thanks to my family who helped me in every way to be able to write this thesis. I would also like to thank the help, support and encouragement of Anna Jánosity.

I would like to thank the expertise, help, criticism and guidance of my Supervisors, Dr. Csilla Mohácsi-Farkas and Dr. Tekla Engelhardt.

I am also grateful to the colleagues at the Hungarian University of Agriculture and Life Sciences, Department of Food Microbiology, Hygiene and Safety. Especially to Dr. Andrea Taczman-Brückner and Dr. Gabriella Kiskó who were always there to help me out in the laboratory.

I am also sincerely grateful for Dr. Konrad Domig for accepting me at his laboratory in the University of Natural Resources and Life Sciences (BOKU), Vienna and for his guidance in the Department of Food Science and Technology. I would like to thank Dr. Benjamin Zwirzitz for his advice and guidance in molecular biology.

I would like to also thank to Dr. Márta Ladányi whose statistical courses helped me to interpret my data.

I acknowledge the Doctoral School of Food Science of the Hungarian University of Agriculture and Life Sciences for supporting this PhD research.

I also acknowledge the Department of Food Science and Technology (DLWT), University of Natural Resources and Life Sciences, Vienna for providing the bacterial strains (*E. coli* ATCC 13706 and *E. coli* DSM 11250) for the experiments presented in the subchapter of 5.4. of the Result section. The experiments presented in the subchapter of 5.5. and 5.6 of Results section were performed at University of Natural Resources and Life Sciences, Vienna (BOKU), Department of Food Science and Technology, Institute of Food Science. The MALDI-TOF MS (Bruker MALDI Biotyper) instrument was provided by the EQ-BOKU-VIBT GmbH.

My work was supported by the CEEPUS (grant numbers: CIII-HR-0306-12-1920-M-139456 - For Safe and Healthy Food in Middle-Europe, M-HR-0306-14-2122-M-150096 – For Safe and Healthy Food in Middle-Europe) and by the EFOP – The Project is supported by the European Union and co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00005).