Targeting the isolation of rumen microbes involved in trimethylamine metabolism using MALDI-TOF MS.

A thesis presented in partial fulfilment of the

requirements for the degree of

Master of Science

in Microbiology

at Massey University, Manawatū,

New Zealand

Priya Soni

2025

Abstract

Trimethylamine (TMA, $N(CH_3)_3$) is produced in gut microbial systems, such as the human gut and the rumen, through the microbial degradation of dietary substrates like choline, betaine, and carnitine. The cultivation of microbes involved in TMA metabolism is of mutual interest for both human health and rumen fermentation studies. TMA is a precursor to trimethylamine N-oxide (TMAO), a metabolite correlated with disease, and TMA is also a precursor for methane (CH_4) production in ruminants. Microbes that carry genes encoding TMA metabolism have been identified from gut microbial metagenomic sequence data. However, confirming TMA metabolism in individual microbes requires experimental validation using pure cultures. Using genes known to be involved in TMA metabolism, a list of candidate TMA-metabolising strains was identified from bacterial isolates within the Hungate1000 culture collection. The study aimed to isolate new microbes involved in TMA metabolism using the Bruker Biotyper® Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry (MALDI-TOF MS) system as a high-throughput screening and identification tool. The MALDI Biotyper® database was expanded by creating mass spectral profiles (MSPs) from rumen-derived candidate strains implicated in TMA metabolism, to assist with the rapid identification of new isolates. Rumen and faecal samples were selectively enriched for microorganisms able to grow on the TMA precursors choline, betaine, and carnitine, and TMA production was confirmed in the choline enrichments. However, purified isolates in this study were unable to produce TMA. Isolated pure cultures were rapidly screened with the MALDI Biotyper® using the expanded targeted database, which resulted in identifications requiring further inclusion of MSPs from the rumen source. MALDI identifications were verified using 16S rRNA gene sequencing of the isolates.

Keywords: TMA, TMAO, gut microbes, human health, rumen, methane, MALDI-TOF Biotyper®.

Acknowledgements

I am grateful for the opportunity to have received supervision from Dr Sara Burgess, Dr Elizabeth Rettedal, Dr Christina Moon, and Dr Graeme Attwood. The consistent guidance from all my supervisors throughout this long journey has enriched my research experience. Dr Sara Burgess, your expert guidance in learning and developing skills with MALDI has been crucial in overcoming the project's challenges. I also appreciate the chance to attend your writing workshops, which have been the most enjoyable learning experiences. Dr Elizabeth Rettedal, I am especially intrigued by your idea for this multi-disciplinary research project that explores new approaches to research. I want to thank Dr Christina Moon for considering my studentship request and making it possible for me to undertake this project. Finally, I hugely acknowledge Dr Graeme Attwood for collaborating on this project, providing access to the Hungate1000 cultures, calf rumen samples for research, and ongoing guidance on understanding TMA metabolism investigated in this study. I am grateful to the AgResearch Strategic Science Investment Fund (SSIF) Microbiomes programme for the funding of this initiative.

I am hugely grateful to Dr Dragana Gagic for believing in my potential. Your support and encouragement inspired me to take on this study. I want to express my sincere gratitude to the Rumen Microbiology team leader during the study, Dr Adrian Molenaar, and my current team leader, Dr Steve Archer, for providing study leave, which was extremely helpful during the writing process. I also appreciate our facility managers, Allison McCarthy, Michelle Kirk, Arvina Ram, and Julie Collins Emerson, whose relentless support was crucial in helping us meet the compliance requirements of the Ministry of Primary Industries. I would like to acknowledge Dr Ahmed Fayaz from Massey University for his valuable advice on protein extraction techniques and MALDI-TOF MS use during the early stages of research, which significantly improved my understanding and execution of the process.

Additionally, I extend my acknowledgement to Dr Nina Butowski and Paul Maclean for bioinformatics analysis of TMAO gene hits from the Hungate1000 rumen genomic dataset. A

huge thanks to Dr Laureen Crouzet and Krsana Rajasekaran for their hands-on assistance in collecting and processing calf rumen and faeces samples for the enrichment experiment. Special thanks to Kay Pilkington, who stepped in during the busiest times to help process samples. I am also grateful to Peter Reid, Dr Mike Tavendale, and Daniel Bernstein for their time and effort in conducting GC and HPLC analyses of the samples, which provided essential results for validating the study.

To the remarkable individuals in the Rumen Microbiology team at AgResearch, Grasslands, and the dedicated Epidemiology team at Hopkirk, Massey University, your support and camaraderie have been a source of inspiration.

Finally, my family has been a great motivation that helped me complete the long-overdue M.Sc. degree, which was a personal development goal. My husband, Manpreet Singh, and my daughter, Gitaaj Kaur; thank you both for your patience during my busy work and study days and for your endless love during difficult times.

Declaration

The Hungate1000 cultures included in this study had previously been identified as potential TMA producers based on their genomic analysis for choline metabolism conducted by Dr William Kelly and have been published previously (Kelly et al., 2019). A bioinformatic analysis in section 2.1.1 was conducted on the Hungate1000 genomic dataset by Dr Elizabeth Rettedal, Dr Nina Butowski, and Paul Maclean, before the project started. The remainder of the work in this thesis was conducted by the candidate with guidance from the supervisors.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

%	Percent
°C	degrees Celsius
ACN	Acetonitrile
AMT	Aminomethyltransferase
BTS	Bacterial Test Standard
cntA	Carnitine monooxygenase
cutC	Choline TMA-lyase
DmgD	dimethylglycine dehydrogenase
DNA	Deoxyribonucleic acid
ENRIA	European Network for the Rapid Identification of Anaerobes
g	Gram
GCS	Glycine cleavage system T protein
grdH	Betaine reductase
GTDB	Genome Taxonomy Database
H1K	Hungate 1000
HCCA	α-Cyano-4-hydroxycinnamic acid
HPLC	High-performance liquid chromatography
HTP	High-throughput
L	Litre
m	Meter
m/z	Mass-to-charge ratio
M2GSC	Modified Medium 2 of Hobson
MALDI-TOF-MS	Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry
mg	Milligram
MGC	Molybdopterin guanine dinucleotide-containing S/N-oxide reductase

mL	Millilitre
mm	Millimeter
mM	Millimolar
MRM	Multiple reaction monitoring
MSPs	Main spectral profiles
nM	Nanomolar
OD_{600}	Optical density at 600 nm
PCR	Polymerase Chain Reaction
ppm	Parts per million
rRNA	Ribosomal ribonucleic acid
RTC	Real-Time Classification
SCFA	Short-chain fatty acid
TMA	Trimethylamines
TMAO	Trimethylamine N-oxide
TMM	Trimethylamine monooxygenase
TOF	Time of flight
TorA	Trimethylamine-N-oxide reductase
TS	4-Hydroxy-tetrahydrodipicolinate synthase
VFA	Volatile Fatty Acid
g	Gravity
µL	Microliter

Chapter 1. Introduction

Microbial identification is critical in microbiology, for both clinical diagnostics and ecological understanding. Traditional microbial identification techniques, including biochemical assays, phenotyping, microscopic methods, and molecular approaches, are often slow and labour-intensive (Jang & Kim, 2018). These conventional microbial identification methods depend on growth and biochemical profiling, which is especially limiting for fastidious or slow-growing microbes (Maurer et al., 2017). The introduction of matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS), particularly through the Bruker MALDI Biotyper® system, has greatly accelerated and refined microbial identification, especially within clinical laboratories (Lagier et al., 2015). The MALDI Biotyper® can rapidly identify microbes by generating protein mass spectra unique to each organism and comparing them against the reference database of mass spectral profiles (MSPs). Clinical microbiology has adopted this tool due to its speed, accuracy, and low cost per sample (Singhal et al., 2015). However, the MALDI Biotyper® identification technique has primarily been used for aerobic microbes, and its application to identify anaerobes is less common.

In recent years, MALDI Biotyper® has been employed to rapidly identify previously elusive anaerobes (Veloo et al., 2018). Anaerobes have historically been challenging to cultivate and identify due to their sensitivity to oxygen (O2) and complex culture requirements (Veloo et al., 2014). Continuous updates to the Bruker reference libraries, such as the European Network for the Rapid Identification of Anaerobes (ENRIA) project, added new clinically relevant anaerobic genera. The ENRIA project has enhanced the capacity for rapid identification of the included anaerobic genera (Veloo et al., 2018). Studies have demonstrated that the adoption of MALDI Biotyper® has led to a significant expansion in the diversity of clinical anaerobes identified within a short period (Bächli et al., 2022). However, its application in targeted culturomics (Diakite et al., 2020), environmental, and, particularly, rumen microbiology has not yet been fully realised.

In the ruminant's forestomach, the rumen is the largest compartment among the pre-gastric sections, which include the rumen, reticulum, and omasum, before reaching the true gastric stomach, the abomasum (Balch, 1959). Rumen fermentation activities are carried out by a diverse array of strictly anaerobic bacteria, archaea, fungi, and protozoa (Hungate, 2013). Rumen microbiota is central to ruminant digestion, productivity, and enteric CH_4 production, a potent greenhouse gas (Hungate, 1966). Many of the rumen-associated microorganisms remain poorly identified due to similar challenges encountered for clinical anaerobes, such as slow growth and O_2 sensitivity (Bächli et al., 2022). Use of the MALDI Biotyper® is hindered in the identification of rumen microbiology due to the lack of curated reference mass spectral profiles from rumen microbes, and the research requiring rapid identification for rumen anaerobes is currently a missed opportunity. Advancements in the field of rapid microbial identification for anaerobic microbes could enhance identification efforts for yet uncultivated members of the complex gut microbiota.

1.1 Principle of the MALDI Biotyper®

MALDI Biotyper® identifies microorganisms by analysing proteomic fingerprints and comparing them against an extensive reference library of mass spectral profiles (MSPs). The MALDI Biotyper® improves microbial identification by targeting ribosomal proteins, enabling species-level classification and continuous library expansion. The basic standard procedure involves sample preparation using a matrix solution, which coats the sample and co-crystalises upon dehydration. An automated laser beam then ionises the crystalised samples into charged molecules. Charged particles are accelerated and separated based on their charge-to-mass ratio (m/z), measured by a time of flight (TOF) analyser, and provide results as generated spectra of peptide mass fingerprinting (PMF) electrographs (Bächli et al., 2022; Singhal et al., 2015). The benefits of the MALDI Biotyper® for microbial detection include rapid real-time identification, accuracy, cost-effectiveness, and minimal training requirements, though initial equipment costs are high, and it relies on comprehensive spectral databases for precise identification (Singhal et al., 2015).

1.2 Role of microbiomes

The microbiome is defined as “a characteristic microbial community occupying a reasonably well-defined habitat which has distinct physico-chemical properties. The term thus not only refers to the microorganisms involved but also encompasses their theatre of activity” (Berg et al., 2020). Whereas microbiota is the community of microbes associated with living organisms and the environment, play a vital function in maintaining the functional ecosystems (Banerjee & van der Heijden, 2022).

The human gut and the rumen are examples of these niches and are colonised by complex anaerobic microbiomes, including bacteria, archaea, fungi, protozoa, and viruses, which perform many metabolic functions to produce essential metabolites (Jandhyala et al., 2015). Ruminants rely on rumen microbiota to produce essential metabolites such as short-chain fatty acids from the fermentation of forage or a low-cost pasture diet (Attwood et al., 2019; Gäbel & Sehested, 1997). Rumen metabolites serve as a form of energy for host growth and the production of milk and meat, which are valuable resources for food security to meet increased demand of a growing global population (Mizrahi et al., 2021).

Microbial metabolites can also be associated with causing adverse effects on hosts and are linked to disease-related metabolism (Alexandrescu et al., 2024; Li et al., 2025). Hyperproduction of disease-associated metabolites can occur because of an imbalance in the microbiota. A harmful disturbance in microbial communities caused by external factors is referred to as dysbiosis (DeGruttola et al., 2016). Numerous recent studies suggest that gut microbiota dysbiosis contributes to the development of various diseases, such as inflammatory bowel disease (IBD), obesity, and allergic conditions, in both humans and animal models (Banerjee & van der Heijden, 2022; DeGruttola et al., 2016).

Figure 1 illustrates how microbial ecosystems, such as soil, plants, ruminants, and the human gut, interact and influence microbiome composition. Studies have linked distinct microbial profiles to an increased risk of diseases and a decrease in genera associated with

anti-inflammatory activities, highlighting the potential of these profiles as biomarkers and therapeutic targets (Alexandrescu et al., 2024). However, within the microbiome, it is challenging to associate the activity of individual microbes with specific metabolic functions until they are experimentally tested *in vitro* and validated using pure cultures (Martínez-del Campo et al., 2015).

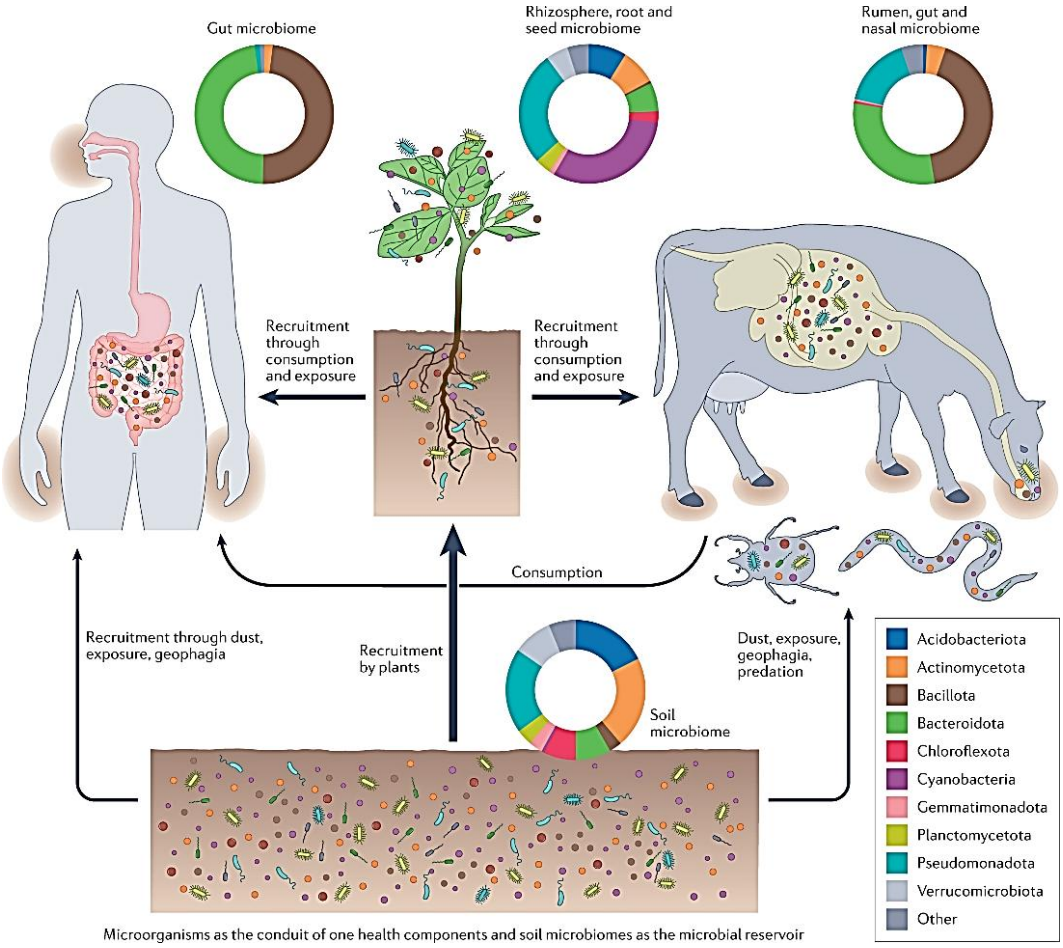


Figure 1. Microbial ecosystems and their influence on microbiome composition. Microbial ecosystems like soil, plants, ruminants, and the human gut occur through direct and indirect pathways, influencing the composition of various microbiomes identified by metagenomics. The figure was obtained from (Banerjee & van der Heijden, 2022) and is used by permission/ license obtained for copyright clearance from the publishers, Nature Reviews Microbiology (https://www.nature.com/articles/s41579-022-00779-w).

1.3 Role of metagenomics

Metagenomics is a culture-independent, sequence-based analysis platform that examines the collective microbial DNA extracted from samples representing ecological environments, such as the rumen, human gut, soil, marine water, and extreme environments (Sleator et al., 2008). This technique provides genomic functional insights into the large proportion of microbial diversity that remains uncultured, highlighting a limitation in traditional culturing methods (Li et al., 2022). However, DNA sequence data alone does not indicate the activity or physiological state of the microorganisms. When integrated with meta-transcriptomic, metabolomic, and metaproteomic analyses, metagenomics can provide insights into microbial populations, gene expression, and biochemical fermentation pathways (Falony et al., 2015).

Despite this, culturing can supplement our understanding by providing evidence of microbial functions and validating them. The strengths of metagenomics include its culture-independent analysis, while its weaknesses lie in being limited by the lack of reference data from the presentation of a largely uncultured proportion of microbial diversity. Metagenomics require integration with other omics techniques for a more comprehensive understanding of microbial function (Tidjani Alou et al., 2017). The microbial community structure and functional analysis of gut microbiomes have been elucidated through gut metagenomics sequencing data (Banerjee & van der Heijden, 2022; Henderson et al., 2015). Rumen bacterial diversity studies reveal many uncultured taxa, indicating a need for further cultivation efforts (Creevey et al., 2014; Seshadri et al., 2018).

1.4 Culturomics

Culturomics is defined as a high-throughput (HTP) culturing approach enabling an elaborated assessment and isolation of the uncultivated or low-abundance microbes from microbiomes. Despite the development of HTP sequencing techniques used in microbiota studies, most microbial content remains undetected, and is known as microbial dark matter

(Greub, 2012; Lagier et al., 2012). Recently, culturomics has gained success in identifying optimal culture conditions for isolation of novel organisms. Culturomics has achieved increased cultivability along with rapid microbial identification as compared to traditional culturing methods (Diakite et al., 2020). Many unassigned sequences obtained from metagenomics are microbes of interest, and targeting their cultivation is a frontier in microbiology. Isolating new microbes will increase our understanding of the microbiology of the human gut and the rumen's microbial repertoire, and provide us with new taxa (Diakite et al., 2020).

1.5 Need for advancements in pure microbial isolations and identifications

Studying pure cultures is crucial for advancing specific disease treatments, such as bacteriotherapy with probiotics or the development of targeted inhibitors for CH_4 inhibition and vaccine therapies (Falony et al., 2015; Janeiro et al., 2018). Despite advancements in HTP metagenomic sequencing analysis, a gap persists in characterising pure microbial isolates due to the challenging nature of cultivating and identifying new organisms (Lagier et al., 2015). To fill this gap, it is essential to make advancements in HTP cultivation and identification techniques to assist targeted isolation of microbes of interest. The complexity of the rumen microbiome and the difficulties in cultivating many of its members is a challenge for microbial research, and these boundaries need to be pushed to improve our understanding of microbial diversity and function in natural ecosystems (Huws et al., 2018).

Earlier efforts to isolate uncultured rumen bacteria using dilution-to-extinction, where samples are diluted until each tube ideally contains a single microbial cell, have successfully led to the discovery of new rumen species (Kenters et al., 2011; Mahoney-Kurpe et al., 2023; Noel, 2013). Additionally, the Hungate1000 (H1K) rumen culture collection is the most diverse culture collection of rumen microbial isolates, with matching genome sequence data availability, representing a valuable resource (Seshadri et al., 2018). These efforts to cultivate plant-adherent rumen bacteria (Noel, 2013) relied heavily on 16S rRNA gene sequencing for taxonomic identification, which, while accurate, is time-consuming and

laborious. A similar reliance on 16S rRNA gene sequencing is also evident in large-scale human gut culturomics initiatives (Lagier et al., 2012; Lagier et al., 2016). These studies aimed to culture a broad diversity of anaerobic microbes and have significantly expanded our understanding of rumen and gut microbial diversity. However, 16S rRNA-based sequencing has limited the throughput of identification of microbes.

Developing MALDI-TOF MS spectral profiles from existing rumen culture collections could revolutionise the identification of rumen microbes, enabling rapid and precise taxonomic classification (Zhou et al., 2021). Microbial identification using the MALDI Biotyper® produces results in less than 1 hour, saving the several days needed to perform PCRs, 16S rRNA gene sequencing and data analysis. Additionally, faster screening of microbial colonies will capture wider diversity and minimise the time spent on repeatedly screening similar. The rapid and accurate identification of species that produce targeted metabolites could accelerate and enhance research projects.

1.6 Trimethylamine (TMA) formation in the gut

The focus of this study was cultivating microorganisms that contribute to the formation of the gut-derived metabolite, TMA. Figure 2 illustrates the role of TMA as a precursor for trimethylamine N-oxide (TMAO), a metabolite linked to various human diseases (Rath et al., 2017), and also as a precursor in the formation of CH_4, a potent greenhouse gas produced by ruminants (Zhou et al., 2021). In the rumen, TMA is produced from choline by bacteria possessing choline trimethylamine-lyase activity (Kelly et al., 2019), through the breakdown of dietary plant phosphatidylcholine, a major component of grass lipids (Dawson & Hemington, 1974). TMA is then converted to CH_4 via methanogenesis by a specific group of methylotrophic methanogens, particularly those belonging to *Methanomassiliicoccales* (Kelly et al., 2019; Zhou et al., 2021).

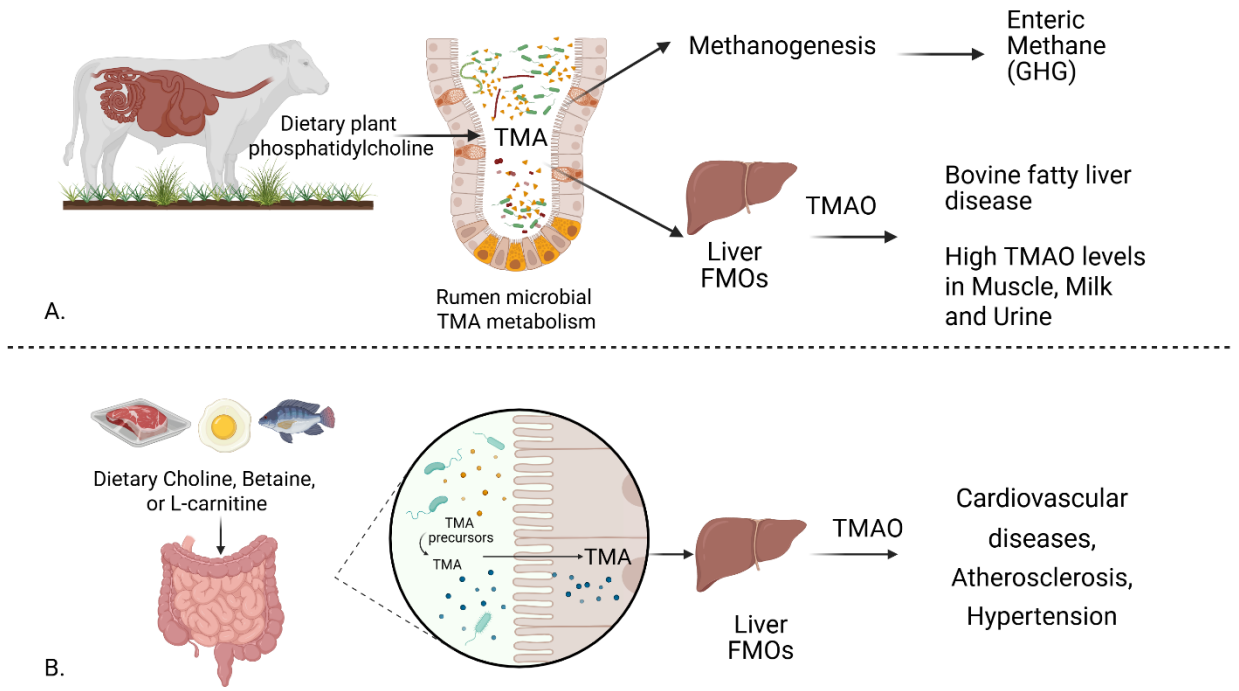


Figure 2. Flow of dietary components to TMA and TMAO production in ruminants and humans. A. Rumen microbiota converts plant phosphatidylcholine into TMA, linked to CH_4 and fatty liver disease. B. Human gut microbiota metabolises choline, betaine, or l-carnitine into TMA, which is oxidised by liver flavin-containing monooxygenases (FMOs) into TMAO, associated with cardiovascular risks. The figure was created using BioRender (https://BioRender.com).

Methanogenesis is a process of anaerobic respiration carried out by a specialised group of archaea known as methanogens (Lyu et al., 2018). Methanogens produce CH_4 as a metabolic end product, using carbon dioxide (CO_2), hydrogen (H_2), acetate, or methylated compounds (Lyu et al., 2018). This process plays a vital role in the global carbon cycle and is especially important in environments such as wetlands, the rumen, and anaerobic digesters (Huws et al., 2018). While exploring CH_4 reduction strategies, it was observed that a potential downside of these strategies can be elevated TMA levels if TMA is not fully converted to CH_4 (Zhou et al., 2021). TMA accumulation could lead to increased TMAO levels in meat and milk, posing potential health risks to consumers (Servillo et al., 2018; Zhou et al., 2021).

TMAO has been linked to bovine fatty liver disease, a considerable economic issue in the dairy industry (Li et al., 2025). Through the gut-liver axis, TMAO has been linked with non-alcoholic fatty liver disease in dairy cows, affecting numerous genes, proteins, and insulin resistance (Li et al., 2025). The same study revealed that TMAO influences liver lipid accumulation through adenosine monophosphate-activated protein kinase (AMPK) signalling and oxidative phosphorylation. Overall, metabolic homeostasis is disrupted, affecting the bile acid signalling, leading to lipid accumulation and inflammation (Li et al., 2025). Hence, regulation of TMA levels is essential for optimal productivity of ruminants.

Quaternary amines, which are precursors of TMA in the human diet, include choline, betaine, and carnitine. These compounds are derived from dietary sources such as fish, meat, and eggs (Rath et al., 2017). TMA is oxidised to TMAO by hepatic enzymes flavin-containing monooxygenases (Falony et al., 2015). Elevated plasma levels of TMAO are linked to inflammation and serious health issues, including cardiovascular disease and atherosclerosis, and may also contribute to neurological disorders (Janeiro et al., 2018).

1.7 Key TMA precursors, microbial enzymes, and metabolic pathways

In the human gut and rumen, the microbial degradation of choline, carnitine, and betaine into TMA involves distinct metabolic pathways (Figure 3). The microbial breakdown of these substrates under anaerobic conditions involves enzymes, including choline TMA-lyase (encoded by *cutC),* the activator enzyme (*cutD*), and betaine reductase (*grdH*) (Jameson et al., 2018). Microbes that possess specific genes encoding enzymes such as carnitine monooxygenase (*cntA* and *cntB*) are known to perform oxygen-dependent catalysis to produce TMA, which occurs mostly in the upper gastrointestinal tract (Arias et al., 2020; Janeiro et al., 2018; Mouné et al., 1999). Kelly et al. (2019) analysed the H1K genomic dataset to detect the presence and expression of genes linked to methyl-compound production in rumen bacteria, aiming to identify potential TMA-producing strains.

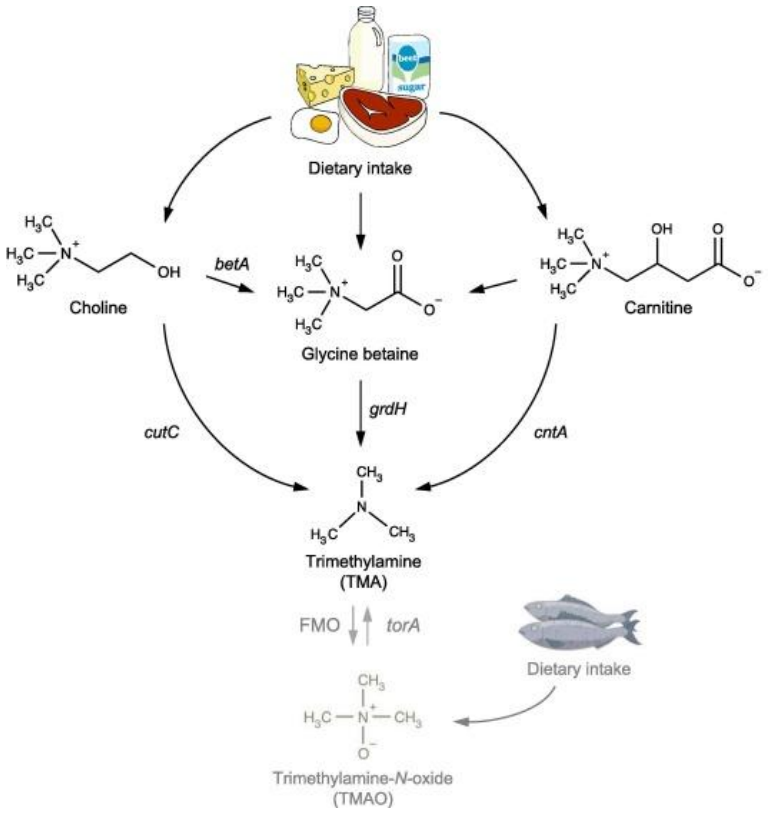


Figure 3. Metabolic pathways and genes for TMA formation from dietary precursors. Key metabolic pathways and identified genes for enzyme production for TMA formation from dietary choline, betaine, and carnitine in the gut. Image obtained from Jameson et al., (2018) and used under Creative Commons Attribution 4.0 International deed (CC BY 4.0) (https://www.sciencedirect.com/science/article/pii/S1046202317304474?via%3Dihub).

1.7.1. Anaerobic choline degradation pathways

The best-characterised pathway is the choline utilisation (*cut*) pathway, where choline TMA-lyase, encoded by *cutC* and the activation enzyme *cutD*, are the key enzymes that catalyse the direct cleavage of choline into TMA and acetaldehyde (Figure 4) (Martínez-del Campo et al., 2015). This enzyme is a glycyl radical enzyme and requires an activation enzyme under anaerobic conditions for the generation of a radical cofactor (Craciun & Balskus, 2012). Additional genes within the *cut* operon, such as *cutE*, *cutF*, and *cutG*, are believed to function in choline transport (Martínez-del Campo et al., 2015), regulation or stabilisation of the enzyme complex, contributing to efficient choline catabolism in anaerobic microbes

such as *Olsenella umbonata, Clostridium sporogenes,* and *Desulfovibrio* spp. (Arias et al., 2020). The choline TMA lyase enzyme complex is located near genes that code for parts of bacterial microcompartments, self-assembling organelles inside bacteria (Craciun & Balskus, 2012; Kerfeld et al., 2018). These microcompartments help contain the choline degradation reactions, preventing harmful byproducts like acetaldehyde from damaging the rest of the cell (Craciun & Balskus, 2012).

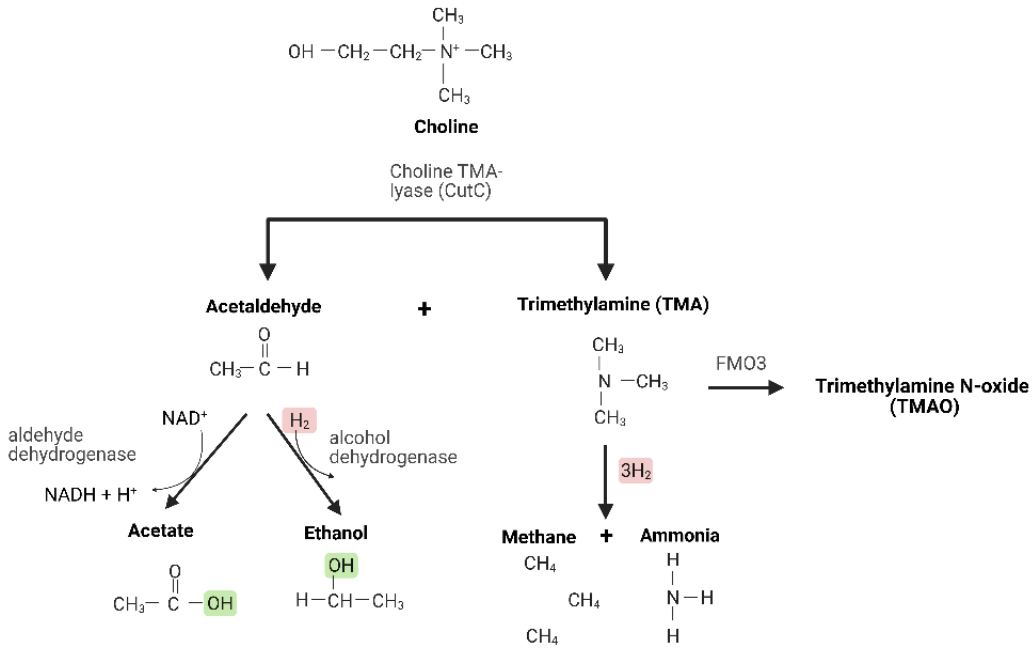
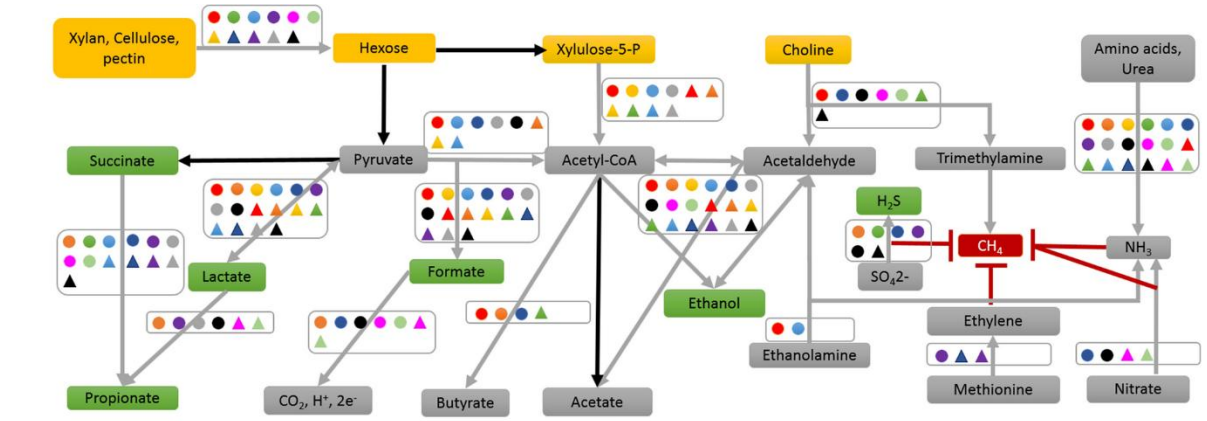


Figure 4. Biochemical pathway and stoichiometry of anaerobic choline degradation. The figure of the biochemical pathway was created using BioRender (https://BioRender.com.).

A study investigating the effects of high doses of choline in rumen *in vitro* models (Li et al., 2021)found inhibition of CH_4 formation and shifts in the rumen microbial community and metabolic pathways and outputs (Figure 5) (Li et al., 2021). The primary fermentation end-products from the breakdown of carbohydrates, choline, and amino acids in a rumen-like environment included short chain fatty acids (SCFA), ethanol, formate, CH_4, and trimethylamine (TMA), with some pathways also contributing to the formation of hydrogen sulfide and ammonia (Li et al., 2021).



Taxa [mOTUs cluster ID]	Ctl	ChCl	ChHC
Enterococcus avium [ref_mOTU_v25_02620]	0%	0%	19.7%
Megasphaera elsdenii [ref_mOTU_v25_01516]	2.22%	7.62%	8.89%
Streptococcus equinus [ref_mOTU_v25_00901]	3.25%	3.47%	10.9%
unknown *Prevotella* [rumen_mOTU_728]	1.19%	15.45%	0%
Enterococcus gallinarum [ref_mOTU_v25_03214]	0%	0.03%	12.2%
unknown *Alkaliphilus* [rumen_mOTU_765]	0%	0.04%	10.5%
Prevotella bryantii [ref_mOTU_v25_03809]	2.04%	5.97%	0.1%
Globicatella sanguinis [ref_mOTU_v25_02265]	0%	0%	7.38%
Proteus mirabilis [ref_mOTU_v25_02513]	0%	0.02%	7.19%
unknown Anaerovoracaceae [rumen_mOTU_731]	0.01%	5.74%,	0%
unknown Anaerovoracaceae [rumen_mOTU_724]	0%	2.73%	0%

Taxa [mOTUs cluster ID]	Ctl	ChCl	ChHC
Lactobacillus ruminis [ref_mOTU_v25_01239]	0.4%	3.23%	1.19%
Lactobacillus salivarius [ref_mOTU_v25_00945]	0.28%	2.32%	2.1%
Lactobacillus agilis [ref_mOTU_v25_04375]	0%	4.2%	0.21%
Olsenella umbonata [ref_mOTU_v25_06395]	0.01%	4.29%	0.06%
Lactobacillus mucosae [ref_mOTU_v25_02550]	1.13%	1.77%	0.63%
Lachnospira multipara [ref_mOTU_v25_03833]	0.07%	1.66%	1.09%
unknown Lachnospiraceae [rumen_mOTU_727]	0.003%	0.15%	0.24%
Bifidobacterium adolescentis [ref_mOTU_v25_02703]	0.14%	2.18%	0.37%
unknown *Lachnotalea* [rumen_mOTU_763]	0%	0%	1.67%
Denitrobacterium detoxificans [ref_mOTU_v25_06442]	0.001%	0.08%	0.08%
Denitrobacterium detoxificans [rumen_mOTU_2272]	0.001%	0.07%	0.06%

Figure 5. Proposed metabolic pathway of VFA production and CH_4 mitigation by abundant species and species of interest rumen microbial species *in vitro*. The yellow boxes represent nutrients from feed bags and supplements supplied to a rumen *in vitro* system (Rusitec), the green boxes represent alternative H_2 sinks, the red box represents CH_4, and the other metabolites are shown in grey boxes. Black arrows represent pathways shared by all species; grey arrows are pathways only the species depicted by the legend are known or predicted to possess. The legend representing each species is shown, along with its average relative abundance from days 10 and 15 in each treatment group. (Li et al., 2021). Image obtained from Li, Y. et al. (2021) used under CC BY 4.0. https://www.nature.com/articles/s41598-021-01031-w)

1.7.2. Anaerobic betaine degradation pathways

Betaine in the rumen ecosystem naturally originates from dietary sources and plays multiple roles in microbial metabolism (Mahmood et al., 2020). Betaine is used as an osmolyte to protect microbes under stress, as a methyl donor in microbial metabolism and as a substrate for fermentation under stress conditions (Ghoneem & El-Tanany, 2023). Recent studies have demonstrated that betaine supplementation enhances milk yield and dry matter intake during lactation in ruminants, particularly under both thermoneutral and heat-

stress conditions (Abhijith et al., 2024). Betaine modulates fermentation pathways, increasing propionate production and short-chain fatty acid (SCFA) concentrations (Mahmood et al., 2020). Certain anaerobic bacteria break down betaine to produce TMA using a pathway that involves betaine reductase GrdHI. Betaine reductase removes methyl groups from betaine in oxygen-free environments (Jameson et al., 2018). Enzyme complex GrdHI in the marine environment enable betaine utilisation for energy production and growth, releasing TMA as a byproduct (Boysen et al., 2022). However, betaine metabolism in the rumen environment is less well understood.

In *Desulfitobacterium hafniense,* an anaerobic gram-positive, organohalide-respiring bacterium, a glycine betaine methyltransferase, MtgB, catalyses the conversion of glycine betaine to dimethylglycine without producing TMA (Ticak et al., 2014). The enzyme works in conjunction with MtgA, a methylcobalamin: tetrahydrofolate methyltransferase, to complete the methyl transfer pathway. Similarly, in the methanogen *Methanolobus volcanii* isolated from estuarine sediment, the betaine to TMA methyltransferase pathway involves a methyltransferase MttB transferring a methyl group from betaine to a corrinoid cofactor, then a secondary methylcorrinoid: coenzyme M methyltransferase, MtbA, transfers that methyl group to CoM (Creighbaum et al., 2019). These alternative metabolic pathways reveal a widespread potential of anaerobic microbes utilising betaine without TMA production.

1.7.3. Carnitine degradation pathways

In the gut, the microbial breakdown of L-carnitine into TMA is facilitated by a two-component oxygenase–reductase system encoded by *cntA* and *cntB*. The *cntA* gene encodes a Rieske-type monooxygenase that hydroxylates L-carnitine, enabling cleavage of the C–N bond and release of TMA (along with malic semialdehyde). While *cntB* encodes an associated reductase that provides the necessary electrons to drive oxygenase activity (Figure 6) (Massmig et al., 2020; Quareshy et al., 2021). Furthermore, *in vitro* assays have demonstrated that activities associated with CntA and CntB are both necessary and sufficient for converting L-carnitine to TMA (Zhu et al., 2014). The CntA/B pathway is not

effective under strict anaerobic conditions due to its dependence on oxygen. Facultative anaerobes like *Escherichia coli* and *Acinetobacter* spp. colonise epithelial surfaces or mucosal biofilms, where low O_2 levels permit the activity of CntA/B-type monooxygenases, enabling L-carnitine derived TMA production even in the largely anaerobic gut environment (Zhu et al., 2014).

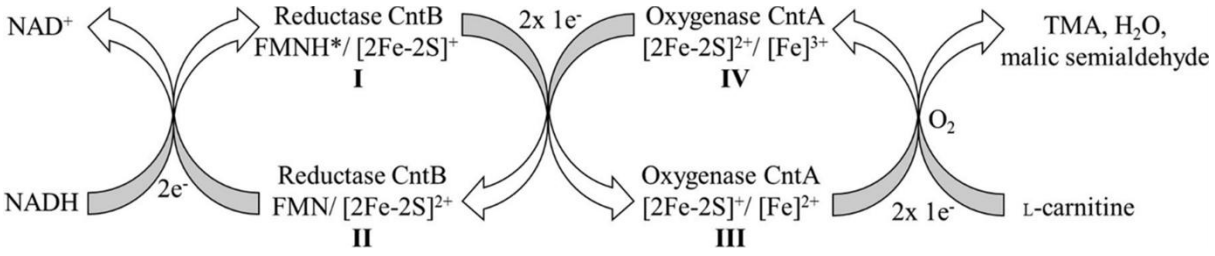


Figure 6. Proposed redox catalytic cycle of carnitine monooxygenase. A redox catalytic cycle for carnitine monooxygenase was proposed based on the experimentally characterised redox states of the present study (Massmig et al., 2020). Image obtained from Massmig et al. (2020), used under CC BY 4.0, with no modifications, and is used in the same manner as the original. https://doi.org/10.1074/jbc.RA120.014266

Carnitine supplementation in ruminants has been studied for its potential benefits in energy metabolism, performance, and recovery during metabolically demanding periods of calving and early lactation (Meyer et al., 2020). During lactation, cows excrete L-carnitine in milk, which may decrease its availability for other metabolic functions (Meyer et al., 2020). Rumen microbes rapidly break down supplemented L-carnitine (Harmeyer, 2003). Therefore, rumen-protected L-carnitine formulations are recommended to ensure bioavailability. Rumen-protected supplementation can enhance mitochondrial energy production, reducing the severity of negative energy balance during calving and early lactation, and may support faster recovery by improving metabolic efficiency (Meyer et al., 2020). Supplementation can also increase L-carnitine concentrations in milk and lean muscle (Harmeyer, 2003). While most studies focus on dairy cows, some evidence suggests potential benefits in growth and reproductive performance across various ruminant species (Ringseis et al., 2018).

1.7.4. Trimethylamine-N-Oxide Reductase (TorA)

The molybdopterin guanine dinucleotide (MGD)-containing S/N-oxide reductases are a family of molybdoenzymes that play a crucial role in the microbial reduction of various N-oxides and S-oxides. A prominent example within this family is TMAO reductase, which facilitates the conversion of TMAO to TMA (Panwar et al., 2024). TorA is an enzyme that catalyses the reduction of TMAO back to TMA, thereby playing a crucial role in the TMA metabolic cycle (Hoyles et al., 2018). This process, known as metabolic retroversion, helps maintain the balance of TMA and TMAO within the host. The conversion of TMAO back to TMA has been well-documented in humans and other monogastric animals, primarily facilitated by gut microbiota such as *Enterobacteriaceae* (Hoyles et al., 2018; Janeiro et al., 2018). However, researchers have not explored the relevance of this pathway in ruminants. Moreover, the strictly anaerobic rumen environment limits the establishment of TorA and may not favour TMAO reduction mechanisms. Table 1 lists additional pathways that explore how methylated amines may be metabolised in the gut, contributing directly or indirectly to the TMA-TMAO interconversion.

Table 1. Enzymes from indirect pathways are potentially linked to microbial TMA metabolism.

Enzymes/ proteins	Function
Aminomethyltransferase (AMT)	Primarily involved in the glycine cleavage system, plays a role in one-carbon metabolism and methyl group transfer, and may influence microbial pathways that contribute to TMA synthesis (De Anda et al., 2021).
Glycine cleavage system T-protein	Part of a system that converts glycine to ammonia and carbon dioxide, potentially linking glycine metabolism to TMA pathways (King, 1984; Lee et al., 2004; Li et al., 2025).
Dimethylglycine dehydrogenase	Catalyses the oxidation of dimethylglycine to glycine, thus contributing to the pool of substrates available for TMA production (Janeiro et al., 2018).

1.8 Strengths and limitations of predicting TMA metabolism using gene sequences

Potential TMA-producing gut microbes have been identified in human faecal samples using metagenomic analysis (Jie et al., 2017; Rath et al., 2017; Romano et al., 2015). TMA-related genes, such *as cutC* and *cntA*, were identified using tools like BLAST and profile Hidden Markov Models (HMMs), each with its own strengths in detecting functional genes (Jameson et al., 2018). However, these methods can yield false positives due to mis-annotation of gene homologues, and false negatives when genes are incorrectly assigned, such as *cutC* being misclassified as pyruvate-formate lyase (Jie et al., 2017). Use of 16S rRNA gene sequencing is a powerful tool for identifying bacteria but does not indicate the presence of specific functional genes such as *cutC, cntA,* or *grdH*, which are directly involved in converting choline, carnitine, and betaine into TMA (Jameson et al., 2018). Genes encoding TMA production are often unevenly distributed among bacterial strains, even within the same species or genus; for example, only certain strains of *Clostridium* or *Enterococcus* contain the *cutC* gene (Romano et al., 2015). This is because such genes may be gained through horizontal gene transfer or lost over time (Rath et al., 2017).

Screening microbial isolates for the presence of genes such as *cutCD* and *cntAB*, through PCR or genome analyses, will aid the identification of putative TMA-metabolising strains. However, the TMA-producing activity of strains requires validation through the detection of TMA as a metabolic endpoint. A gene-targeted assay of human faecal samples revealed that TMA-producing bacteria exist in low abundance, and pure microbial isolates are yet to be obtained for *in vivo* experimental validation of TMA production (Rath et al., 2017).

1.9 Research gaps

There is a gap in the availability of pure cultures that can produce TMA isolated from the rumen and gut. There is also a lack of a clear understanding of the microbial enzymes that create methyl compounds (Kelly et al., 2019). A research gap is also identified in the Bruker

Biotyper® database with limited representation of mass spectral profiles (MSPs) for rumen strains.

1.10 Project aim and objectives

This project aims to cultivate TMA-producing rumen isolates from TMA-precursor enrichment cultures and identify these isolates using the MALDI Biotyper®. This will be achieved by undertaking the following objectives:

- Develop a custom database expansion for MALDI-Biotyper® using candidate bacterial strains from H1K culture collection, which are associated with TMA-production.
- To enrich microbial cultures from rumen samples using choline, betaine, and carnitine as selective substrates to favour the growth of potential TMA producers.
- Isolate novel TMA-producing bacterial strains from selective enrichments based on their ability to grow as isolated microbial colonies on solid agar medium containing TMA precursors under anaerobic conditions and use MALDI-TOF MS to identify new isolates using a customised expanded database.
- Develop methods to validate TMA production experimentally using high-performance liquid chromatography (HPLC) to confirm functional activity of the isolates.

Chapter 2. Materials and Methods

2.1 Strain selection for database expansion

Bacterial strains from the H1K rumen culture collection, identified as potential TMA producers, were selected to generate targeted MSPs for expanding the MALDI Biotyper® database. These strains are primarily associated with the presence of choline degradation pathways, leading to TMA production (Kelly et al., 2019). H1K strains associated with the retro-conversion of TMAO metabolism were pre-identified using the method explained in Section 2.1.1. identified stains were also included in the list of candidate strains for database expansion.

2.1.1. Rumen-derived gut bacteria screened for their ability to utilise TMAO

Predicted protein sequences for the H1K strains were searched against the KEGG prokaryotic protein database using the *blastp* function of BLAST+ version 2.13.0 with an e-value cut-off of 1E-50 but otherwise default values (Camacho et al., 2009). The resulting hits in the KEGG prokaryotic protein database were matched with KEGG terms and filtered down to those associated with TMAO metabolism, K07811, and K07812: trimethylamine-N-oxide reductase (cytochrome c), K18277: trimethylamine monooxygenase, and K00605: aminomethyltransferase, non-oxidative demethylation of trimethylamine N-oxide by *Pseudomonas aminovorans*.

2.2 Bruker MALDI Biotyper® methods

2.2.1. Preparation of the HCCA matrix solution

The α-cyano-4-hydroxycinnamic acid matrix (HCCA, Bruker Bremen, Germany) or HCCA from Sigma-Aldrich (Missouri, USA) was used. To 2.5 mg HCCA, 250 µL of standard solvent (50% acetonitrile, 47.5% water, and 2.5% trifluoroacetic acid; Bruker Standard Solvent) to a final concentration of 10 mg HCCA/mL was added. The solution was vortexed at room

temperature until all crystals were dissolved and stored at room temperature in an airtight tube. The HCCA matrix was used within the same week of preparation.

2.2.2. Bacterial test standard (BTS)

To prepare the BTS solution, BTS (Bruker) was dissolved in 50 µL of standard solvent. The solution was centrifuged at 18,000 x *g* for 2 minutes at room temperature. Aliquots of 5 µL of the suspensions were transferred into 0.5 mL airtight microcentrifuge tubes and stored at -20°C.

2.2.3. Sample preparation using the direct transfer protein extraction method

A single, well-isolated bacterial colony was picked from an agar plate using a sterile toothpick or loop and smeared onto a MALDI target plate Biotarget 96 Plate (Bruker, Bremen, Germany) and air-dried. Once dry, 1 µL of 70% (v/v) formic acid was added to lyse the cells and release proteins, which was followed by another drying period. Next, 1 µL of HCCA matrix solution was applied, and the sample was allowed to air-dry completely. The prepared Biotarget 96 Plate was then used for analysis.

2.2.4. Sample preparation using extended protein extraction

Isolated bacterial colonies were transferred to 300 µL of HPLC-grade water and mixed thoroughly. The suspension was vortexed for 1 minute, then 900 µL of absolute ethanol was added and vortexed again. Samples were stored at –20 °C or processed immediately. The ethanol suspension was centrifuged at 18,000 × *g* for 2 minutes, the supernatant discarded, and the pellet was further centrifuged to remove residual liquid. The pellet was air-dried for at least 5 minutes and then resuspended in 70% formic acid, adjusting the volume based on colony size, which is 5 µL for a single colony and 10-25 µL for two to four, pooled colonies. An equal volume of acetonitrile was added, and the mixture was vortexed and centrifuged again. Finally, 1.5 µL of the supernatant was spotted onto a MALDI target plate, air-dried, and overlaid with HCCA matrix solution for MALDI-TOF MS analysis.

2.3 Candidate strain preparation and identification using the MALDI Biotyper®

The candidate rumen strains were first screened for identification using the MALDI-TOF MS Microflex™ LT/SH bench top systems (Bruker, Bremen, Germany). Commercially available agar plates of Brain Heart Infusion (BHI) Agar (Fort Richard, City, Country), Columbia Sheep Blood Agar (Fort Richard, City, Country), Fastidious Anaerobic Agar (Fort Richard, City, Country) and M2GSC agar plates (Miyazaki et al., 1997) were used for cultivation. Plates were stored in the COY anaerobic chamber for two days to make them anaerobic before use. Cultures were chosen to grow on M2GSC agar medium under strict anaerobic conditions using 96% CO_2, 4% H_2 atmosphere in a Coy Vinyl Type-B anaerobic chamber (Coy Laboratory Products, Grass Lake, Michigan, USA), incubated at 39 °C for 3–5 days to obtain well-isolated colonies. Isolated colonies were preserved as ethanol cell suspensions (Section 2.2.4). The suspensions were processed using the extended protein extraction method (Section 2.2.4) to obtain high-quality protein preparations that were suitable for spectral generation. For MALDI-TOF MS analysis, the protein extract was analysed using a Bruker Daltonics MALDI Biotyper® Compass (v4.1.100) and monitored using the Bruker Daltonics 22 flexControl (v3.4) software. Spectra were matched against the Bruker Daltonic Revision H database (Revision M (July 2021) Doc. No. 5026705). Identification scores ≥2.0 indicated reliable species-level identification, and scores ≥1.7 indicated genus-level identification.

2.4 Library creation for database expansion and verification

To create a targeted custom library for anaerobic rumen bacteria that are not present in the Bruker reference database for the MALDI Biotyper®, the required four key software components are flexControl (v3.4), flexAnalysis (v3.4), MALDI Biotyper® Compass Explorer (v4.1.100) and MALDI Biotyper® RTC. These tools work together to support sample processing, spectra acquisition, library creation, and identification.

2.4.1. Spectra measurement using flexControl

The target plate was prepared by overlaying a minimum of 12 - 15 target spots for the protein sample and 1 spot for the BTS using 1.5 µL of extract (Section 2.2) per spot. The flexControl system was calibrated using the BTS control sample each time before creating a library, following the manufacturer's guidelines (Bruker Daltonik GmbH, 2020). The laser power was optimised to ensure efficient ionisation of the crystallised proteins. Ionised particles were accelerated through the TOF analyser, and their mass-to-charge ratios were recorded as distinct peaks in the mass spectrum. A strain-specific, customised spectra acquisition method was saved for each isolate. The inbuilt auto-execute function was used to acquire three spectra per spot, resulting in a total of 36 spectra from a total of 12 target spots overlayed from extracted protein of each candidate strain.

2.4.2. Quality control using FlexAnalysis.

All captured spectra per sample were analysed using flexAnalysis. The process involved downloading the spectra, performing baseline subtraction, and applying smoothing. For quality control, the calibration constants of both the sample and the BTS control were compared and were required to be identical to confirm that the sample and the BTS were captured under the same conditions. Spectra with flatline profiles and outlier peaks were removed. Peak shifts were evaluated for seven major peaks within the m/z range of 3,000 to 10,000, assessed at one peak every 1,000 steps. Any spectrum with peak shifts exceeding the permitted mass tolerance of 500 ppm was excluded.

2.4.3. MSP creation using MALDI Biotyper® Compass Explorer

Twenty or more high-quality spectra that have passed quality control (Section 2.4.1) were downloaded using MALDI Biotyper® Compass Explore. The MSP was created for each strain. The MSP was tested against the spectrum to obtain the highest identification score of 3 and then integrated into the taxonomy tree under a new node with a unique project name.

2.4.4. User-created library verification using MALDI Biotyper® RTC

User-created libraries were accessed by MALDI Biotyper® RTC under a user-created project. To ensure the accuracy and reliability of the custom MALDI-TOF MS spectral libraries developed for the selected H1K bacterial strains (Section 2.1), each MSP was verified. Each strain was freshly cultured on solid M2GSC agar medium, and proteins were extracted directly on the target plate from isolated colonies. BTS was used as a control each time the target plates were analysed. Verification was by spectral acquisition and real-time identification using Real-Time Classification (RTC) software. The resulting spectra were then matched against the newly created MSPs within the user-defined library. Identification scores generated by the RTC software were interpreted according to standard thresholds. A score of ≥2.0 was considered a reliable species-level match, confirming that the MSP accurately represented the strain. Libraries that consistently produced scores of ≥2.0 were deemed verified and valid for inclusion in the H1K database. Conversely, any MSPs that yielded identification scores below 2.0 were flagged as unacceptable, indicating the spectral quality was inadequate or an incorrect representation of the strain. These MSPs were excluded from the final library and were generated again as described.

Verification was crucial for maintaining the integrity of the custom spectral library and ensuring high confidence in the identification of rumen microbial strains to the species level in downstream applications.

2.5 Anaerobic cultivation methods

2.5.1. Anaerobic cultivation

Anaerobic microbial cultures were prepared using the Hungate technique (Hungate, 1969) with modifications, which enable strict anaerobic conditions through the use of gas-tight culture tubes and oxygen-free gas flushing. Media were dispensed into sterile Hungate tubes under a continuous flow of 100% CO_2 gas (BOC, Catalogue No. G013), to displace oxygen. Tubes were sealed with butyl rubber stoppers and screw caps to maintain anaerobiosis.

Media were autoclaved and cooled under anaerobic conditions, and reducing agents were added as required (Hungate, 1969). Inoculation was performed using sterile syringes through the rubber septa to prevent O_2 exposure. Cultures were incubated at 39°C for microbial growth under strictly anaerobic conditions.

2.5.2. Anaerobic cultivation using the COY anaerobic chamber

Anaerobic cultures were prepared and maintained in a Coy anaerobic chamber (Coy Laboratory Products, Grass Lake, Michigan, USA) containing a gas mix of ~95% CO_2, <4% H_2 and <50 ppm O_2. O_2 was eliminated using a palladium catalyst system, and gas levels were monitored using a CAM-12S monitor (Coy Laboratory Products, Grass Lake, Michigan, USA). Sample transfers into the chamber were performed through a vacuum airlock with automated purge cycles. All handling and incubation of cultures occurred within the chamber under sterile, O_2-free conditions.

2.5.3. Modified RM02 with selective substrates

A bicarbonate-buffered mineral salts solution was composed of 950 mL distilled water containing 1.4 g KH_2PO_4, 0.6 g $(NH_4)_2SO_4$, 1.5 g KCl, and 1 mL trace element solution SL1 (Widdel et al., 1983), 1 mL of selenite/tungstate solution (Tschech & Pfennig, 1984), and 0.4 mL of 0.1% (w/v) resazurin solution. The solution was mixed thoroughly, boiled, gassed under an O_2-free atmosphere using 100% CO_2 and cooled in an ice bath while continuously bubbled with 100% CO_2. Once cooled, 50 mL of clarified rumen fluid (Section 2.5.4) and choline chloride at a final concentration of 25 mM were added and dissolved. To make a selective medium for betaine or carnitine, choline was replaced by each of the substrates at the same concentration of 25 mM. The medium continued to be gassed with CO_2 for another 30 minutes or until the substrates had dissolved. Then, 4.2 g $NaHCO_3$ and 0.5 g L-cysteine·HCl·H_2O were added per litre while maintaining the CO_2 gas flow in the headspace of the vessel containing the reduced medium. The final pH of the medium was adjusted to 6.5.

For RM02 broth, the prepared medium was dispensed into Hungate tubes (13 mm diameter, 100 mm length; Bellco Glass, Vineland, New Jersey, USA) under a CO_2 atmosphere, with 5 mL of medium per tube. Tubes were sealed using black butyl rubber stoppers and perforated plastic caps, ensuring a CO_2 headspace. Sterilisation was performed by autoclaving at 121 °C for 20 minutes.

For RMO2 agar, a 1.5% final concentration was made. 7.5 g of agar was added to high-pressure 1L Schott bottles (Duran® Pressure Plus bottles volume 1,000 mL), and 500 mL of prepared RM02 broth was added while gassing with CO_2 and sealed with a large rubber bung and screw cap. Sterilisation was performed by autoclaving at 121 °C for 20 minutes. Molten, sterilised agar was cooled to 55°C, then poured into petri dishes under anaerobic conditions inside a COY anaerobic chamber.

2.5.4. Clarified rumen fluid

Rumen fluid was collected from fistulated, hay-fed cattle and squeezed through cheesecloth to remove large particulate matter. The filtrate was then centrifuged at 10,000 × *g* for 20 minutes at 4 °C, and the resulting supernatant was stored at −20 °C. For media preparation, the frozen rumen fluid was thawed in hot water, transferred to a Duran® Pressure Plus bottle, and heated in a microwave until boiling. The fluid was flushed with 100% nitrogen gas for at least 20 minutes, loosely capped, and sterilised by autoclaving. After cooling, the sterilised rumen fluid was transferred to a clean container and frozen overnight at −20 °C. The next day, the fluid was thawed and centrifuged at 15,000 × *g* for 15 minutes at 4 °C. The final supernatant was used for media preparation.

2.5.5. Anaerobic glycerol solution

For a 500 mL anaerobic glycerol solution (40%), 85 mL of Salt Solution A was combined with 85 mL of Salt Solution B (Bauchop, 1979), and 130 mL of distilled water with 200 mL of glycerol. Solution A contained (wt./vol): KH2PO4, 0.3%; NaCl, 0.6%; (NH4)2SO4, 0.3%; CaCl2, 0.03%; and MgSO4, 0.03%. Solution B contained 0.3% (wt./vol) K2HPO4. $NaHCO_3$

(2.5 g) was added and 2 drops of 0.1% resazurin solution. All components were mixed in an Erlenmeyer flask and heated in a microwave oven until boiling. The solution was sparged with O_2-free 100% CO_2, cooled in an ice bath under a stream of CO_2 and 250 mg of L-cyst-HCl was added. Aliquots (75 mL) of the solution were dispensed into CO_2-filled 125 mL serum bottles and autoclaved for 20 minutes at 121°C and 15 psi. Anaerobic glycerol solution was added to cultures to give a final glycerol concentration of 10%.

2.6 Cultivation of rumen microbes using TMA precursors as substrates

2.6.1. Optimising growth medium for selective growth

Modified RM02 medium (Kenters et al., 2011) was optimised to support the selective growth of TMA-producing microbes. The medium was modified to include clarified rumen fluid (Kenters et al., 2011), added at 5% of the final volume before dispensing. This simplified the process of adding the fluid to each tube by filter sterilisation into autoclaved medium tubes. TMA precursors, including choline chloride, betaine, and L-carnitine (all from Sigma-Aldrich), were added at a final concentration of 25 mM to the medium before dispensing and autoclaving. The media were dispensed in 5 mL volumes into small-sized 10 mL Hungate tubes, sealed with rubber bungs and screw caps, while maintaining anaerobic conditions, and then autoclaved.

2.6.2. Animal ethics and sample collection

Rumen and faecal samples were collected under the approval of the AgResearch Animal Ethics Committee (AE 2074). Rumen fluid was obtained *via* stomach tubing from pasture-fed 6-month-old dairy heifer calves, (tagged 126 and 130), participating in trial CCA 01-W26. Hereafter, the rumen samples from these calves are referred to as R126 and R130, respectively. Faecal samples were also collected as part of the same study from calf F124. All sampling procedures were conducted at the AgResearch Grasslands Ulyatt-Reid Facility in Palmerston North.

2.6.3. Enrichment experiment

To enrich for putative TMA-producing microbes using the medium described above (Section 2.5.3), the modified RM02 was supplemented with 25 mM choline, betaine, or L-carnitine. Rumen samples were kept on ice and added to anaerobic media immediately after transferring to the laboratory. For each sample, 1 g of rumen content was diluted 10-fold by adding 9 mL of the anaerobic RM02 base (Kenters et al., 2011) without clarified rumen fluid or substrates, followed by 10-fold serial dilution to 10^{-5} in Hungate tubes. Due to the low abundance of TMA producers expected (Kelly et al., 2019; Rath et al., 2020), dilutions 10^{-2} to 10^{-5} were performed for rumen and faecal samples and inoculated into RM02 with the substrate. Tubes were incubated at 39 °C, and the optical density at 600 nm (OD_{600}) was measured the next day after 18 hours of incubation to assess growth. From enrichments, 2 mL of sample was collected from each dilution per treatment for TMA, short-chain fatty acid (SCFA), and alcohol quantification, and the remaining sample was preserved as a 10% glycerol stock at –80 °C.

2.6.4. Isolating pure cultures from enrichment cultures

Microbial isolations were performed using RM02 agar supplemented with choline or betaine. To isolate colonies, selective agar was inoculated using a glycerol stock of preserved enrichment culture. The 10^{-2} stock tube was diluted in a serial 10-fold dilution up to a 10^{-10} dilution, and 100 µL of each dilution was plated. Plates were incubated at 39°C under anaerobic conditions and monitored for colony development. Plates exhibiting isolated colonies were selected for further analysis, while those showing a lawn of bacterial growth were excluded. Individual colonies were picked using a sterile 1 µL loop, with a thin smear applied to a designated spot on a MALDI target plate. The remaining sample on the loop was patched onto agar plates and labelled to correspond with the MALDI plate location. Direct protein extraction was performed on the MALDI plate samples for profiling via MALDI-TOF MS, and the resulting spectra were analysed using reference databases for strain identification (Figure 7).

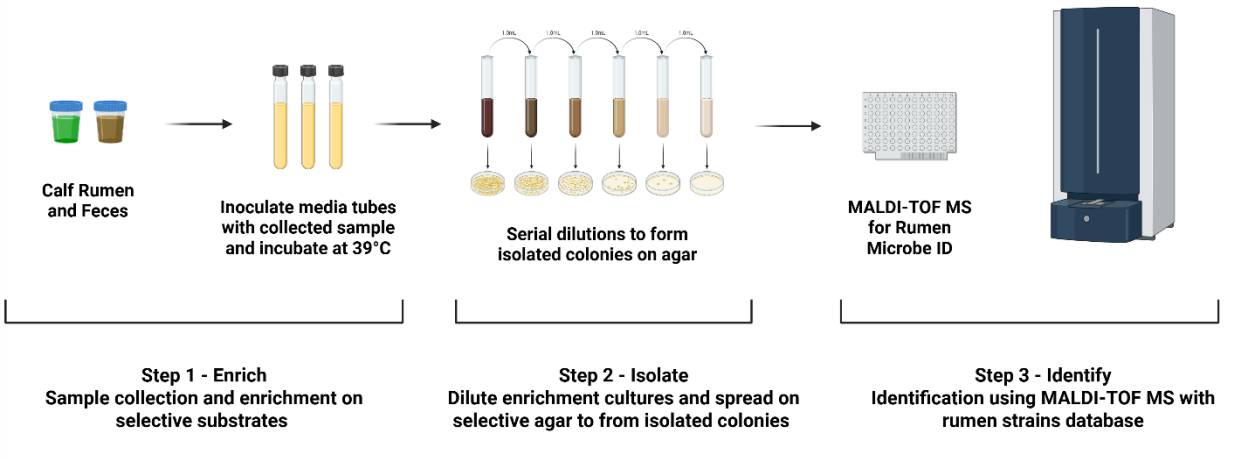


Figure 7. Schematic representation of microbial enrichment, isolation, and high-throughput screening for microbial identification of new microorganisms that can grow on TMA precursors.

2.7 16S rRNA gene sequencing and analysis

Pure cultures were subjected to 16S rRNA gene amplification and sequencing for taxonomic identification. Amplification was performed using the Platinum™ Hot Start PCR Master Mix (2X) (Invitrogen, Thermo Fisher Scientific, Auckland, NZ, Cat. No. 13000014) following the manufacturer's protocol. Universal bacterial primers 27F (GAG TTT GAT CMT GGC TCA G) (Lane, 1991) and 1492R (GGY TAC CTT GTT ACG ACT T) (Loy et al., 2002) were used to amplify approximately 1465 bp of the 16S rRNA gene. Two µl of culture was added directly to the PCR reaction as template, which included an initial cell lysis step at 95°C for 10 minutes. Then, 30 PCR cycles for denaturation at 94 °C (15 sec), annealing at 60 °C (15 sec) and extension at 68 °C (30 sec).

PCR products were resolved in 1.5% agarose gels in 1× TAE buffer containing SYBR® Safe DNA stain (Invitrogen, Thermo Fisher Scientific, Auckland, NZ) and run for 30 minutes at 110 V . Amplified DNA bands were visualised under UV light, and a 1 Kb Plus DNA ladder (Invitrogen, Thermo Fisher Scientific, Auckland, NZ) was used for size estimation.

Amplicons were purified using the QIAquick PCR Purification Kit (Qiagen Ltd, Auckland, NZ) and sequenced at the Massey Genome Service (Massey University, Palmerston North, NZ)

using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Thermo Fisher Scientific, Auckland, NZ). Raw ab1 files were quality-trimmed and assembled into consensus sequences using Geneious Prime® 2023.0.1, an integrated bioinformatics software platform used for sequence analysis. Final sequences were queried using the NCBI BLAST sequence server against the GenBank nr/nt database and the 16S rRNA database using BLASTN to identify closely related taxa excluding uncultured/environmental sample sequences.

2.8 Quantification of TMA, VFA, and alcohol compounds

2.8.1. High-performance liquid chromatography (HPLC)

The HPLC (Shimadzu, LC-2040C 3D) chromatographic separation was performed using a Coresep 100 column (4.6 × 150 mm, 2.7 µm, 90 Å), employing both reversed-phase and cation-exchange modes. The mobile phase consisted of 15% ACN with 0.3% trifluoroacetic acid (TFA), delivered at a flow rate of 0.6 mL/minute. A standard concentration range from 5 to 25 mM was calibrated with an injection volume of 2 µL. Sample detection was carried out using a refractive index detector maintained at 40 °C. One mL of each sample was centrifuged at 15,000 × *g* at 4°C for 5 minutes and filtered through 0.2-µm nylon membrane filters into 1.5 mL vials.

2.8.2. Volatile Fatty Acid (VFA) and alcohol analysis by gas chromatography

For VFA analysis using gas chromatography (GC) (Rajasekaran, 2023) 1 mL of sample was centrifuged at 10,000 × *g* for 20 minutes at 4 °C, and 900 µL of supernatant was mixed with 100 µL of 20 mM 2-ethylbutyrate in 20% phosphoric acid as an internal standard. Samples were frozen for a minimum of 24 hours before analysis. A 750 µL aliquot was transferred to a 1.5 mL glass GC vial and analysed using a Shimadzu GC-2010 Plus equipped with Flame Ionisation Detection and a Phenomenex Zebron ZB-FFAP column (30 m × 0.53 mm I.D. × 1.00 µm film thickness). The injection volume was 1 µL, with an injection temperature of 90 °C and a detector temperature of 240 °C. Helium was used as the carrier gas. VFAs measured

included acetic, propionic, butyric, caproic, valeric, *iso*-butyric, and *iso*-valeric acids. The column temperature program involved sequential ramps from 60 °C to 200 °C, with a final hold at 200 °C.

Alcohols (methanol, ethanol, 1-propanol, 2-propanol, and butanol) were analysed using the same sample vial. A Shimadzu GC-2010 Plus with an AOC-6000 autosampler and a Restek Rtx-Volatiles column (60 m × 0.32 mm I.D. × 1.5 µm film thickness) was used. Samples were injected at 200 °C, and the FID detector was set to 230°C. The column temperature program ramped from 50°C to 230°C, with a final hold at 230°C. Helium was used as the carrier gas throughout (Rajasekaran, 2023).

Chapter 3. Results: Library creation for database expansion

3.1 MALDI Biotyper® identification for putative choline-utilising candidate strains

Choline-utilising candidate strains (Kelly et al., 2019) were screened using MALDI Biotyper®. This screen evaluated the efficiency of target preparation and identification status using the Bruker reference database, MALDI Biotyper® Compass Library Revision H (2021). A total of 11 strains, summarised in Table 2, present the identification results for evaluating the need for library expansion. All samples prepared using the extended protein extraction method (Section 2.2.4) resulted in the successful generation of spectra required for comparing against the MSPs in the database.

The identification analysis revealed that seven strains were correctly identified at the species level. MALDI screening of *Olsenella umbonata* and two strains of *Proteus mirabilis* showed consistency with their nomenclature. However, some nomenclature discrepancies were noted due to recent reclassifications of certain genera. The identified discrepancies were verified by searching the Genome Taxonomy Database (GTDB) for updated taxonomy, which was then compared with the MALDI results, as shown in Table 2. The meaning of the Bruker MALDI Biotyper® identification scores is described in Appendix 3.

Concerningly, *Eubacterium* sp. AB3007 matched to an unrelated taxon, *Pasteurella multocida*, with a score above 1.7, resulting in misidentification, because of high spectral similarity or poor spectral representation in the current database. Additionally, the isolates *Desulfovibrio desulfuricans, Desulfovibrio legallii,* and *Lachnoclostridium lavalense* failed to yield any reliable identification, highlighting gaps in the current reference database. Hence, database expansion with new MSP libraries will be required for these strains.

Table 2. Identification of the choline-utilising candidate strains using MALDI Biotyper®

Candidate strains (Kelly et al., 2019)	1st best match	1st match score*	2nd best match	2nd match score*	GTDB Taxonomy
***Lachnoclostridium clostridioforme* AGR2157**	*Enterocloster clostridioformis*	2.13	*Enterocloster bolteae*	2.12	*Enterocloster clostridioformis*_A
***Eubacterium* sp. AB3007**	*Pasteurella multocida*	1.84	*Pasteurella multocida*	1.74	*Hornefia* sp000688015
***Lachnoclostridium aerotolerans* X8A62 (DSM 5434)**	*Lacrimispora aerotolerans*	2.20	Unreliable	1.38	*Lacrimispora aerotolerans*
***Olsenella umbonata* A2 (DSM 22619)**	*Olsenella umbonata*	2.27	*Olsenella umbonata*	1.83	*Parafannyhessea umbonata*
***Enterococcus* sp. KPPR-6**	*Enterococcus malodoratus*	2.32	*Enterococcus malodoratus*	2.16	*Enterococcus*_A *malodoratus*
***Desulfovibrio desulfuricans* G11 (DSM 7057)**	Unreliable	1.48	Unreliable	1.44	*Desulfovibrio* sp900243745
***Proteus mirabilis* NLAE-zl-G534**	*Proteus mirabilis*	2.57	*Proteus mirabilis*	2.56	*Proteus mirabilis*
***Lachnoclostridium lavalense* NLAE-zl-G277**	Unreliable	1.43	Unreliable	1.36	*Enterocloster lavalensis NLAE-zl*-G277
***Lachnoclostridium citroniae* NLAE-zl-G70**	*Enterocloster citroniae*	2.35	*Enterocloster citroniae*	2.26	*Enterocloster citroniae*
***Desulfovibrio legallii* KHC7**	Unreliable	1.28	Unreliable	1.25	*Desulfovibrio legallii_A*
***Proteus mirabilis* NLAE-zl-C285**	*Proteus mirabilis*	2.60	*Proteus mirabilis*	2.45	*Proteus mirabilis*
*Score interpretation: Species level: A score of ≥2.0, Genus level: A score between ≥1.7 and ≤1.99, Unreliable: A score of <1.7 (Panda et al., 2014).					

3.2 Retroversion potential of TMAO back to TMA in H1K culture collection

In work carried out before this study (unpublished), the H1K genomic dataset was screened for potential candidates involved in TMA and TMAO utilisation. The heat map in Figure 8, illustrates the distribution and relative abundance of key enzymes involved in TMAO metabolism within the H1K genomic dataset. Among the analysed KEGG Orthology (KO) identifiers K07811 and K07812, enzymes with potential for TMAO reversion were rarely present among H1K genomic database. K18277 (Trimethylamine monooxygenase, tmm) enzyme, which catalyses the oxidation of TMA to TMAO, was also rarely present. In contrast, K00605 (Aminomethyltransferase, AMT), which plays a role in glycine degradation and broader amino acid metabolism, was widely distributed across the dataset. A total of 111 strains were identified during this analysis and are listed in the heatmap (Figure 8), with *Prevotella* and *Bacteroides* being the most prevalent genera identified to carry AMT enzyme-coding genes.

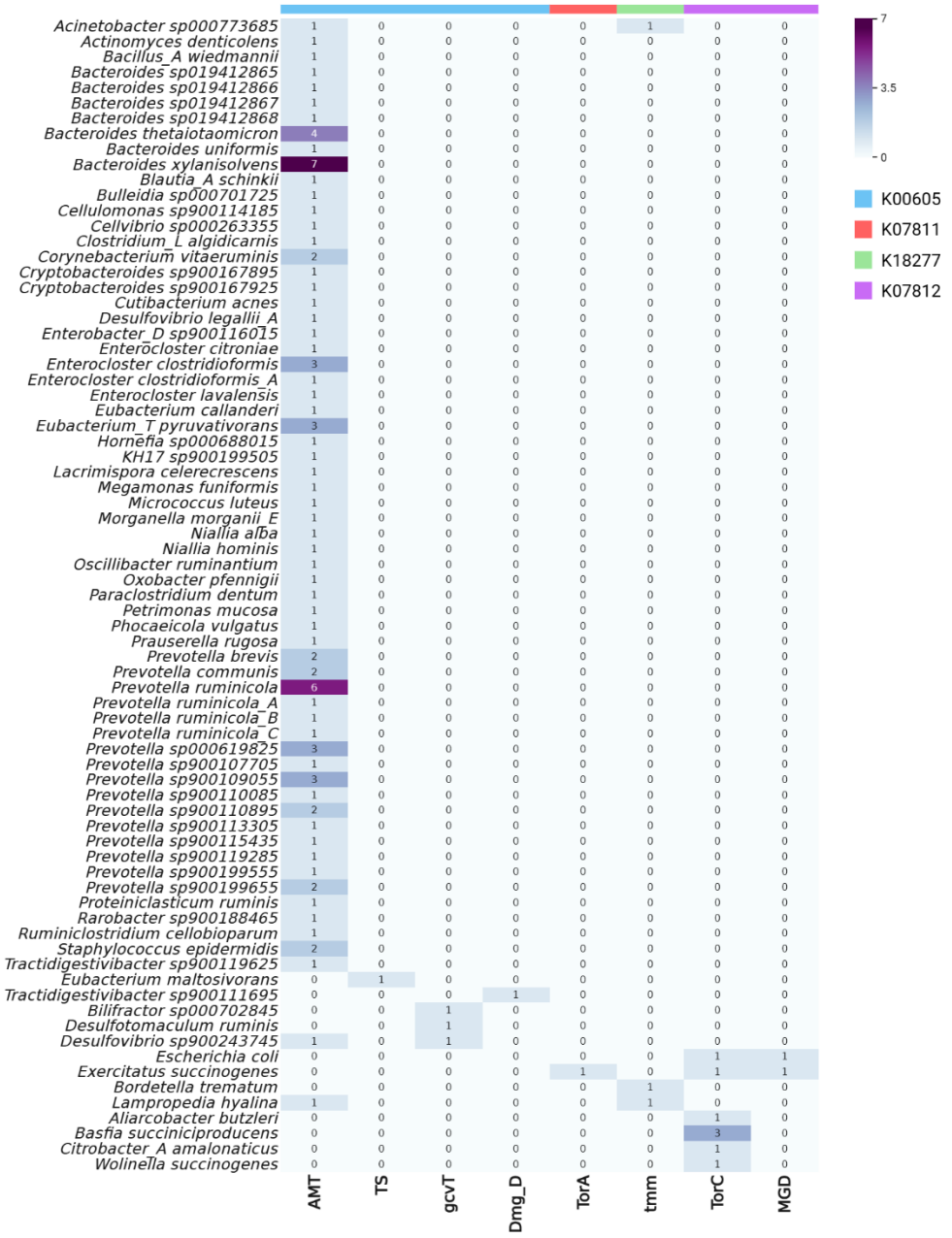


Figure 8. Heat map depicting the distribution of TMAO metabolising enzyme proteins in H1K rumen microbial cultures. The count represents the number of copies of the genes present for each enzyme. Enzymes presented on the x-axis include aminomethyltransferase (AMT), dimethylglycine dehydrogenase (Dmg_D), glycine cleavage system T protein (gcvT), molybdopterin guanine dinucleotide-containing S/N-oxide reductase (MGD), 4-hydroxy-tetrahydrodipicolinate synthase (TS), trimethylamine monooxygenase (tmm), and trimethylamine-N-oxide reductase (TorA and TorC), and bacterial species are at the y-axis. The heatmap was created in https://BioRender.com.

3.3 Gaps in identifying strains from rumen origin

The growth of candidate strains was tested on anaerobic commercially-prepared agar plates (BHI, CSBA, and FAA), used in other studies that created a protein fingerprint database for anaerobic strains with clinical relevance (Asare et al., 2023). These media did not consistently support good growth of the rumen candidate strains (Sections 3.1 and 3.2), which resulted in insufficient protein extraction for MALDI-TOF analysis. M2GCS agar medium (Miyazaki et al., 1997) a rumen bacteria-specific growth medium supported the growth of most candidate strains and as a result enabled successful protein extraction and spectral generation.

The identification of the candidate strains using the MALDI Biotyper® system with the revision H database is presented in Table 3. Fifteen strains were identified at the species level with scores >2 and another five strains obtained match scores >2 but with nomenclature discrepancies. All strains were searched for updated taxonomy using GTDB which verified MALDI identifications for reclassified strains. One strain was identified at the genus level with a score above 1.7. Nomenclature discrepancies were found because Hungate strains were classified in 2018 based on 16S rRNA databases at the time, while the revision H Bruker database was updated in July 2021. However, fifteen rumen strains remained unidentified, as summarised in Figure 9.

Table 3. Screening results for the identification of candidate strains using the Bruker database revision H

Candidate strains	Assession/ GTDB Taxonomy	Identification with Bruker database	Score*	Match level/ Interpretation
***Bacteroides ovatus* NLAE-zl-C500**	*GCA*_019412865.1/ *Bacteroides sp019412865*	*Bacteroides ovatus*	1.77	Genus-level
***Lachnoclostridium citroniae* NLAE-zl-G70**	GCA_900115855.1/ *Enterocloster citroniae*	*Enterocloster citroniae*	2.35	Species-level
***Lachnoclostridium clostridioforme* NLAE-zl-G208**	GCF_900100685.1/ *Enterocloster clostridioformis*	*Enterocloster clostridioformis*	2.45	Species-level
***Lachnoclostridium clostridioforme* NLAE-zl-C196**	GCF_900100685.1/ *Enterocloster clostridioformis*	*Enterocloster clostridioformis*	2.38	Species-level
***Lachnoclostridium clostridioforme* AGR2157**	GCF_000424325.1/ *Enterocloster clostridioformis_A*	*Enterocloster clostridioformis*	2.13	Species-level
***Lachnoclostridium aerotolerans* X8A62 (DSM 5434)**	GCF_000687555.1/ *Lacrimispora aerotolerans*	*Lacrimispora aerotolerans*	2.20	Species-level
***Olsenella umbonata* A2**	GCA_016901295.1/ *Parafannyhessea umbonata*	*Olsenella umbonata*	2.27	Species-level
***Bacteroides thetaiotaomicron* KPPR-3**	GCA_900109385.1/ *Bacteroides thetaiotaomicron*	*Bacteroides thetaiotaomicron*	2.36	Species-level
***Proteus mirabilis* NLAE-zl-G534**	GCF_900113495.1/ *Proteus mirabilis*	*Proteus mirabilis*	2.57	Species-level
***Bacteroides thetaiotaomicron* NLAE-zl-C579**	GCF_900108155.1/ *Bacteroides thetaiotaomicron*	*Bacteroides thetaiotaomicron*	2.44	Species-level

***Proteus mirabilis* NLAE-zl-C285**	GCF_900101725.1/ *Proteus mirabilis*	*Proteus mirabilis*	2.60	Species-level
***Bacteroides ovatus* NLAE-zl-C57**	GCF_900100465.1/ Bacteroides sp019412865	*Bacteroides ovatus*	2.04	Species-level
***Morganella morganii* NLAE-zl-C84**	GCF_900142745.1/ *Morganella morganii*_E	*Morganella morganii*	2.27	Species-level
***Megamonas* sp. Calf98-2**	GCF_900110235.1/ *Megamonas funiformis*	*Megamonas* sp[2]	2.12	Species-level
***Shigella sonnei* NLAE-zl-G496**	GCF_900103855.1/ *Escherichia coli*	*Escherichia coli*	2.44	Species-level
***Bacteroides xylanisolvens* NLAE-zl-C202**	GCF_900114865.1/ *Bacteroides xylanisolvens*	*Bacteroides ovatus*	2.34	Species-level
***Bacteroides* sp. AR20**	GCF_900109965.1/ *Bacteroides thetaiotaomicron*	*Bacteroides thetaiotaomicron*	2.45	Species-level
***Enterococcus* sp. KPPR-6**	GCF_900111415.1/ *Enterococcus*_A *malodoratus*	*Enterococcus malodoratus*	2.32	Species-level
***Enterobacter* sp. KPR-6**	GCF_900116015.1/ *Enterobacter*_D sp900116015	*Enterobacter kobei*	2.02	Species-level
***Olsenella* sp. KH3B4**	GCF_900111695.1/ *Tractidigestivibacter* sp900111695	*Olsenella scatoligenes*	2.13	Species-level
***Citrobacter* sp. NLAE-zl-C269**	GCF_900112055.1/ *Citrobacter*_A *amalonaticus*	*Citrobacter amalonaticus*	2.34	Species-level
***Eubacterium* sp. AB3007**	GCF_000688015.1/ *Hornefia* sp000688015	None		Unidentified
***Lachnospiraceae bacterium* FE2018**	GCA_000702845.1/ *Bilifractor* sp000702845	None		Unidentified
***Prevotella brevis* P6B11**	GCF_000621825.1/	None		Unidentified

	Prevotella brevis			
***Blautia schinkii* DSM 10518**	GCA_000702025.1/ *Blautia_A schinkii*	None		Unidentified
Proteiniclasticum ruminis DSM 24773	GCF_000701905.1/ *Proteiniclasticum ruminis*	None		Unidentified
Prevotella sp. RM4	GCF_000701965.1/ *Prevotella ruminicola*	None		Unidentified
***Basfia succiniciproducens* KPR-2**	GCF_900110905.1/ *Basfia succiniciproducens*	None		Unidentified
***Desulfovibrio desulfuricans* G11 (DSM 7057)**	*Desulfovibrio desulfuricans* G11 (DSM 7057)	None		Unidentified
***Bacteroidales bacterium* WCE2008**	GCA_900167925.1/ *Cryptobacteroides* sp900167925	None		Unidentified
***Cellulomonas* sp. KH9**	GCF_900114185.1/ *Cellulomonas* sp900114185	None		Unidentified
***Prevotellaceae bacterium* KH2P17**	GCA_900199655.1/ *Prevotella* sp900199655	None		Unidentified
***Lachnoclostridium lavalense* NLAE-zl-G277**	GCF_900102595.1/ *Enterocloster* lavalensis	None		Unidentified
***Desulfovibrio legallii* KHC7**	GCF_900102485.1/ *Desulfovibrio* legallii_A	None		Unidentified
***Porphyromonadaceae bacterium* KHP3R9**	GCA_900116745.1/ *Petrimonas mucosa*	None		Unidentified
***Prevotella ruminicola* D31d**	GCF_900107775.1/ *Prevotella ruminicola*_C	None		Unidentified
*Score interpretation: Species level: A score of ≥2.0, Genus level: A score between ≥1.7 and ≤1.99, Unreliable: A score of <1.7 (Panda et al., 2014)				

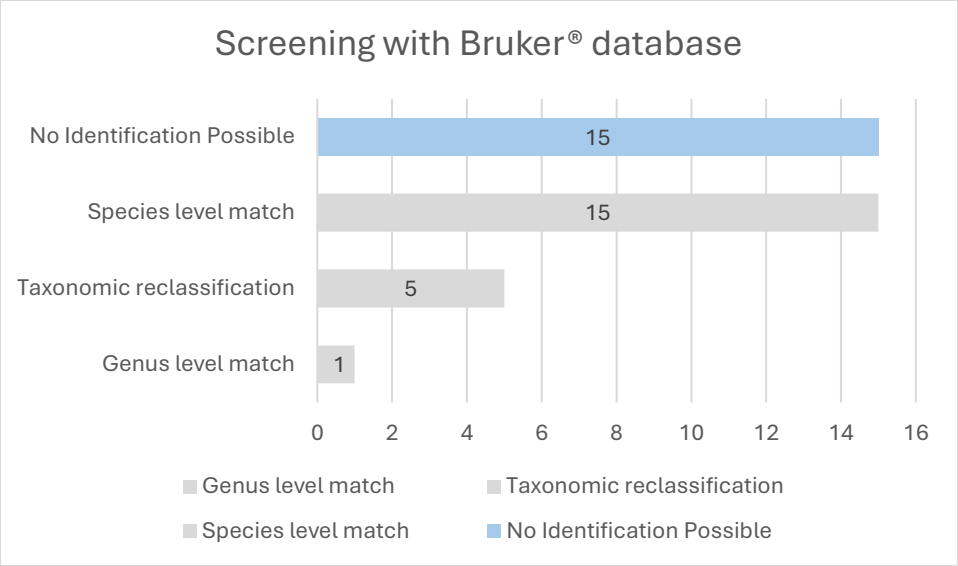


Figure 9. Summary of candidate strains identification matches using Bruker MALDI-Biotyper® database revision H.

3.4 New mass spectral profiles and verification

To expand the database by addressing gaps in spectral profiles, closely related genera and species absent from the existing library were selected. New spectral profiles were created using high-quality protein extracts obtained via the extended protein extraction method (Section 2.2.4). Each MSP underwent quality control and verification before inclusion in the reference library. Appendix 1 presents quality control data for passed allowed mass tolerance for seven peaks from a mass range of 2000 Da to 10,000 Da for each MSP. Verification was performed using the direct transfer method (Section 2.23), which resulted in successful generation of spectra and species-level identification scores detailed in Table 4.

Table 4. Validated spectral profiles of newly added candidate strains using the direct protein extraction method

Organism	Identification score before creating a new MSP	Verification Score with New MSP
***Eubacterium* sp. AB3007**	unidentified	2.06
***Lachnospiraceae bacterium* FE2018**	unidentified	2.19
***Prevotella brevis* P6B11**	unidentified	2.54
***Blautia schinkii* DSM 10518**	unidentified	2.54
***Proteiniclasticum ruminis* DSM 24773**	unidentified	2.27
***Prevotella* sp. RM4**	unidentified	2.46
***Basfia succiniciproducens* KPR-2**	unidentified	2.37
***Desulfovibrio* desulfuricans G11 (DSM 7057)**	unidentified	2.71
***Bacteroidales bacterium* WCE2008**	unidentified	2.39
***Cellulomonas* sp. KH9**	unidentified	2.29
***Prevotellaceae bacterium* KH2P17**	unidentified	2.41
***Lachnoclostridium lavalense* NLAE-zl-G277**	unidentified	2.42
***Desulfovibrio legallii* KHC7**	unidentified	2.69
***Porphyromonadaceae bacterium* KHP3R9**	unidentified	2.47
***Prevotella ruminicola* D31d**	unidentified	2.49
***Citrobacter* sp. NLAE-zl-C269**	2.27	2.36
***Bacteroides ovatus* NLAE-zl-C500**	1.98	2.49
***Olsenella* sp. KH3B4**	1.98	2.38
***Enterobacter* sp. KPR-6**	1.91	2.74

3.5 Integration of new spectral profiles into the MSP dendrogram

To determine the taxonomic relatedness between the candidate strains and reference strains, a dendrogram was generated using the spectral profiles (Figure 10). The MSPs from rumen isolates listed in Table 4 are positioned among selected MSPs that belong to the same genus present in the Bruker database. The placement of MSPs from rumen origin isolates indicates the presence of novel spectral profiles that were previously missing in the taxonomy tree of Bruker reference database revision H. The dendrogram shows that the new MSPs, particularly those clustered at near-zero distance levels, exhibited high similarity within a species taxon. The hierarchical structure is based on protein spectral similarity, with closely related strains forming tight, clusters such as *Desulfovibrio* (shown in red), *Cellulomonas* (green), and *Porphyromonas/ Prevotella* (various colours), emphasising intra-cluster similarity and clear separation between different clusters. The rumen-based *Prevotella* strains form a distinct spectral cluster from the existing spectral profiles of non-rumen *Prevotella* strains.

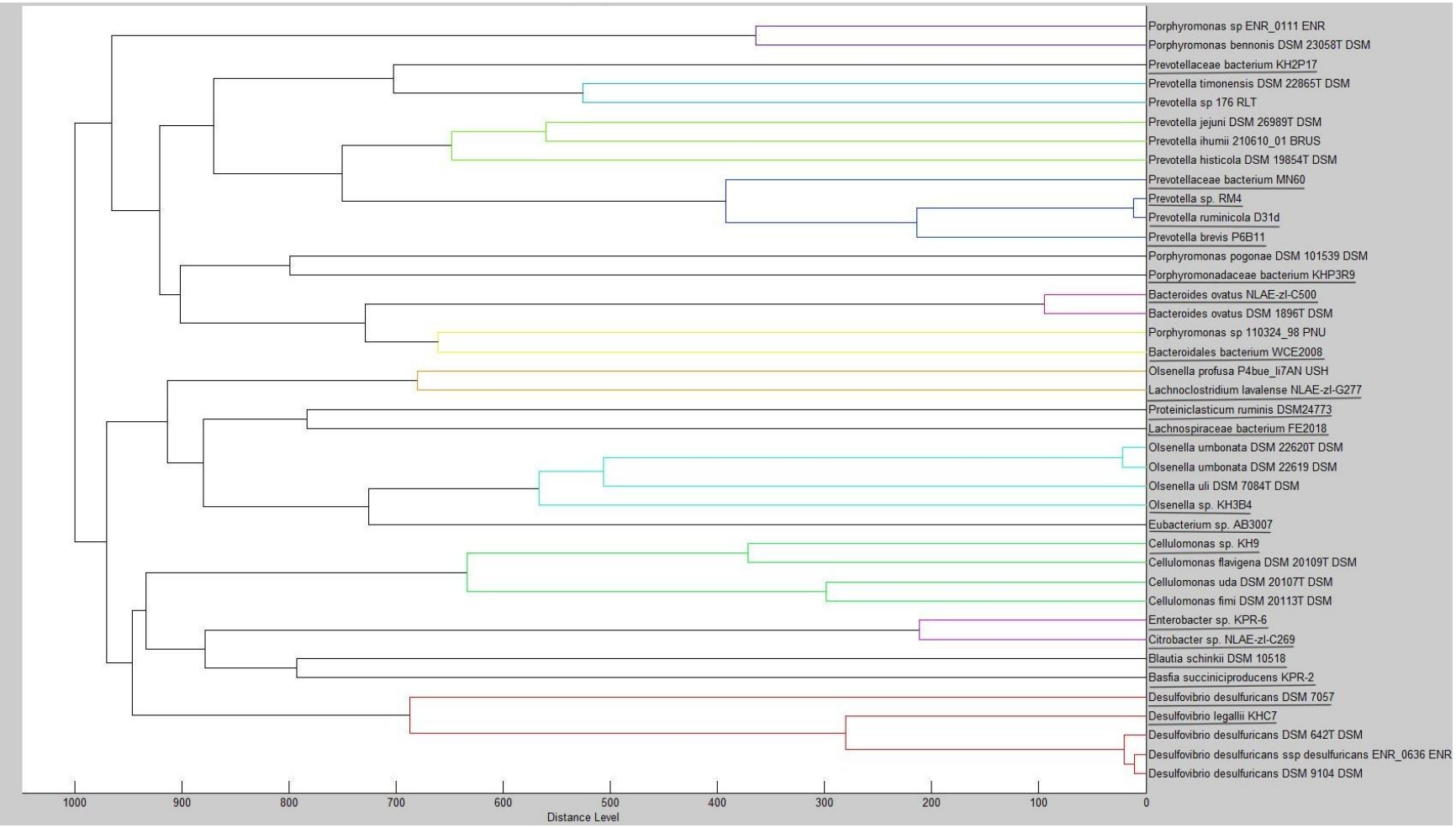


Figure 10. Hierarchical dendrogram illustrates distinctly placed spectral profiles of candidate strains (underlined) compared to existing MSPs in the Bruker database. The X-axis represents distance levels, and the Y-axis presents microbial taxa. Their colour-coded branches show clustering based on spectral similarity.

Chapter 4. Results: Selective cultivation of TMA-producing microbes

4.1 Choline selective growth confirmation by candidate strains

To evaluate substrate-specific microbial growth, candidate strains were cultivated on modified RM02 medium supplemented with either choline or betaine. Among the tested strains, *Olsenella umbonata* A2 demonstrated the highest growth on choline, indicating its ability to utilise this substrate (Figure 11). This observation aligns with previous genomic findings by Kelly et al. (2019), which identified the presence of the choline trimethylamine lyase (*cutC*) and its activating protein (*cutD*) in *O. umbonata*. However, confirmation of TMA production from choline metabolism requires further validation through metabolite quantification.

In contrast, strains not previously associated with choline metabolism, such as *Blautia schinkii* DSM 10518, *Basfia succiniciproducens* strains DSM 22022 and KPR-2, *Megamonas* sp. Calf98-2, *Shigella sonnei* NLAE-zl-G496, and *Lachnoclostridium clostridioforme* strains NLAE-zl-G208 and NLAE-zl-C196 exhibited low growth levels comparable to the no-substrate control, suggesting minimal or no choline utilisation under the tested conditions. Betaine-supplemented cultures showed no growth enhancement across all strains, with growth levels similar to the no-substrate control. This suggests limited utilisation of betaine under the experimental conditions. Cultures grown in a mixture of sugars–choline–betaine exhibited elevated optical density, attributed primarily to sugar metabolism rather than utilisation of choline or betaine. This was evident when compared to the no-substrate control, which showed minimal growth in the absence of fermentable sugars.

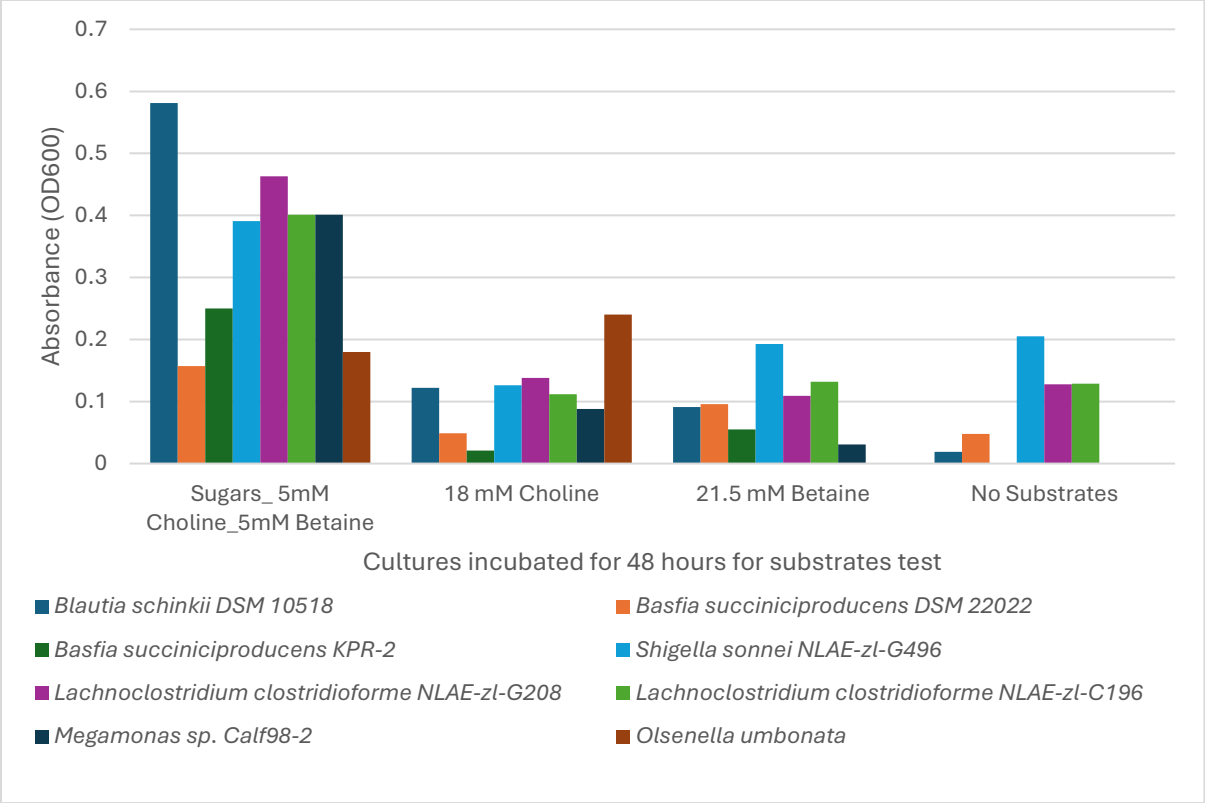


Figure 11. Bar graph depicting growth on modified RM02 medium with choline and betaine. The growth assay included a positive control with medium containing a sugar mix with choline and betaine, and a negative control with no substrates. Cultures were incubated for 24 hours at 39 °C. The optical density was measured at OD_{600}. This pilot growth test was performed using a single tube of each culture, with the main objective to test growth on the chosen medium and substrates.

4.2 Selection for rumen and faecal microbiota that utilise choline, betaine and carnitine

Growth measurements were obtained from rumen and faecal material dilutions inoculated into RM02 containing the selective substrates choline, betaine, and carnitine. The rumen samples exhibited higher growth across all the substrates as compared to the faecal samples. The absorbance measurements at OD_{600} were higher in the rumen sample tubes with choline substrate after 24 hours of incubation, as compared to betaine (except for R126 betaine replicates at 10^{-2} dilution) and carnitine. Growth measurement in rumen sample tubes was observed to decrease with dilution for the choline substrate, but not much dilution effect was observed with betaine and carnitine substrates (Figure 12A and 12B). In contrast, the faeces samples showed low growth and less variation across the dilutions and substrates (Figure 12C).

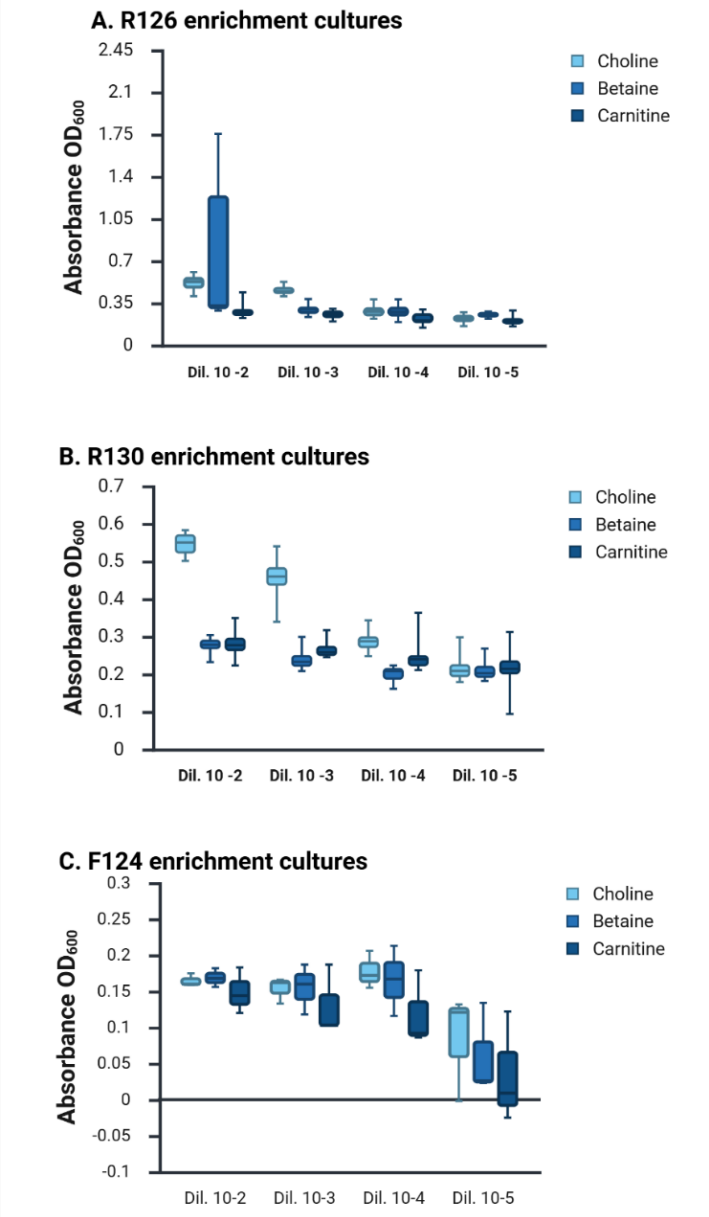


Figure 12. Box plots show the growth of two rumen samples (A) R126 and (B) R130, and one faecal sample (C) F124 across serial dilutions (ranging from 10^{-2} to 10^{-5}). The Y-axis shows absorbance (OD_{600}) and the X-axis shows rumen or faecal sample dilution for 10 replicates. Box plots were created using https://BioRender.com, and data is presented with the horizontal line on each box representing the median of the 10 replicates, and the edges of the box represent the lower and upper quarters, and the whiskers represent the extremes.

4.3 Pure culture isolations on choline and betaine agar

Pure isolates were obtained from the choline and betaine enrichment cultures, where each isolate was derived from the lowest dilution of the rumen sample. I could not purify isolates from carnitine enrichment cultures due to time constraints.

To obtain pure isolates, choline agar plates were inoculated with serial 10-fold dilutions of the glycerol-preserved enrichment stock. Well-separated colonies were observed at dilutions ranging from 10^{-4} to 10^{-10}. A total of 48 colonies were screened using the MALDI Biotyper®; 25 failed to generate detectable spectral peaks, and 17 produced spectra that could not be identified. Six colonies were successfully identified to the species level, with identification scores ≥ 2.0. Two isolates were classified as *Prevotella* species using the expanded custom-built reference database, while four matched *Streptococcus* species in the Bruker database. Table 5 summarises the MALDI-TOF identification results.

The results for betaine-selected isolates are also presented in Table 5. Betaine agar plates inoculated with enrichment stock dilutions showed overgrowth at lower dilutions, while well-isolated colonies were obtained at dilutions of 10^{-6} and higher. A total of 36 colonies were screened using the MALDI Biotyper®. Among these, 18 failed to produce spectral peaks, and 16 spectra were unidentifiable; only two targets were successfully identified, one at the species level and the other at the genus level.

Table 5. Identification results for choline and betaine selective isolates using the MALDI Biotyper® with the expanded reference database

MALDI identification database	Identification outcome (score)	Count of choline isolated targets	Count of betaine isolated targets
No identification		17	16
No peaks found		25	18
Bruker database	*Streptococcus infantarius* (1.97)	1	(1.82) 1
	Streptococcus lutetiensis (2.00)	2	
	Streptococcus lutetiensis (2.05)	1	
Expanded database	*Prevotella ruminicola* (2.01)	1	
	Prevotella ruminicola (2.23)	1	
	Bacteroidales bacterium (2.13)		1
Total targets screened		48	36

4.4 Grouping unidentified spectra based on similarity score

A substantial number of spectral profiles remained unidentified; there were 33 unidentified spectra out of 41 targets that generated spectra from both choline and betaine isolated colonies. To identify if the remaining unidentified colonies belonged to different genera, five colonies were chosen to obtain representative spectra using MALDI-TOF and labelled A to E. Then, the remaining unidentified spectra from the previous screens were re-analysed, with new add-on spectra along with the Bruker and expanded databases. Those spectra with a similarity score above 2 to these additional reference spectra (A to E) were grouped accordingly (Table 6).

Table 6. Grouping unidentified targets based on similarity score using novel spectra

Group of 5 chosen novel spectra labelled A to E	**Identification after re-analysis with added novel spectra**	
	Choline isolated matches	**Betaine isolated matches**
Matched spectrum A	6	none
Matched spectrum B	6	3
Matched spectrum C	none	2
Matched spectrum D	none	none
Matched spectrum E	1	8
Unidentified	4	3

Figure 13 illustrates the improvement in MALDI identification as the database was expanded further with add-on novel spectra of new isolates. Using only the Bruker MSP library resulted in the fewest number of isolates identified and expanding the database with MSPs for targeted TMA-metabolising H1K candidate strains only marginal improvement. Further inclusion of the spectra from the new reference isolates enhanced identification, as summarised in Figure 13. These findings demonstrate that adding new spectral profiles enhances the proportion of identified isolates through similarity-based grouping. However, 16S rRNA gene sequencing remains essential for confirming the identity of previously unidentified new isolates used as add-on spectra.

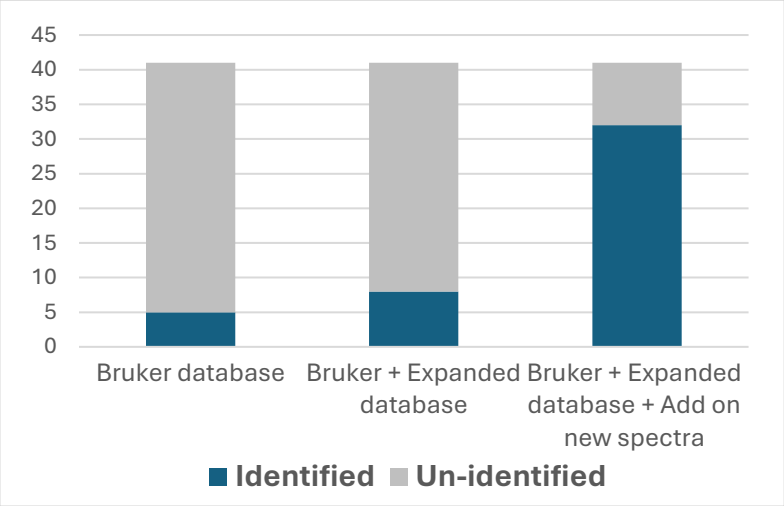


Figure 13. Grouping unknown samples by matching their spectra to improve identification.

4.5 Identification of pure isolates using 16S rRNA gene sequencing

To verify the MALDI identification results, 16S rRNA gene sequencing analyses were performed and are summarised in Table 7. Sequences were successfully generated using the universal forward primer 27F except for RCH22b, where the reverse primer 1492R was used due to poor-quality reads from the forward primer (Appendix 2). MALDI Biotyper® identification of the new isolates, including RB06, RCH10, and RCH11, was confirmed using 16S rRNA sequencing. Two isolates were identified as *Butyrivibrio hungatei* (RB04 and RB05)

and one isolate as *Slackia heliotrinireducens* (RB50_D). The identification of isolates using 16S rRNA gene sequencing with short sequence lengths (less than 200 bp) resulted in different hits using both databases (16S_ribosomal_RNA and core_nt). These sequences (<200bp) produced matching top blast hits to *Pseudobutyrivibrio ruminis* using the 16S rRNA database for the isolates RB47_A, RB48_B, RB49_C and RB51_E. In contrast, the core_nt database produced top blast hits to match various unclassified bacteria.

Table 7. Summarising the 16S rRNA gene sequencing results for the new rumen isolates.

Isolate ID	MALDI identification source	NCBI Standard Nucleotide BLAST	Top BLAST hit against the database*	% Pairwise Identity	E value	Accession	Query sequence (bases)
Rumen_Betaine_RB04**	Un-identified	16S_ribosomal_RNA	*Butyrivibrio proteoclasticus* strain B316 16S ribosomal RNA, partial sequence	98.55%	0	NR_102893.1	487
Rumen_Betaine_RB04**	Un-identified	core_nt	*Butyrivibrio* sp. CA23 gene for 16S ribosomal RNA, partial sequence	99.17%	0	AB849434.1	487
Rumen_Betaine_RB05**	Un-identified	16S_ribosomal_RNA	*Butyrivibrio proteoclasticus* strain B316 16S ribosomal RNA, partial sequence	98.14%	0	NR_102893.1	485
Rumen_Betaine_RB05**	Un-identified	core_nt	*Butyrivibrio* sp. P-18 16S rRNA gene, strain P-18	99.38%	0	AM039827.1	485
Rumen_Betaine_RB06**	Bruker database	16S_ribosomal_RNA	*Streptococcus lutetiensis* strain CIP 106849 16S ribosomal RNA, partial sequence	99.81%	0	NR_115719.1	515
Rumen_Betaine_RB06	Bruker database	core_nt	*Streptococcus* sp. strain Nyingchi-S2 16S ribosomal RNA gene, partial sequence	100.00%	0	OR072953.1	515
Rumen_Betaine_RB47_A**	Unidentified, novel spectra A	16S_ribosomal_RNA	*Pseudobutyrivibrio ruminis* strain DSM 9787 16S ribosomal RNA, partial sequence	97.24%	3.00E-65	NR_026315.1	143

Rumen_Betaine_RB47_A**	Unidentified, novel spectra A	core_nt	*Bacterium* MD2005 16S ribosomal RNA gene, partial sequence	97.90%	4.00E-62	KF698219.1	143
Rumen_Betaine_RB48_B**	Unidentified, novel spectra B	16S_ribosomal_RNA	*Pseudobutyrivibrio ruminis* strain DSM 9787 16S ribosomal RNA, partial sequence	97.96%	2.00E-67	NR_026315.1	146
Rumen_Betaine_RB48_B**	Unidentified, novel spectra B	core_nt	*Bacterium* LE2007 16S ribosomal RNA gene, partial sequence	98.65%	2.00E-65	KF698103.1	146
Rumen_Betaine_RB49_C**	Unidentified, novel spectra C	16S_ribosomal_RNA	*Pseudobutyrivibrio ruminis* strain DSM 9787 16S ribosomal RNA, partial sequence	100.00%	1.00E-58	NR_026315.1	121
Rumen_Betaine_RB49_C**	Unidentified, novel spectra C	core_nt	*Pseudobutyrivibrio ruminis* strain pC-XS7 16S ribosomal RNA gene, partial sequence	100.00%	2.00E-54	AF202262.1	121
Rumen_Betaine_RB50_D**	Unidentified, novel spectra D	16S_ribosomal_RNA	*Slackia heliotrinireducens* strain RHS1 16S ribosomal RNA, partial sequence	99.83%	0	NR_028733.1	588
Rumen_Betaine_RB50_D**	Unidentified, novel spectra D	core_nt	*Slackia heliotrinireducens* DSM 20476, complete genome	100.00%	0	CP001684.1	588
Rumen_Betaine_RB51_E**	Unidentified, novel spectra E	16S_ribosomal_RNA	*Pseudobutyrivibrio ruminis* strain DSM 9787 16S ribosomal RNA, partial sequence	96.07	1.00E-78	NR_026315.1	177
Rumen_Betaine_RB51_E*	Unidentified, novel spectra E	core_nt	*Bacterium* YGD2003 16S ribosomal RNA gene, partial sequence	98.88%	1.00E-82	KF698524.1	177

Rumen_Choline_RCH10**	Expanded database	16S_ribosomal_RNA	*Xylanibacter ruminicola* strain Bryant 23 16S ribosomal RNA, partial sequence	94.64%	0	NR_102887.1	444
Rumen_Choline_RCH10**	Expanded database	core_nt	*Bacterium* MA2016 16S ribosomal RNA gene, partial sequence	96.17%	0	KF698137.1	444
Rumen_Choline_RCH11**	Expanded database	16S_ribosomal_RNA	*Xylanibacter ruminicola* strain Bryant 23 16S ribosomal RNA, partial sequence	98.32%	0	NR_102887.1	595
Rumen_Choline_RCH11**	Expanded database	core_nt	*Prevotella ruminicola* gene for 16S rRNA, complete sequence, strain: BP1-76	98.83%	0	AB501161.1	595
Rumen_Choline_RCH22b**	Unidentified	16S_ribosomal_RNA	*Pseudobutyrivibrio ruminis* strain DSM 9787 16S ribosomal RNA, partial sequence	98.35%	6e-56	NR_026315.1	121
Rumen_Choline_RCH22b**	Unidentified	core_nt	*Rumen bacterium* NK4A555 16S ribosomal RNA gene, partial sequence	99.17%	3e-53	GU124464.1	121
Rumen_Choline_RCH22b_1492R***	Unidentified	16S_ribosomal_RNA	*Pseudobutyrivibrio ruminis* strain DSM 9787 16S ribosomal RNA, partial sequence	98.99%	0	NR_026315.1	397
Rumen_Choline_RCH22b_1492R***	Unidentified	core_nt	*Butyrivibrio* sp. strain FBDS c19 16S ribosomal RNA gene, partial sequence)	99.50%	0	KY643822.1	397

*Exclude: uncultured/environmental sample sequences
**Sequence derived from the universal 16S rRNA gene forward primer 27F
***Sequence derived from the universal 16S rRNA gene reverse primer 1492R

4.6 Alignment of novel spectra with MSPs in the expanded reference database

A MALDI-based dendrogram was generated to compare the unidentified spectral profiles, revealing distinct clustering patterns among existing MSPs in both the commercial Bruker and expanded databases. New additional spectra were underrepresented in the databases. In the dendrogram (Figure 14), spectrum D showed greater similarity to the *Bacteroidales bacterium* WCE2008 reference MSP and other novel spectra A–E were distinctly placed from the rest of the reference MSPs.

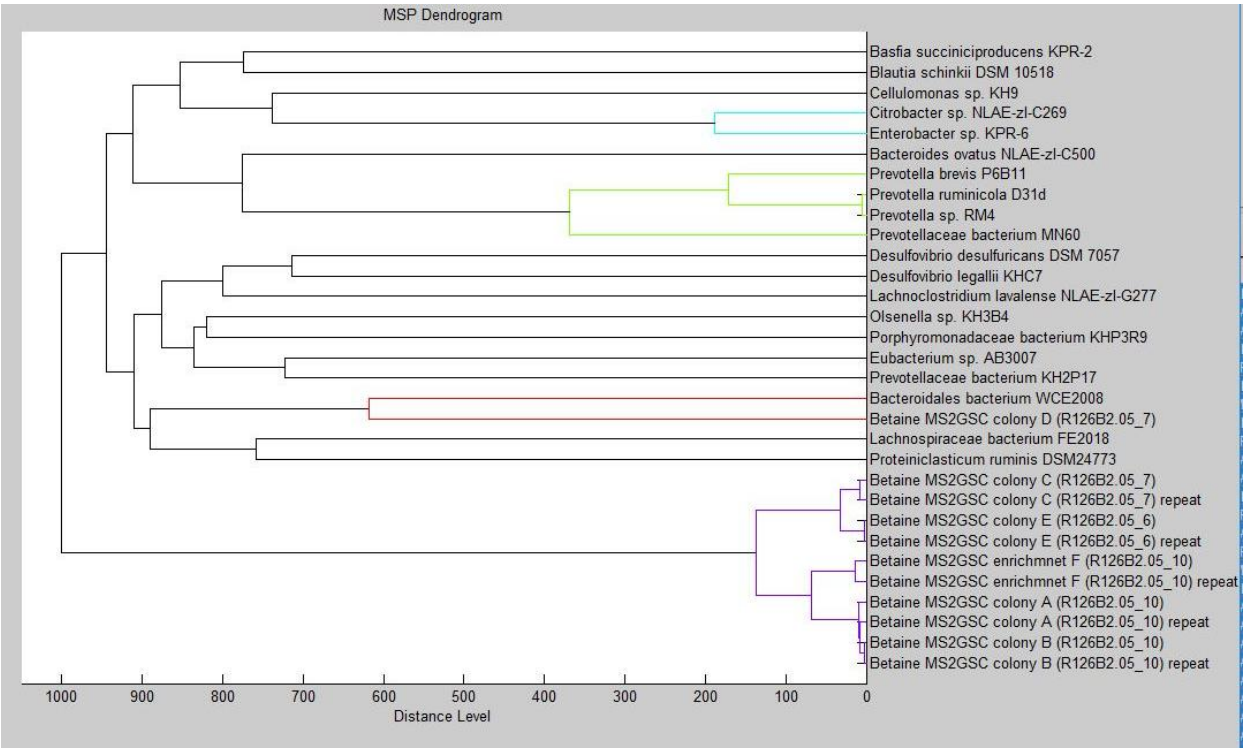


Figure 14. Alignment of unidentified additional spectral profiles with MSPs in the expanded reference database

4.7 Substrate utilisation and degradation products of rumen enrichment samples

Rumen microbial enrichment samples were quantified for substrate utilisation of choline, betaine, and carnitine using HPLC. Choline was completely utilised across all sample dilutions, but in contrast, betaine and carnitine showed minimal reduction in their initial concentrations (Figure 15).

To evaluate TMA-lyase activity in microbial enrichments, the enzymatic degradation products TMA, acetate, and ethanol were quantified. The results of HPLC quantification of TMA and GC quantification of ethanol and acetate are shown in Figure 15. HPLC revealed detectable levels of TMA in cultures supplemented with choline, whereas betaine enrichments showed only minimal TMA production at lower dilutions, and carnitine enrichments produced no detectable TMA.

Acetate was detected across all substrate conditions and was confirmed with GC analysis. The highest concentrations were observed in choline enrichments, followed by betaine and carnitine, which produced comparable levels of acetate in both substrates. GC analysis also confirmed ethanol production, with the highest levels found in microbial enrichments on choline. Betaine and carnitine treatments resulted in significantly lower ethanol concentrations. This result supports that choline is the primary precursor for TMA production in the tested rumen sample enrichments.

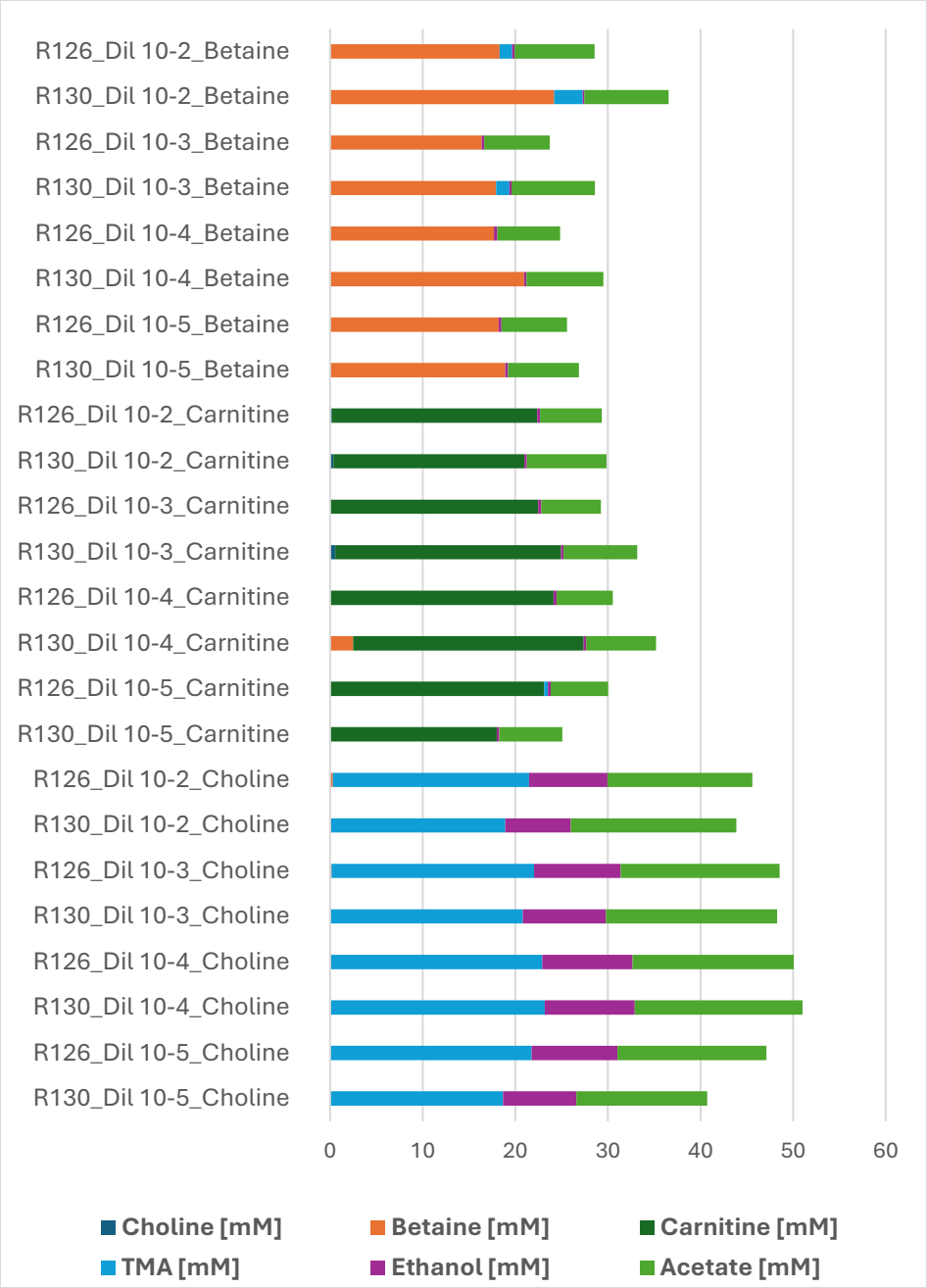


Figure 15. Bar graph illustrating substrate utilisation and metabolites of rumen microbial enrichment from R126 and R130. Methylated amines, including TMA, choline, betaine, and carnitine, were quantified using HPLC, and ethanol and acetate were quantified using GC. Concentrations are reported in mM.

4.8 Utilisation of choline and betaine in pure cultures

Supernatants from broth cultures of pure isolates obtained after colony purification were analysed using HPLC to assess substrate utilisation (Figure 16A and 16B). Isolates CH01, CH02, CH03, CH04 and CH17 from choline enrichments showed notable substrate depletion measured below 20 mM. Most samples exhibited only minor reductions in choline concentration compared to the original level, as shown in the control sample with medium and choline and no TMA was detected in any of the pure cultures. The same results were obtained for pure isolates obtained from betaine, where utilisation was minimal across most culture samples.

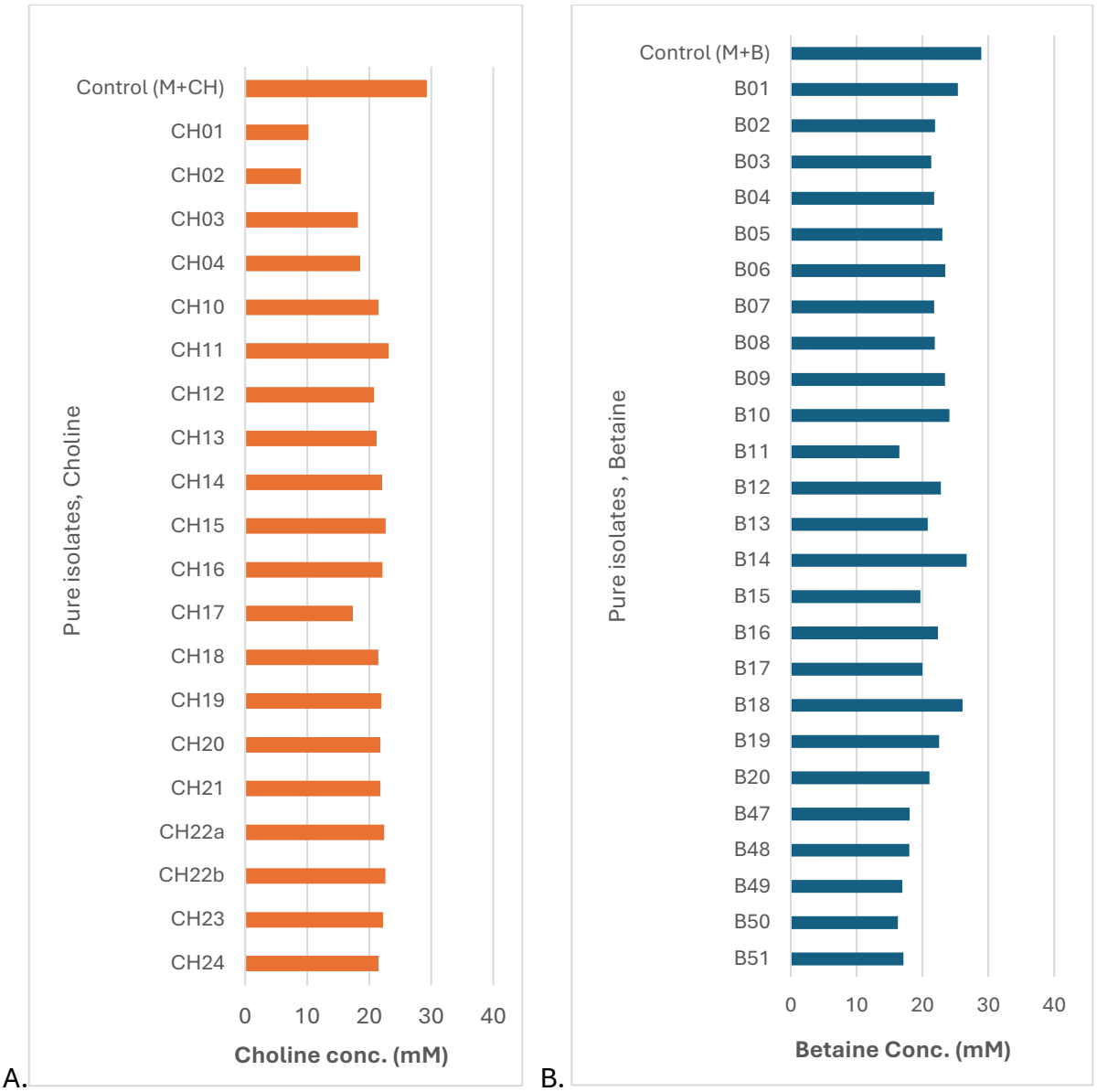


Figure 16. Quantification of substrate utilisation and TMA production in pure cultures of isolates from (A) choline and (B) betaine enrichment cultures. The bar chart displays residual choline and betaine concentrations (mM) after 24-hour incubation in modified RM02 medium supplemented with 25 mM of each TMA precursor.

Chapter 5. Discussion

This study aimed to isolate TMA-producing microbes from the rumen, facilitated using the MALDI Biotyper® databases that had been expanded to incorporate rumen microbes for rapid identification. This study resulted in successful enrichment for TMA-producing microbiota from choline, but TMA production was not detected in any of the purified isolates examined. The MALDI Biotyper® database was expanded with the MSPs created for selected rumen strains from the H1K culture collection (Seshadri et al., 2018), possessing genes involved in TMA metabolism. This study confirmed gaps in the current version of the Bruker database revision H, which lacked representation of anaerobic isolates from rumen origin. Discussion in this chapter highlights the key findings of this study, including successes, limitations, challenges, and future directions for the developed methodology.

5.1 The importance of MALDI Biotyper® database expansion

The use of the MALDI Biotyper® proved highly effective for rapidly identifying new isolates. The accurate microbial identification depended on the presence of the reference MSPs, which were previously absent for those rumen genera and species listed in Table 4. The ability to expand the MALDI Biotyper® database made it most suited for rapid microbial identification compared to 16S rRNA sequencing. Previous cultivation and H1K culture collection efforts represent the largest number of cultured representatives of rumen microbes available, along with their whole genome sequences (Seshadri et al., 2018). However, the remaining uncultured fraction of rumen microbial diversity needs to be cultivated. Cultivation efforts can be enhanced by the development of a high-throughput microbial identification method.

In this study, a targeted cultivation approach was explored, selecting strains from the H1K culture collection that are linked to TMA metabolism pathways. Initial screening of candidate strains confirmed significant gaps in the commercial Bruker database; the identification results presented in Table 3 highlighted under-represented rumen strains. The pilot study

successfully created a database specifically for rumen microbes to bridge this gap. The expanded database with 19 candidate strains successfully identified three new isolates, but most remained unidentified, suggesting the need to further expand the database to cover more diverse rumen genera and species.

To my knowledge, this study represents the first effort to expand Bruker's database by testing library creation methods on strictly anaerobic rumen strains. A previous study expanded the Bruker database with anaerobic strains of human clinical origin included in the revision H (Alcala et al., 2021; Asare et al., 2023). In this study, using the Bruker database revision H, a total of 20 candidate strains were identified in Table 3. However, further expansion by creating MSPs from rumen origin represented distinct spectral profiles presented in the dendrogram in Figure 10.

The strain preparation methods (section 2.3) used in other studies for clinical anaerobic strains produce sufficient microbial biomass as a single colony or with a pool of two to three isolated colonies for protein extraction, which was suitable for generating high-quality spectra (Asare et al., 2023). Since I was testing the application of the MALDI Biotyper® on fastidious microbes, challenges arose when using strictly anaerobic microbes of rumen origin. Selected rumen candidate strains grew poorly on Brain Heart Infusion (BHI Agar) recommended by Asare et al. (2023) and Colombia sheep blood agar recommended by Bakhtiary et al. (2018) for anaerobes from clinical studies (Asare et al., 2023; Bakhtiary et al., 2018). However, the rumen microbe-specific medium, M2GCS agar (Miyazaki et al., 1997), supported the growth of most candidate bacterial strains for generating protein spectral profiles using the MALDI Biotyper®. For some strains that formed small colonies, it was necessary to pool multiple colonies (more than four) to obtain sufficient biomass for protein extraction to generate spectra. However, spectra generated from pooled colonies failed quality control due to peak variability that exceeded the acceptable mass tolerance range. Quality control results for MSPs included in the expanded database are presented in Appendix 1. To obtain 20 valid spectra that met quality control criteria, at least 36 spectra

had to be captured from pooled (two to four) colonies that showed minimal variation, allowing 20 valid spectra for MSP creation.

When new isolates were identified using MALDI Biotyper®, most target spots failed to generate a spectrum from colonies grown on minimal medium modified RM02 agar with an added TMA-precursor. Isolated microbes produced tiny colonies with insufficient microbial biomass to generate a spectrum. However, of the targets for which successful spectra were generated, most remained unidentified, which may be for various reasons. One possibility is impure colonies, because isolated colonies have not been put through the purification step of a single colony streak. Secondly, this highlighted the absence of these strains in the targeted rumen reference database and overall under-representation of MSPs from rumen strains in Bruker's database.

Identified isolates using the MALDI Biotyper® confirmed matches to *Prevotella ruminicola* (from choline medium) and *Bacteroidales* bacterium (from betaine) using the targeted reference database. The MALDI Biotyper® identification was verified by 16S rRNA gene sequencing, which confirmed that new isolates matched *Prevotella ruminicola* using both identification methods. However, I could not verify the MALDI identification matching *Bacteroidales bacterium*, because of accidental loss of the colony. But the validation by rapid identification of a species-level match to *Prevotella ruminicola* highlighted the importance of expanding the database with spectral libraries from rumen strains. Other isolates were identified as *Streptococcus infantarius* and *Streptococcus lutetiensis* using the Bruker database and these were also verified using 16S rRNA sequencing, again validating the reliability and reproducibility of MALDI Biotyper® for microbial identification.

To further improve the identification of isolates in real time, spectral grouping based on the similarity scores proved highly effective. These unidentified spectra were grouped based on their protein profile similarities (Table 6). A representative strain from each cluster was identified using 16S rRNA gene sequencing (Table 7), which confirmed the utility of MALDI Biotyper® in reducing redundancy and prioritising phylogenetic diversity. This semi-targeted

approach improved throughput and ensured the inclusion of a broader range of taxa. Through this process, previously unidentified isolates were matched to *Butyrivibrio proteoclasticus, Butyrivibrio hungatei, Pseudobutyrivibrio ruminis,* and *Slackia heliotrinireducens* by 16S rRNA gene sequencing.

5.2 Microbial growth responses to substrate specificity

The TMA precursor, choline, is abundant in the rumen (Kelly et al., 2019). While betaine and carnitine are also TMA precursors and have been included in my study, there is limited research that has examined their contribution to TMA formation in the rumen (Servillo et al., 2018). Betaine has, however, been extensively studied for its osmo-protective roles in the rumen (Mahmood et al., 2020), and carnitine for its role in lipid metabolism during lactation (Meyer et al., 2020). Growth of microbiota from the calf rumen and faecal samples showed that they were enriched on choline, betaine, and carnitine supplemented minimal medium, and presumably selected for choline, betaine, or carnitine-utilising microbes.

The sample source had a significant influence on growth in the selective enrichments, as illustrated in Figure 12. The rumen sample exhibited an active microbial community that is adept at metabolising choline substrates compared with betaine and carnitine. This could be due to rumen microbiota exposure to an abundance of choline (105 ± 59 μM in the rumen fluid) (Foroutan et al., 2020; O'Callaghan et al., 2018) from plant phosphatidylcholine (Dawson & Hemington, 1974), as compared to the availability of betaine (15 ± 40 μM in rumen fluid) (Foroutan et al., 2020; O'Callaghan et al., 2018) and carnitine (3 ± 1 μM in rumen fluid) (Foroutan et al., 2020). Conversely, the lower growth in faecal enrichments suggests limited microbial enrichment occurred for these selective substrates. Several factors could explain this observation. Firstly, the bioavailability of the choline, betaine, and carnitine substrates could be limited in the lower intestinal tract because of susceptibility to rapid ruminal degradation (Davidson et al., 2008; Pirestani & Aghakhani, 2018). Secondly, the modified RM02 medium, containing 5% clarified rumen fluid, is specifically designed for cultivating rumen microbes, making it potentially less effective for faecal microbes.

Pure culture isolations were restricted to stocks of only the lowest dilution enrichments because of time limitations. The choline-enriched, identifiable pure isolates identified two *Prevotella ruminicola,* three *Streptococcus lutetiensis,* and a *Streptococcus infantarius*. Additionally, isolates originating from the betaine enrichment were identified as *Bacteroidales* sp. and *Streptococcus infantarius,* indicating a diversity of microbes obtained from both substrates. Interestingly, although the choline-enriched cultures showed higher absorbance values for OD_{600} compared to the betaine-enriched cultures, the betaine-enriched cultures produced a greater colony-forming unit (CFU) count than the choline-enriched ones. A possible explanation for the higher CFU in the betaine-enriched cultures could be the osmolyte properties of betaine, which help microbes cope with stress (Mahmood et al., 2020).

5.3 Source of TMA in the rumen

The production of TMA, ethanol, and acetate suggests TMA-lyase activity occurred in the choline enrichment cultures derived from rumen samples. The breakdown of 1 mol of choline produces 1 mol of trimethylamine, along with 0.5 mol each of acetic acid and ethanol (Müller et al., 1981). TMA quantification in samples from the choline enrichments confirmed full conversion of choline into TMA. However, TMA was low in a few of the betaine samples and no ethanol was detected in the betaine and carnitine enrichments, but acetate was detected in all sets of substrates. The highest levels of acetate were detected in the choline enrichment cultures, followed by the betaine and carnitine enrichment cultures. The production of acetate without TMA and ethanol in the betaine and carnitine enrichments does not involve TMA-lyase activity. Hence, betaine and carnitine are likely not the main sources of TMA production in the rumen ecosystem, while choline can be confirmed as a source of TMA.

5.4 Microbial cross-feeding may result in isolation of non-TMA-producing microbes from enrichment cultures

The pure culture isolates obtained in this study did not produce TMA. Some isolates were identified as *Bacteroides* and *Prevotella* species, which play an indirect yet ecologically significant role in the metabolism of both betaine and choline within the gut. Although they are not known to produce TMA from choline and betaine, nor are they known to possess TMA-lyase genes (e.g., *cutC* for choline and *grdH* for betaine), they may utilise choline, betaine and carnitine using other metabolic pathways (Creighbaum et al., 2019). Betaine utilising microbes may convert it into intermediates like dimethylglycine or use it for osmoprotection, aided by transporters and enzymes such as betaine aldehyde dehydrogenase (Munoz-Clares et al., 2010). Choline-utilising microbes may contribute indirectly by breaking down phosphatidylcholine, releasing free choline that TMA-producing bacteria can then utilise, or choline can be converted irreversibly to betaine via the enzyme choline dehydrogenase (Swartz et al., 2022). In both cases, the metabolic activity of choline and betaine-utilising microbes in the rumen enrichment tubes shaped substrate availability and microbial cross-feeding interactions, which favoured the growth of non-selective microbes during cultivations.

Streptococcus infantarius and *Streptococcus lutetiensis* were also identified as matches to isolates from the choline and betaine enrichments. However, these species are not known to metabolise choline or betaine directly, but this could be growing because of cross-feeding interactions. For instance, studies with *Lactococcus lactis* have shown that in co-culture with *Bacteroides* species, lactate produced by *L. lactis* is consumed by *Bacteroides*, which convert it into short-chain fatty acids like propionate and acetate, demonstrating an example of metabolic cross-feeding (Schultz & Breznak, 1979). Similarly, *Streptococcus* spp. likely benefits from these interactions by using fermentation end-products released by primary degraders of choline and betaine, such as acetogens or methanogens. This cross-feeding supports their growth and contributes to the stability and functionality of the microbial

ecosystem in anaerobic environments by linking different metabolic niches and facilitating efficient substrate utilisation within the microbial community.

5.5 Limitations and challenges of the study

There are limitations to this study that should be considered when interpreting the findings of the targeted identification of TMA-metabolising microorganisms. The candidate strains were not experimentally confirmed to produce TMA, and the presence of the *cutC/D* gene cluster is widely distributed but discontinuously present across bacterial phyla, likely due to horizontal gene transfer (Craciun & Balskus, 2012). Hence, the confirmation of the presence of genes alone is a poor indicator to use for strain selection for this study.

One limitation of the study is that the purifications were conducted after 24 hours of incubation of enrichment cultures using the lowest dilution of the rumen sample. This approach did not yield pure isolates that produced TMA. However, if sequential subculturing had been performed over a week, with subcultures every 48 hours before isolation, it would have facilitated the selection of microbes of interest on specific substrates.

5.6 Conclusion

In conclusion, the rumen microbiota adept at utilising choline, betaine, and carnitine was successfully enriched and produced TMA only by utilising choline. However, none of the obtained pure isolates produced TMA, concluding that alternative targeted cultivations are required. Rapid identification was facilitated using a MALDI Biotyper® with successful expansion of the database, combined with add-on spectra. Integrating the spectral profiles from the entire Hungate 1000 culture collection into the MALDI-Biotyper® reference databases will enhance throughput. Given the ecological and agricultural importance of rumen microbiota, there is a critical need to expand mass spectral profiling efforts to build robust databases for rumen-derived anaerobic strains. This would enable accurate and rapid identification of rumen bacteria.

5.7 Future studies

Findings in this study could be applied in future research to selectively grow specific microbial populations from rumen or faeces microbiomes. The insights from data showed favourable microbial growth conditions that can help further refinement in targeted enrichment protocols for cultivating microbes of interest. Since the isolation efforts were only performed on the lowest diluted enrichments from rumen samples, future research could also focus on isolating from the remaining dilutions of enrichments produced in this study.

This pilot study has important implications for the development of microbiota-targeted therapies in humans, especially in the context of faecal microbiota transplantation, gastrointestinal health, and mitigation of cardiometabolic risks associated with TMAO accumulation. Future efforts should focus on novel TMA-producing anaerobes from the rumen to identify new enzymes and pathways involved in metabolism.

Chapter 6. References

Abhijith, A., Dunshea, F. R., Chauhan, S. S., Sejian, V., & DiGiacomo, K. (2024). A Meta-Analysis of the Effects of Dietary Betaine on Milk Production, Growth Performance, and Carcass Traits of Ruminants. *Animals*, *14*(12), 1756.

Alcala, L., Marin, M., Ruiz, A., Quiroga, L., Zamora-Cintas, M., Fernandez-Chico, M. A., Munoz, P., & Rodriguez-Sanchez, B. (2021). Identifying Anaerobic Bacteria Using MALDI-TOF Mass Spectrometry: A Four-Year Experience. *Front Cell Infect Microbiol*, *11*, 521014. https://doi.org/10.3389/fcimb.2021.521014

Alexandrescu, L., Suceveanu, A. P., Stanigut, A. M., Tofolean, D. E., Axelerad, A. D., Iordache, I. E., Herlo, A., Nelson Twakor, A., Nicoara, A. D., & Tocia, C. (2024). Intestinal Insights: The Gut Microbiome's Role in Atherosclerotic Disease: A Narrative Review. *Microorganisms*, *12*(11), 2341.

Arias, N., Arboleya, S., Allison, J., Kaliszewska, A., Higarza, S. G., Gueimonde, M., & Arias, J. L. (2020). The Relationship between Choline Bioavailability from Diet, Intestinal Microbiota Composition, and Its Modulation of Human Diseases. *Nutrients*, *12*(8), 2340.

Asare, P. T., Lee, C. H., Hürlimann, V., Teo, Y., Cuénod, A., Akduman, N., Gekeler, C., Afrizal, A., Corthesy, M., Kohout, C., Thomas, V., de Wouters, T., Greub, G., Clavel, T., Pamer, E. G., Egli, A., Maier, L., & Vonaesch, P. (2023). A MALDI-TOF MS library for rapid identification of human commensal gut bacteria from the class Clostridia. *Front Microbiol*, *14*, 1104707. https://doi.org/10.3389/fmicb.2023.1104707

Attwood, G. T., Wakelin, S. A., Leahy, S. C., Rowe, S., Clarke, S., Chapman, D. F., Muirhead, R., & Jacobs, J. M. (2019). Applications of the soil, plant and rumen microbiomes in pastoral agriculture. *Frontiers in nutrition*, *6*, 107.

Bächli, P., Baars, S., Simmler, A., Zbinden, R., & Schulthess, B. (2022). Impact of MALDI-TOF MS identification on anaerobic species and genus diversity in routine diagnostics. *Anaerobe*, *75*, 102554.

Bakhtiary, F., Sayevand, H. R., Remely, M., Hippe, B., Indra, A., Hosseini, H., & Haslberger, A. G. (2018). Identification of Clostridium spp. derived from a sheep and cattle slaughterhouse by matrix-assisted laser desorption and ionization-time of flight mass spectrometry (MALDI-TOF MS) and 16S rDNA sequencing. *J Food Sci Technol*, *55*(8), 3232–3240. https://doi.org/10.1007/s13197-018-3255-2

Balch, C. (1959). Structure of the ruminant stomach and the movement of its contents. *Proceedings of the Nutrition Society*, *18*(2), 97–102.

Banerjee, S., & van der Heijden, M. G. (2022). Soil microbiomes and one health. *Nature Reviews Microbiology*, 1–15.

Bauchop, T. (1979). Rumen anaerobic fungi of cattle and sheep. *Applied and Environmental Microbiology*, *38*(1), 148–158.

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C. C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G. H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J. A., Maguin, E., Mauchline, T., McClure, R.,...Schloter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, *8*(1), 103. https://doi.org/10.1186/s40168-020-00875-0

Boysen, A. K., Durham, B. P., Kumler, W., Key, R. S., Heal, K. R., Carlson, L. T., Groussman, R. D., Armbrust, E. V., & Ingalls, A. E. (2022). Glycine betaine uptake and metabolism in marine microbial communities. *Environ Microbiol*, *24*(5), 2380–2403. https://doi.org/10.1111/1462-2920.16020

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC bioinformatics*, *10*, 1–9.

Craciun, S., & Balskus, E. P. (2012). Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proceedings of the National Academy of Sciences*, *109*(52), 21307–21312.

Creevey, C. J., Kelly, W. J., Henderson, G., & Leahy, S. C. (2014). Determining the culturability of the rumen bacterial microbiome. *Microbial biotechnology*, *7*(5), 467–479.

Creighbaum, A. J., Ticak, T., Shinde, S., Wang, X., & Ferguson, D. J. (2019). Examination of the Glycine Betaine-Dependent Methylotrophic Methanogenesis Pathway: Insights Into Anaerobic Quaternary Amine Methylotrophy [Original Research]. *Frontiers in Microbiology*, *Volume 10 - 2019*. https://doi.org/10.3389/fmicb.2019.02572

Davidson, S., Hopkins, B. A., Odle, J., Brownie, C., Fellner, V., & Whitlow, L. W. (2008). Supplementing Limited Methionine Diets with Rumen-Protected Methionine, Betaine, and Choline in Early Lactation Holstein Cows. *Journal of dairy science*, *91*(4), 1552–1559. https://doi.org/https://doi.org/10.3168/jds.2007-0721

Dawson, R. M., & Hemington, N. (1974). Digestion of grass lipids and pigments in the sheep rumen. *Br J Nutr*, *32*(2), 327–340. https://doi.org/10.1079/bjn19740086

De Anda, V., Chen, L., Dombrowski, N., Hua, Z., Jiang, H., Banfield, J., Li, W., & Baker, B. (2021). Brockarchaeota, a novel archaeal phylum with unique and versatile carbon cycling pathways. Nat Commun 12: 2404. In.

DeGruttola, A. K., Low, D., Mizoguchi, A., & Mizoguchi, E. (2016). Current Understanding of Dysbiosis in Disease in Human and Animal Models. *Inflamm Bowel Dis*, *22*(5), 1137–1150. https://doi.org/10.1097/mib.0000000000000750

Diakite, A., Dubourg, G., Dione, N., Afouda, P., Bellali, S., Ngom, I. I., Valles, C., Lagier, J.-C., & Raoult, D. (2020). Optimization and standardization of the culturomics technique for human microbiome exploration. *Scientific Reports*, *10*(1), 1–7.

Falony, G., Vieira-Silva, S., & Raes, J. (2015). Microbiology meets big data: the case of gut microbiota–derived trimethylamine. *Annual review of microbiology*, *69*, 305–321.

Foroutan, A., Fitzsimmons, C., Mandal, R., Piri-Moghadam, H., Zheng, J., Guo, A., Li, C., Guan, L. L., & Wishart, D. S. (2020). The bovine metabolome. *Metabolites*, *10*(6), 233.

Gäbel, G., & Sehested, J. (1997). SCFA transport in the forestomach of ruminants. *Comparative Biochemistry and Physiology Part A: Physiology*, *118*(2), 367–374.

Ghoneem, W. M. A., & El-Tanany, R. R. A. (2023). Impact of natural betaine supplementation on rumen fermentation and productive performance of lactating Damascus goats. *Tropical Animal Health and Production*, *55*(2), 123. https://doi.org/10.1007/s11250-023-03524-4

Greub, G. (2012). Culturomics: a new approach to study the human microbiome. *Clin Microbiol Infect*, *18*(12), 1157–1159. https://doi.org/10.1111/1469-0691.12032

Harmeyer, J. (2003). Use of L-carnitine additions in domestic animal feeds. *Lohmann information*, *28*, 1–9.

Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., Abecia, L., Angarita, E., Aravena, P., Arenas, G. N., & Ariza, C. (2015). Rumen microbial community composition varies with diet and host,

but a core microbiome is found across a wide geographical range. *Scientific Reports*, *5*, 14567.

Hoyles, L., Jiménez-Pranteda, M. L., Chilloux, J., Brial, F., Myridakis, A., Aranias, T., Magnan, C., Gibson, G. R., Sanderson, J. D., & Nicholson, J. K. (2018). Metabolic retroconversion of trimethylamine N-oxide and the gut microbiota. *Microbiome*, *6*, 1–14.

Hungate, R. (1969). Chapter IV A roll tube method for cultivation of strict anaerobes. In *Methods in microbiology* (Vol. 3, pp. 117–132). Elsevier.

Hungate, R. E. (2013). *The rumen and its microbes*. Elsevier.

Huws, S. A., Creevey, C. J., Oyama, L. B., Mizrahi, I., Denman, S. E., Popova, M., Muñoz-Tamayo, R., Forano, E., Waters, S. M., & Hess, M. (2018). Addressing global ruminant agricultural challenges through understanding the rumen microbiome: past, present, and future. *Frontiers in Microbiology*, *9*, 2161.

Jameson, E., Quareshy, M., & Chen, Y. (2018). Methodological considerations for the identification of choline and carnitine-degrading bacteria in the gut. *Methods*, *149*, 42–48. https://doi.org/10.1016/j.ymeth.2018.03.012

Jandhyala, S. M., Talukdar, R., Subramanyam, C., Vuyyuru, H., Sasikala, M., & Nageshwar Reddy, D. (2015). Role of the normal gut microbiota. *World J Gastroenterol*, *21*(29), 8787–8803. https://doi.org/10.3748/wjg.v21.i29.8787

Janeiro, M. H., Ramírez, M. J., Milagro, F. I., Martínez, J. A., & Solas, M. (2018). Implication of Trimethylamine N-Oxide (TMAO) in Disease: Potential Biomarker or New Therapeutic Target. *Nutrients*, *10*(10). https://doi.org/10.3390/nu10101398

Jang, K.-S., & Kim, Y. (2018). Rapid and robust MALDI-TOF MS techniques for microbial identification: a brief overview of their diverse applications. *Journal of Microbiology*, *56*. https://doi.org/10.1007/s12275-018-7457-0

Jie, Z., Xia, H., Zhong, S. L., Feng, Q., Li, S., Liang, S., Zhong, H., Liu, Z., Gao, Y., Zhao, H., Zhang, D., Su, Z., Fang, Z., Lan, Z., Li, J., Xiao, L., Li, J., Li, R., Li, X.,...Kristiansen, K. (2017). The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun*, *8*(1), 845. https://doi.org/10.1038/s41467-017-00900-1

Kelly, W. J., Leahy, S. C., Kamke, J., Soni, P., Koike, S., Mackie, R., Seshadri, R., Cook, G. M., Morales, S. E., & Greening, C. (2019). Occurrence and expression of genes encoding methyl-compound production in rumen bacteria. *Animal Microbiome*, *1*(1), 1–13.

Kenters, N., Henderson, G., Jeyanathan, J., Kittelmann, S., & Janssen, P. H. (2011). Isolation of previously uncultured rumen bacteria by dilution to extinction using a new liquid culture medium. *J Microbiol Methods*, *84*(1), 52–60. https://doi.org/10.1016/j.mimet.2010.10.011

Kerfeld, C. A., Aussignargues, C., Zarzycki, J., Cai, F., & Sutter, M. (2018). Bacterial microcompartments. *Nature Reviews Microbiology*, *16*(5), 277–290. https://doi.org/10.1038/nrmicro.2018.10

King, G. M. (1984). Metabolism of trimethylamine, choline, and glycine betaine by sulfate-reducing and methanogenic bacteria in marine sediments. *Appl Environ Microbiol*, *48*(4), 719–725. https://doi.org/10.1128/aem.48.4.719-725.1984

Lagier, J. C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., Bittar, F., Fournous, G., Gimenez, G., Maraninchi, M., Trape, J. F., Koonin, E. V., La Scola, B., & Raoult, D. (2012). Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect*, *18*(12), 1185–1193. https://doi.org/10.1111/1469-0691.12023

Lagier, J. C., Hugon, P., Khelaifia, S., Fournier, P. E., La Scola, B., & Raoult, D. (2015). The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev*, *28*(1), 237–264. https://doi.org/10.1128/cmr.00014-14

Lagier, J. C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., Caputo, A., Cadoret, F., Traore, S. I., Seck, E. H., Dubourg, G., Durand, G., Mourembou, G., Guilhot, E., Togo, A., Bellali, S., Bachar, D., Cassir, N., Bittar, F.,...Raoult, D. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol*, *1*, 16203. https://doi.org/10.1038/nmicrobiol.2016.203

Lee, H. H., Kim, D. J., Ahn, H. J., Ha, J. Y., & Suh, S. W. (2004). Crystal structure of T-protein of the glycine cleavage system. Cofactor binding, insights into H-protein recognition, and molecular basis for understanding nonketotic hyperglycinemia. *J Biol Chem*, *279*(48), 50514–50523. https://doi.org/10.1074/jbc.M409672200

Li, C., Luan, Z., Zhao, Y., Chen, J., Yang, Y., Wang, C., Jing, Y., Qi, S., Li, Z., & Guo, H. (2022). Deep insights into the gut microbial community of extreme longevity in south Chinese centenarians by ultra-deep metagenomics and large-scale culturomics. *NPJ biofilms and microbiomes*, *8*(1), 1–12.

Li, C., Wang, F., Mao, Y., Ma, Y., & Guo, Y. (2025). Multi-omics reveals the mechanism of Trimethylamine N-oxide derived from gut microbiota inducing liver fatty of dairy cows. *BMC Genomics*, *26*(1), 10. https://doi.org/10.1186/s12864-024-11067-7

Li, Y., Kreuzer, M., Clayssen, Q., Ebert, M.-O., Ruscheweyh, H.-J., Sunagawa, S., Kunz, C., Attwood, G., Amelchanka, S., & Terranova, M. (2021). The rumen microbiome inhibits methane formation through dietary choline supplementation. *Scientific Reports*, *11*(1), 1–15.

Lyu, Z., Shao, N., Akinyemi, T., & Whitman, W. B. (2018). Methanogenesis. *Current Biology*, *28*(13), R727–R732.

Mahmood, M., Petri, R. M., Gavrău, A., Zebeli, Q., & Khiaosa-Ard, R. (2020). Betaine addition as a potent ruminal fermentation modulator under hyperthermal and hyperosmotic conditions in vitro. *J Sci Food Agric*, *100*(5), 2261–2271. https://doi.org/10.1002/jsfa.10255

Mahoney-Kurpe, S. C., Palevich, N., Noel, S. J., Gagic, D., Biggs, P. J., Soni, P., Reid, P. M., Koike, S., Kobayashi, Y., & Janssen, P. H. (2023). Aristaeella hokkaidonensis gen. nov. sp. nov. and Aristaeella lactis sp. nov., two rumen bacterial species of a novel proposed family, Aristaeellaceae fam. nov. *International journal of systematic and evolutionary microbiology*, *73*(5), 005831.

Martínez-del Campo, A., Bodea, S., Hamer, H. A., Marks, J. A., Haiser, H. J., Turnbaugh, P. J., & Balskus, E. P. (2015). Characterization and detection of a widely distributed gene cluster that predicts anaerobic *MBio*, *6*(2), e00042–00015.

Massmig, M., Reijerse, E., Krausze, J., Laurich, C., Lubitz, W., Jahn, D., & Moser, J. (2020). Carnitine metabolism in the human gut: characterization of the two-component carnitine monooxygenase CntAB from Acinetobacter baumannii. *J Biol Chem*, *295*(37), 13065–13078. https://doi.org/10.1074/jbc.RA120.014266

Maurer, F. P., Christner, M., Hentschke, M., & Rohde, H. (2017). Advances in rapid identification and susceptibility testing of bacteria in the clinical microbiology laboratory: implications for patient care and antimicrobial stewardship programs. *Infectious disease reports*, *9*(1), 6839.

Meyer, J., Daniels, S. U., Grindler, S., Tröscher-Mußotter, J., Alaedin, M., Frahm, J., Hüther, L., Kluess, J., Kersten, S., & von Soosten, D. (2020). Effects of a dietary L-carnitine supplementation on performance, energy metabolism and recovery from calving in dairy cows. *Animals*, *10*(2), 342.

Miyazaki, K., Martin, J. C., Marinsek-Logar, R., & Flint, H. J. (1997). Degradation and Utilization of Xylans by the Rumen AnaerobePrevotella bryantii (formerlyP. ruminicolasubsp. brevis) B14. *Anaerobe*, *3*(6), 373–381.

Mizrahi, I., Wallace, R. J., & Moraïs, S. (2021). The rumen microbiome: balancing food security and environmental impacts. *Nature Reviews Microbiology*, *19*(9), 553–566. https://doi.org/10.1038/s41579-021-00543-6

Mouné, S., Manac'h, N., Hirschler, A., Caumette, P., Willison, J., & Matheron, R. (1999). NEW TAXA-Proteobacteria-Haloanaerobacter salinarius sp nov, a novel halophilic fermentative bacterium that reduces glycine-betaine to trimethylamine with hydrogen or serine as electron donors. *International Journal of Systematic Bacteriology*, *49*(1), 103–112.

Müller, E., Fahlbusch, K., Walther, R., & Gottschalk, G. (1981). Formation of N,N-Dimethylglycine, Acetic Acid, and Butyric Acid from Betaine by Eubacterium limosum. *Applied and Environmental Microbiology*, *42*(3), 439–445. https://doi.org/doi:10.1128/aem.42.3.439-445.1981

Munoz-Clares, R. A., Diaz-Sanchez, A. G., Gonzalez-Segura, L., & Montiel, C. (2010). Kinetic and structural features of betaine aldehyde dehydrogenases: mechanistic and regulatory implications. *Arch Biochem Biophys*, *493*(1), 71–81. https://doi.org/10.1016/j.abb.2009.09.006

Noel, S. (2013). *Cultivation and community composition analysis of plant-adherent rumen bacteria: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Microbiology and Genetics at Massey University, Palmerston North, New Zealand* Massey University].

O'Callaghan, T. F., Vázquez-Fresno, R., Serra-Cayuela, A., Dong, E., Mandal, R., Hennessy, D., McAuliffe, S., Dillon, P., Wishart, D. S., & Stanton, C. (2018). Pasture feeding changes the bovine rumen and milk metabolome. *Metabolites*, *8*(2), 27.

Panda, A., Kurapati, S., Samantaray, J. C., Srinivasan, A., & Khalil, S. (2014). MALDI-TOF mass spectrometry proteomic based identification of clinical bacterial isolates. *Indian J Med Res*, *140*(6), 770–777.

Panwar, A., Martins, B. M., Sommer, F., Schroda, M., Dobbek, H., Iobbi-Nivol, C., Jourlin-Castelli, C., & Leimkühler, S. (2024). Purification and Electron Transfer from Soluble c-Type Cytochrome TorC to TorA for Trimethylamine N-Oxide Reduction. *Int J Mol Sci*, *25*(24). https://doi.org/10.3390/ijms252413331

Pirestani, A., & Aghakhani, M. (2018). The effects of rumen-protected choline and l-carnitine supplementation in the transition period on reproduction, production, and some metabolic diseases of dairy cattle. *Journal of Applied Animal Research*, *46*(1), 435–440.

Quareshy, M., Shanmugam, M., Townsend, E., Jameson, E., Bugg, T. D. H., Cameron, A. D., & Chen, Y. (2021). Structural basis of carnitine monooxygenase CntA substrate specificity, inhibition, and intersubunit electron transfer. *J Biol Chem*, *296*, 100038. https://doi.org/10.1074/jbc.RA120.016019

Rajasekaran, K. (2023). *Phenotypic characterisation of members of the Lachnospiraceae family isolated from ruminants: a thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Microbiology at Massey University, Manawatū, New Zealand* Massey University].

Rath, S., Heidrich, B., Pieper, D. H., & Vital, M. (2017). Uncovering the trimethylamine-producing bacteria of the human gut microbiota. *Microbiome*, *5*(1), 54. https://doi.org/10.1186/s40168-017-0271-9

Rath, S., Rud, T., Pieper, D. H., & Vital, M. (2020). Potential TMA-producing bacteria are ubiquitously found in mammalia. *Frontiers in Microbiology*, *10*, 2966.

Ringseis, R., Keller, J., & Eder, K. (2018). Regulation of carnitine status in ruminants and efficacy of carnitine supplementation on performance and health aspects of ruminant livestock: a review. *Archives of animal nutrition*, *72*(1), 1–30.

Romano, K. A., Vivas, E. I., Amador-Noguez, D., & Rey, F. E. (2015). Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. *MBio*, *6*(2), e02481. https://doi.org/10.1128/mBio.02481-14

Schultz, J., & Breznak, J. A. (1979). Cross-feeding of lactate between Streptococcus lactis and Bacteroides sp. isolated from termite hindguts. *Applied and Environmental Microbiology*, *37*(6), 1206–1210.

Servillo, L., D'Onofrio, N., Giovane, A., Casale, R., Cautela, D., Castaldo, D., Iannaccone, F., Neglia, G., Campanile, G., & Balestrieri, M. L. (2018). Ruminant meat and milk contain δ-valerobetaine, another precursor of trimethylamine N-oxide (TMAO) like γ-butyrobetaine. *Food Chem*, *260*, 193–199. https://doi.org/10.1016/j.foodchem.2018.03.114

Seshadri, R., Leahy, S. C., Attwood, G. T., Teh, K. H., Lambie, S. C., Cookson, A. L., Eloe-Fadrosh, E. A., Pavlopoulos, G. A., Hadjithomas, M., Varghese, N. J., Paez-Espino, D., Palevich, N., Janssen, P. H., Ronimus, R. S., Noel, S., Soni, P., Reilly, K., Atherly, T., Ziemer, C.,...Hungate project, c. (2018). Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nature biotechnology*, *36*(4), 359–367. https://doi.org/10.1038/nbt.4110

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.

Sleator, R. D., Shortall, C., & Hill, C. (2008). Metagenomics. *Letters in applied microbiology*, *47*(5), 361–366.

Swartz, T., Bradford, B., Malysheva, O., Caudill, M., Mamedova, L., & Estes, K. (2022). Effects of dietary rumen-protected choline supplementation on colostrum yields, quality, and choline metabolites from dairy cattle. *JDS communications*, *3*(4), 296–300.

Ticak, T., Kountz, D. J., Girosky, K. E., Krzycki, J. A., & Ferguson, D. J. (2014). A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. *Proceedings of the National Academy of Sciences*, *111*(43), E4668–E4676. https://doi.org/doi:10.1073/pnas.1409642111

Tidjani Alou, M., Million, M., Traore, S. I., Mouelhi, D., Khelaifia, S., Bachar, D., Caputo, A., Delerce, J., Brah, S., Alhousseini, D., Sokhna, C., Robert, C., Diallo, B. A., Diallo, A., Parola, P., Golden, M., Lagier, J. C., & Raoult, D. (2017). Gut Bacteria Missing in Severe Acute Malnutrition, Can We Identify Potential Probiotics by Culturomics? *Front Microbiol*, *8*, 899. https://doi.org/10.3389/fmicb.2017.00899

Tschech, A., & Pfennig, N. (1984). Growth yield increase linked to caffeate reduction in Acetobacterium woodii. *Archives of microbiology*, *137*(2), 163–167.

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. *Clin Microbiol Infect*, *20*(12), O1091–1097. https://doi.org/10.1111/1469-0691.12644

Veloo, A. C. M., Jean-Pierre, H., Justesen, U. S., Morris, T., Urban, E., Wybo, I., Kostrzewa, M., Friedrich, A. W., & workgroup, E. (2018). Validation of MALDI-TOF MS Biotyper database optimized for anaerobic bacteria: The ENRIA project. *Anaerobe*, *54*, 224–230. https://doi.org/10.1016/j.anaerobe.2018.03.007

Widdel, F., Kohring, G.-W., & Mayer, F. (1983). Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids: III. Characterization of the filamentous gliding Desulfonema limicola gen. nov. sp. nov., and Desulfonema magnum sp. nov. *Archives of microbiology*, *134*(4), 286–294.

Zhou, Y., Jin, W., Xie, F., Mao, S., Cheng, Y., & Zhu, W. (2021). The role of Methanomassiliicoccales in trimethylamine metabolism in the rumen of dairy cows. *Animal*, *15*(7), 100259.

Zhu, Y., Jameson, E., Crosatti, M., Schäfer, H., Rajakumar, K., Bugg, T. D. H., & Chen, Y. (2014). Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota. *Proceedings of the National Academy of Sciences*, *111*(11), 4268–4273. https://doi.org/doi:10.1073/pnas.1316569111

Appendix 1. Quality control for created Libraries

Quality control is a crucial step in ensuring the reproducibility of MSP creation in MALDI-TOF MS-based microbial identification. It is necessary to measure relative mass deviation (RMD) across the spectra before creating an MSP from 20 replicate spectra. Quality control parameters recorded in Appendix 1 for each created MSP step ensure mass accuracy and alignment across replicates, confirm consistent performance across the full m/z range, and guarantee high-quality, reproducible MSP generation.

To calculate relative mass deviation for each peak:

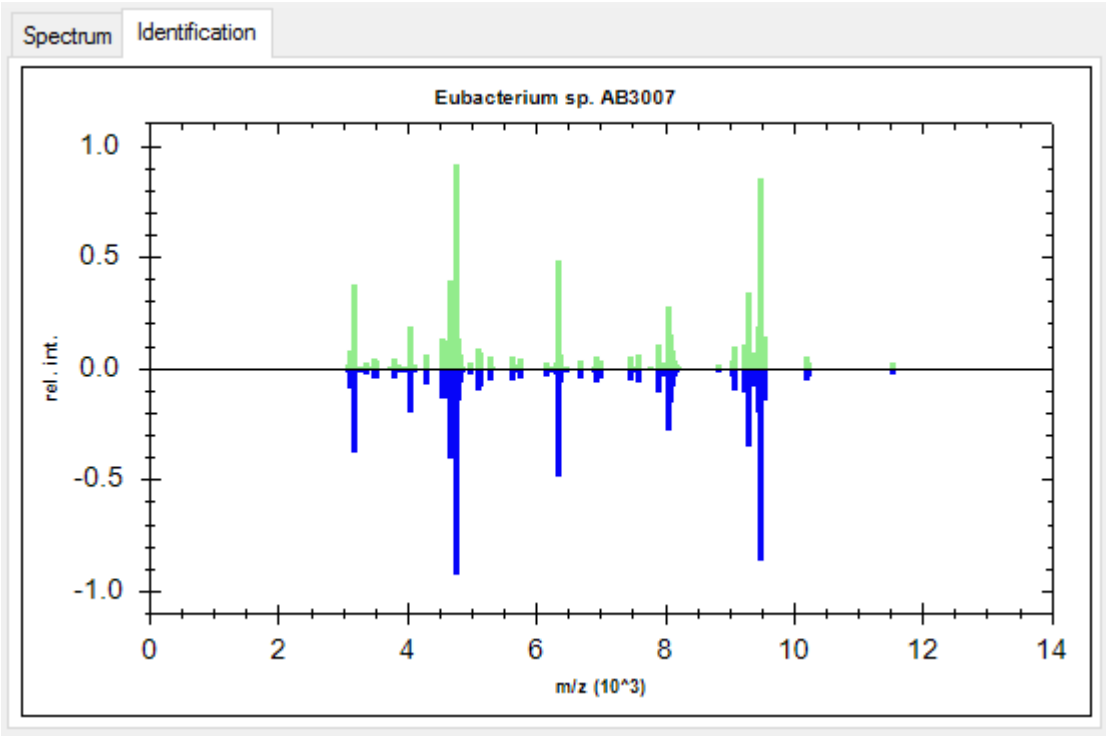
RMD (ppm) = ((upper mass – lower mass)/ upper mass) *1,000,000

The recommended peak shift between the spectra with the lower (smallest) and the upper (largest) mass is 500 ppm.

Lower mass = peak mass with the smallest m/z (most left on the electrograph)

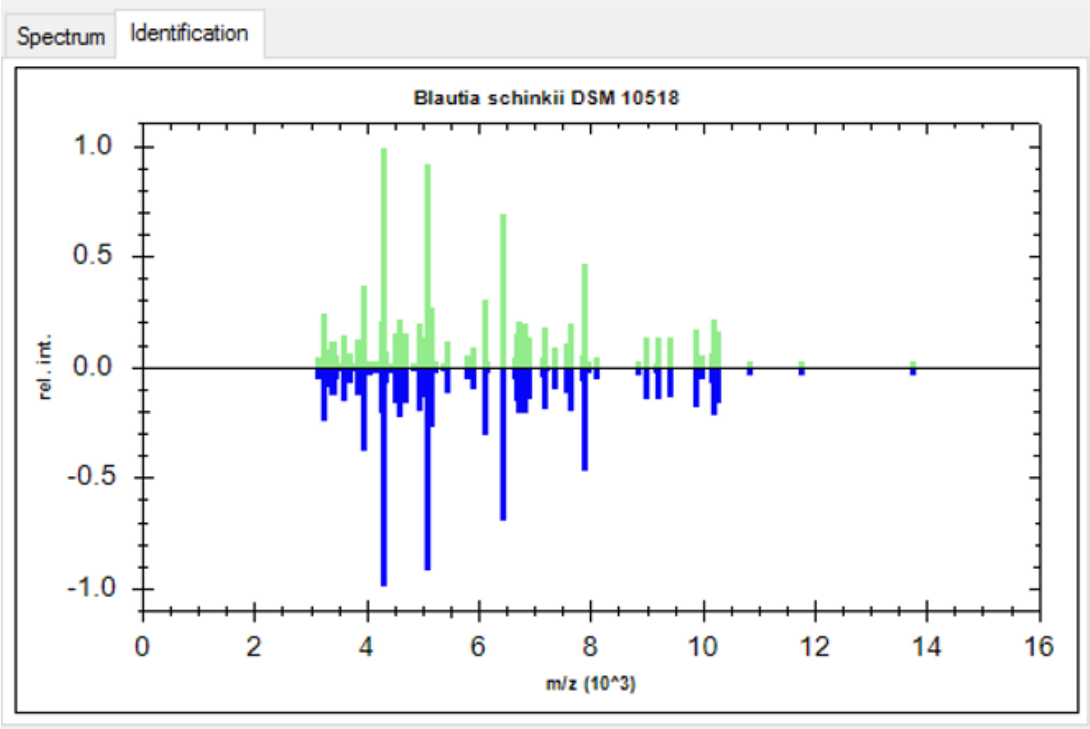
Upper mass = the comparative peak mass with the highest m/z value (most right on the electrograph)

A.1. *Eubacterium* sp. AB3007 (Hun105)



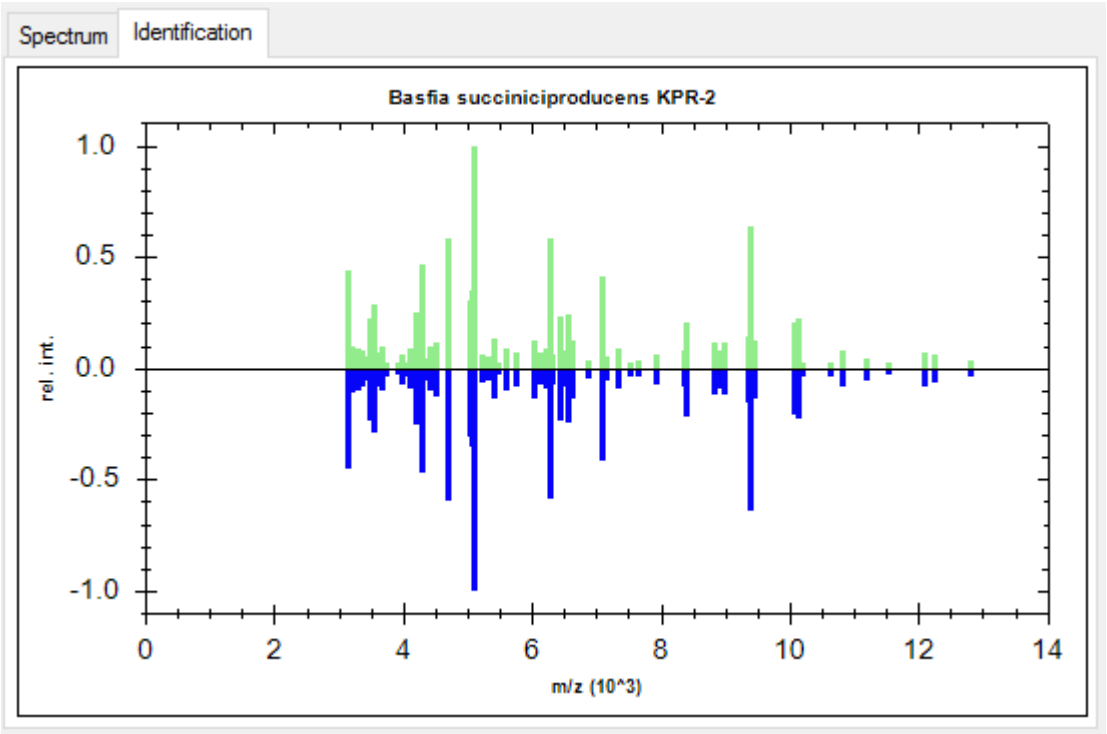
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3097 Da	3098 Da	1.1 Da	339 ppm	PASS
4000 Da to 5000 Da	4648 Da	4650 Da	1.3 Da	275 ppm	PASS
5000 Da to 6000 Da	5091 Da	5093 Da	2.0 Da	393 ppm	PASS
6000 Da to 7000 Da	6993 Da	6997 Da	3.1 Da	447 ppm	PASS
7000 Da to 8000 Da	7591 Da	7592 Da	0.8 Da	107 ppm	PASS
8000 Da to 9000 Da	8043 Da	8046 Da	3.4 Da	418 ppm	PASS
9000 Da to 10000 Da	9293 Da	9296 Da	2.7 Da	292 ppm	PASS

A.2. *Blautia schinkii* DSM 10518 (Hun196)



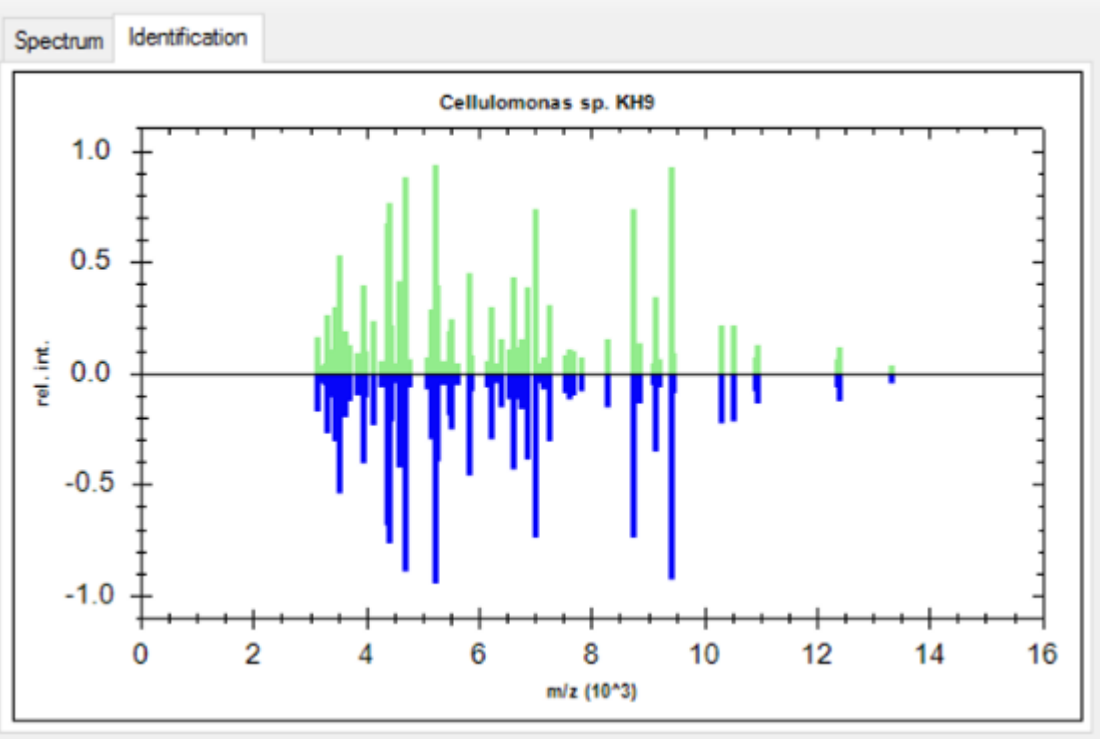
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3217 Da	3219 Da	1.6 Da	494 ppm	PASS
4000 Da to 5000 Da	4307 Da	4309 Da	1.9 Da	429 ppm	PASS
5000 Da to 6000 Da	5071 Da	5073 Da	2.0 Da	394 ppm	PASS
6000 Da to 7000 Da	6437 Da	6440 Da	3.0 Da	466 ppm	PASS
7000 Da to 8000 Da	7166 Da	7169 Da	3.2 Da	442 ppm	PASS
8000 Da to 9000 Da	8101 Da	8104 Da	3.4 Da	416 ppm	PASS
9000 Da to 10000 Da	9176 Da	9180 Da	4.5 Da	489 ppm	PASS

A.3. *Basfia succiniciproducens* KPR-2 (Hun298)



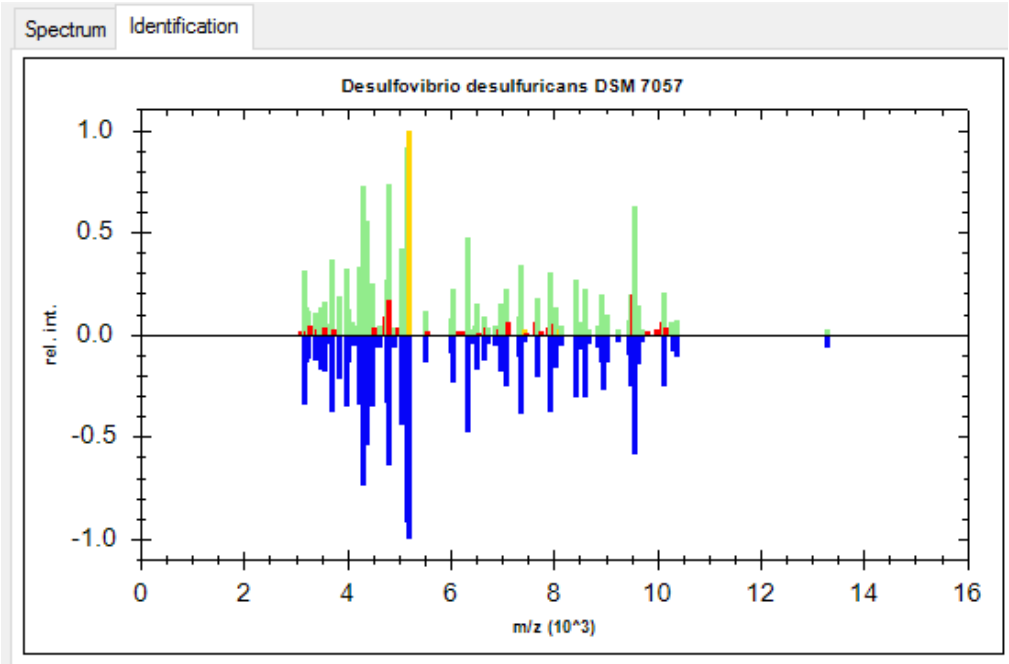
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3180 Da	3181 Da	1.1 Da	330 ppm	PASS
4000 Da to 5000 Da	4280 Da	4282 Da	1.8 Da	430 ppm	PASS
5000 Da to 6000 Da	5101 Da	5103 Da	2.0 Da	392 ppm	PASS
6000 Da to 7000 Da	6264 Da	6267 Da	3.0 Da	474 ppm	PASS
7000 Da to 8000 Da	7093 Da	7096 Da	3.2 Da	444 ppm	PASS
8000 Da to 9000 Da	8378 Da	8381 Da	3.4 Da	409 ppm	PASS
9000 Da to 10000 Da	9375 Da	9379 Da	4.5 Da	483 ppm	PASS

A.4. *Cellulomonas* sp. KH9 (Hun336)



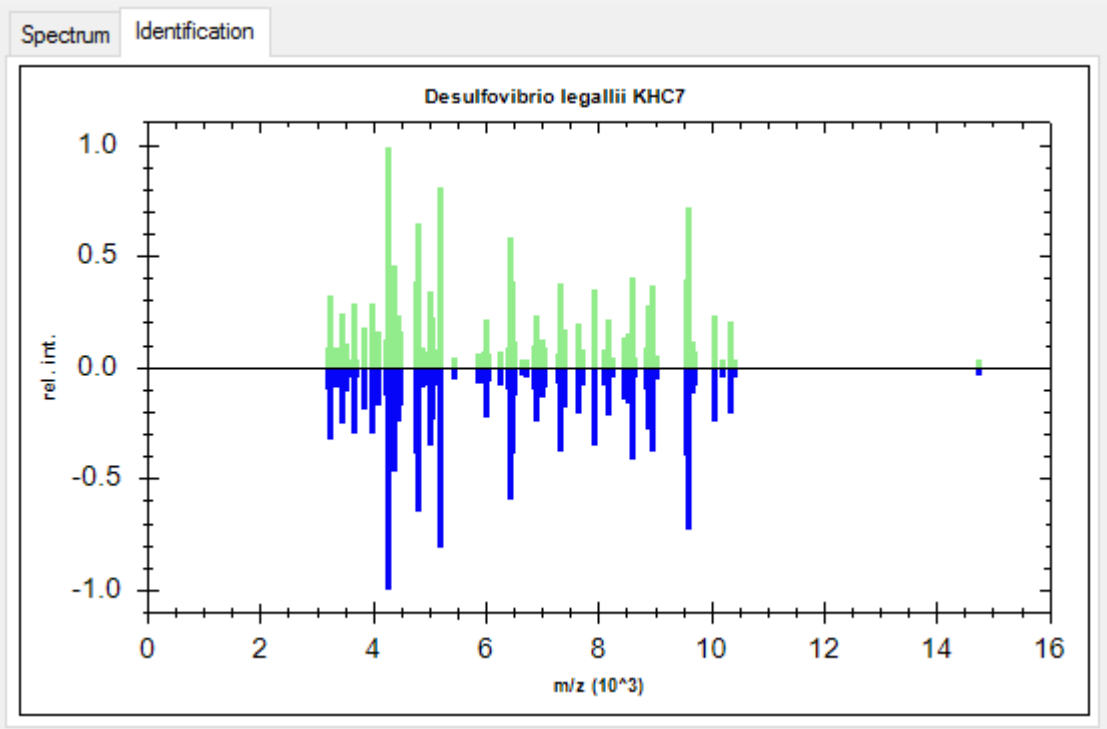
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3955 Da	3957 Da	1.8 Da	445 ppm	PASS
4000 Da to 5000 Da	4402 Da	4404 Da	1.9 Da	422 ppm	PASS
5000 Da to 6000 Da	5224 Da	5226 Da	2.0 Da	388 ppm	PASS
6000 Da to 7000 Da	6194 Da	6197 Da	2.9 Da	474 ppm	PASS
7000 Da to 8000 Da	7222 Da	7225 Da	3.2 Da	440 ppm	PASS
8000 Da to 9000 Da	8265 Da	8268 Da	3.4 Da	412 ppm	PASS
9000 Da to 10000 Da	9388 Da	9392 Da	4.5 Da	483 ppm	PASS

A.5. *Desulfovibrio desulfuricans* G11 (DSM 7057) (Hun310)



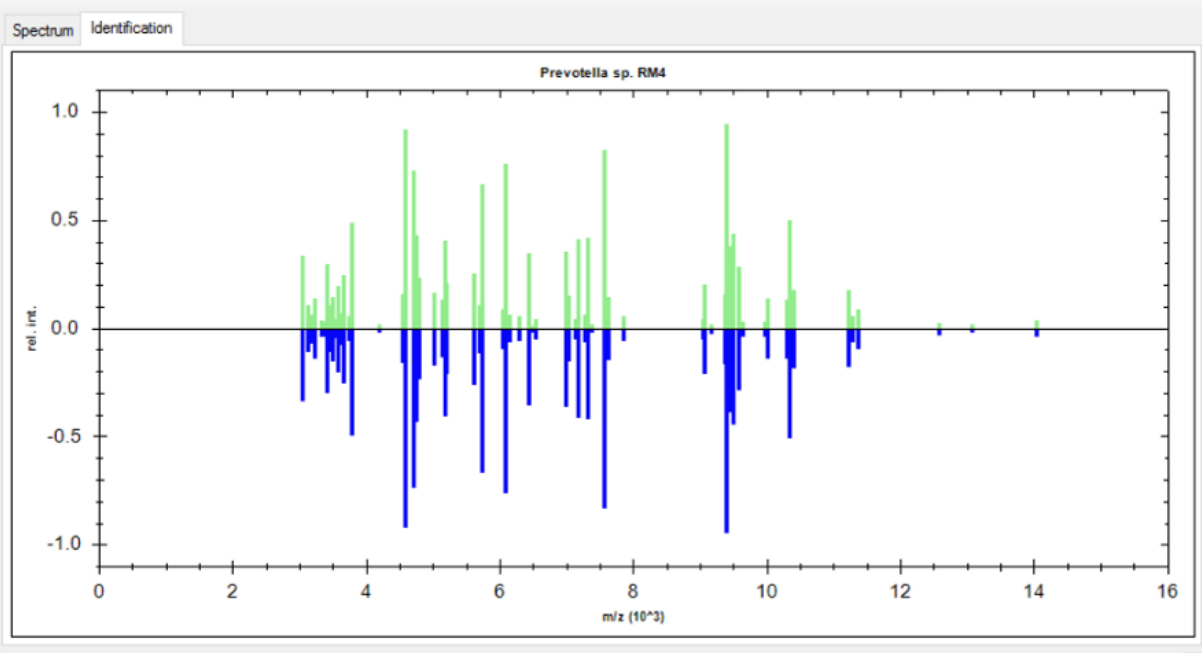
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3162 Da	3163 Da	1.6 Da	499 ppm	PASS
4000 Da to 5000 Da	4211 Da	4213 Da	1.8 Da	432 ppm	PASS
5000 Da to 6000 Da	5167 Da	5169 Da	2.0 Da	391 ppm	PASS
6000 Da to 7000 Da	6326 Da	6328 Da	2.2 Da	354 ppm	PASS
7000 Da to 8000 Da	7353 Da	7356 Da	3.2 Da	438 ppm	PASS
8000 Da to 9000 Da	8421 Da	8425 Da	3.4 Da	408 ppm	PASS
9000 Da to 10000 Da	9553 Da	9557 Da	4.6 Da	479 ppm	PASS

A.6. *Desulfovibrio* legallii KHC7 (Hun434)



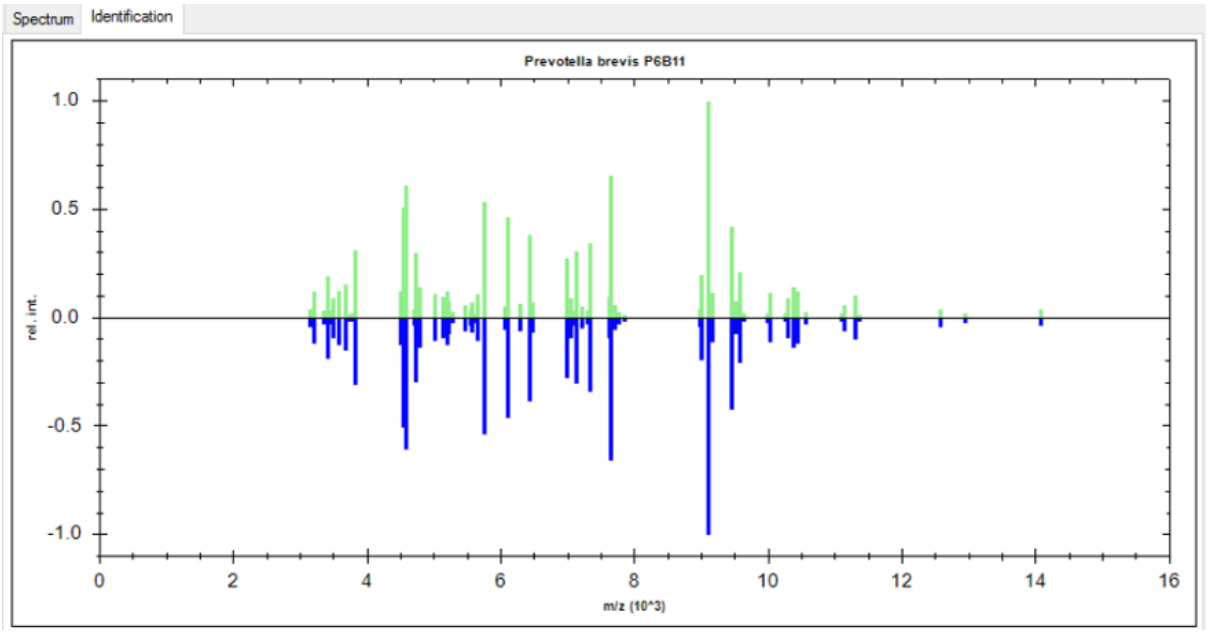
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3005 Da	3006 Da	1.0 Da	343 ppm	PASS
4000 Da to 5000 Da	4265 Da	4267 Da	1.2 Da	288 ppm	PASS
5000 Da to 6000 Da	5182 Da	5183 Da	1.3 Da	260 ppm	PASS
6000 Da to 7000 Da	6421 Da	6423 Da	1.5 Da	234 ppm	PASS
7000 Da to 8000 Da	7302 Da	7305 Da	2.4 Da	329 ppm	PASS
9000 Da to 10000 Da	9593 Da	9597 Da	3.7 Da	383 ppm	PASS
Above 10000 Da	10033 Da	10036 Da	2.8 Da	281 ppm	PASS

A.7. *Prevotella* sp. RM4 (Hun203)



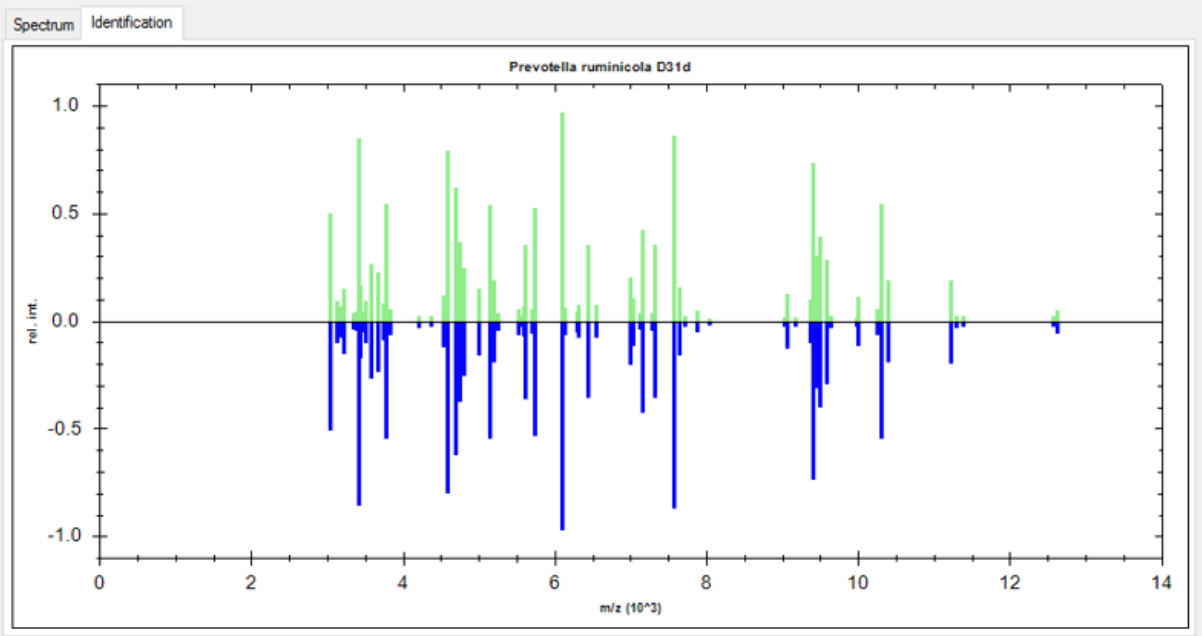
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3041 Da	3042 Da	1.0 Da	339 ppm	PASS
4000 Da to 5000 Da	4577 Da	4579 Da	1.9 Da	415 ppm	PASS
5000 Da to 6000 Da	5168 Da	5170 Da	2.0 Da	389 ppm	PASS
6000 Da to 7000 Da	6085 Da	6087 Da	2.2 Da	360 ppm	PASS
7000 Da to 8000 Da	7159 Da	7162 Da	3.2 Da	443 ppm	PASS
7000 Da to 8000 Da	7849 Da	7853 Da	3.3 Da	422 ppm	PASS
9000 Da to 10000 Da	9392 Da	9397 Da	4.5 Da	483 ppm	PASS

A.8. *Prevotella brevis* P6B11 (Hun167)



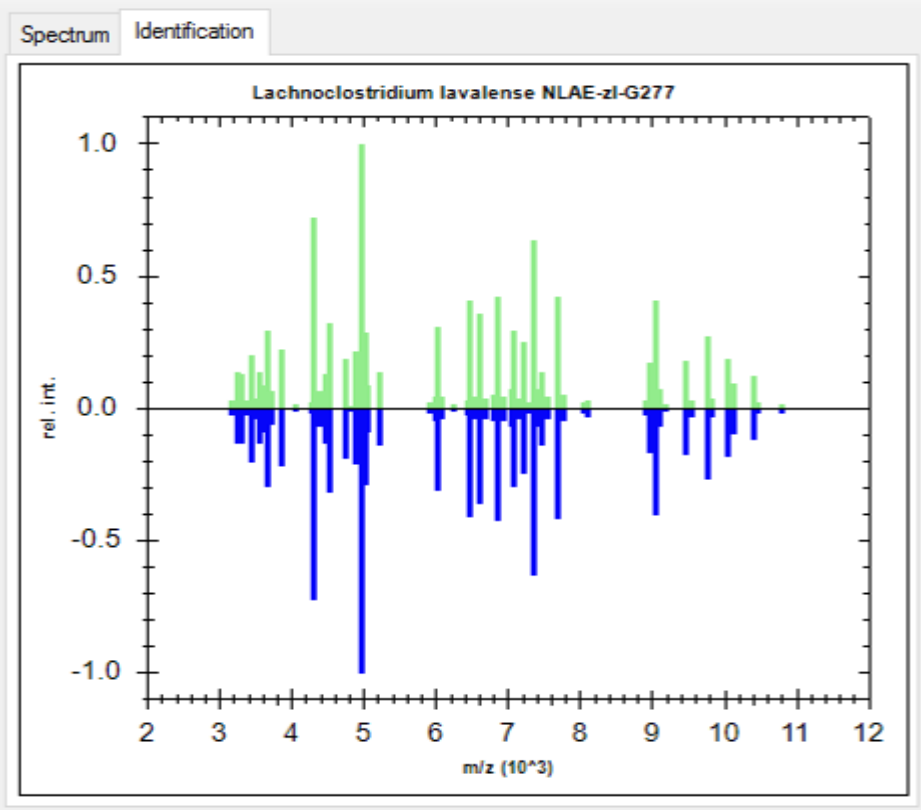
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3049 Da	3050 Da	1.0 Da	338 ppm	PASS
4000 Da to 5000 Da	4546 Da	4547 Da	1.3 Da	277 ppm	PASS
5000 Da to 6000 Da	5013 Da	5015 Da	1.3 Da	263 ppm	PASS
6000 Da to 7000 Da	6100 Da	6102 Da	2.2 Da	359 ppm	PASS
7000 Da to 8000 Da	7137 Da	7140 Da	2.4 Da	332 ppm	PASS
7000 Da to 8000 Da	7640 Da	7643 Da	3.3 Da	428 ppm	PASS
9000 Da to 10000 Da	9090 Da	9093 Da	2.7 Da	295 ppm	PASS

A.9. *Prevotella ruminicola* D31d (Hun463)



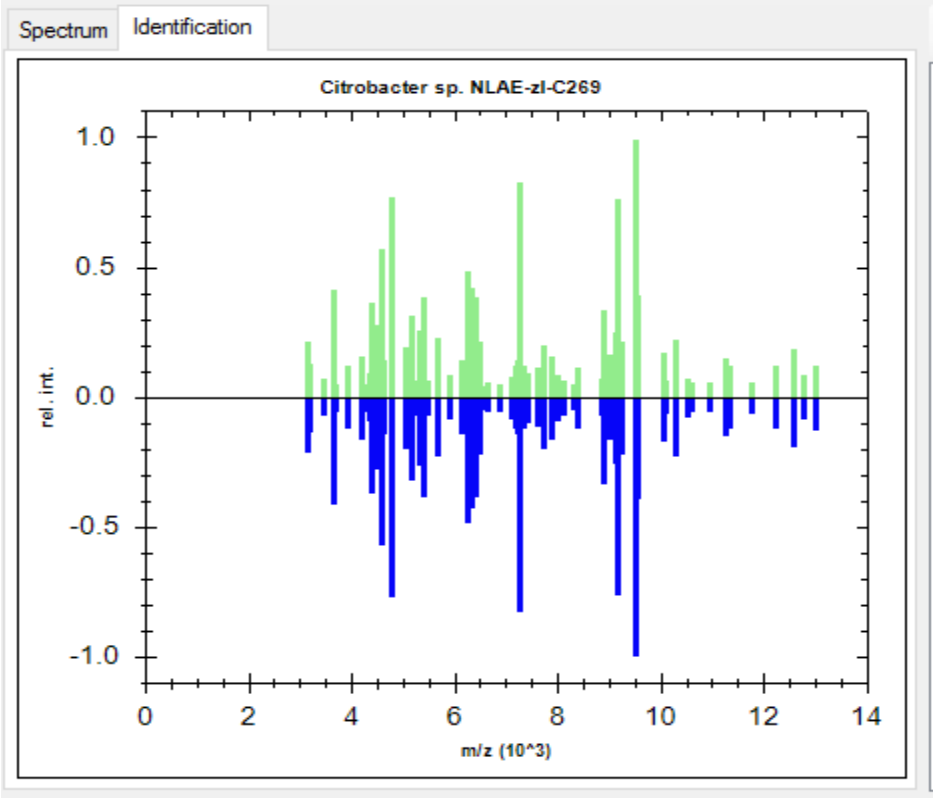
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3041 Da	3042 Da	0.5 Da	168 ppm	PASS
4000 Da to 5000 Da	4577 Da	4578 Da	1.3 Da	277 ppm	PASS
5000 Da to 6000 Da	5150 Da	5151 Da	1.3 Da	260 ppm	PASS
6000 Da to 7000 Da	6085 Da	6087 Da	2.2 Da	360 ppm	PASS
7000 Da to 8000 Da	7159 Da	7161 Da	2.4 Da	332 ppm	PASS
7000 Da to 8000 Da	7865 Da	7868 Da	3.3 Da	422 ppm	PASS
9000 Da to 10000 Da	9392 Da	9396 Da	3.6 Da	387 ppm	PASS

A.10. *Lachnoclostridium lavalense* NLAE-zl-G277 (Hun391)



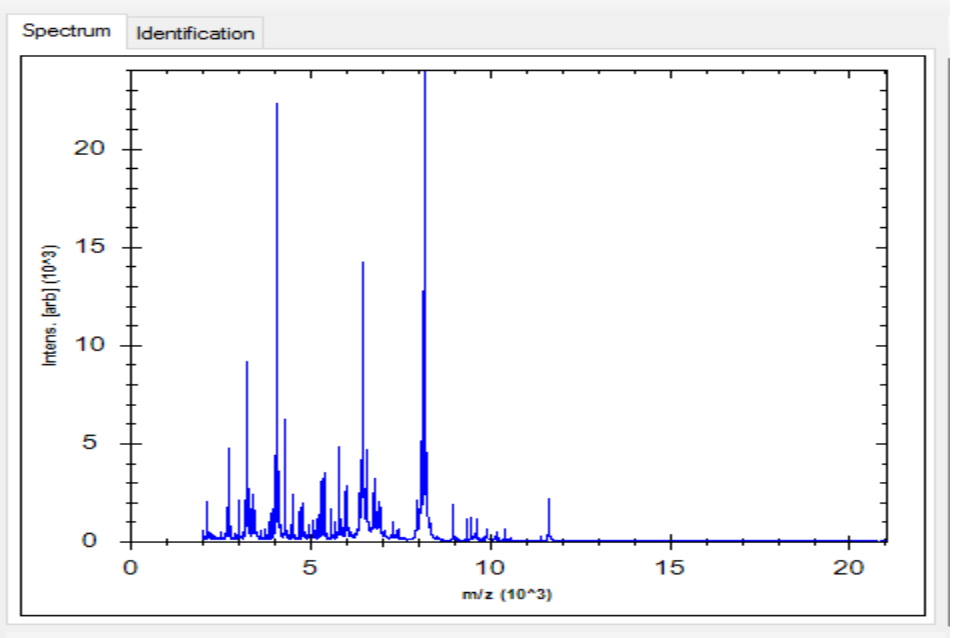
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3002 Da	3003 Da	1.0 Da	340 ppm	PASS
4000 Da to 5000 Da	4308 Da	4310 Da	1.2 Da	283 ppm	PASS
4000 Da to 5000 Da	4958 Da	4960 Da	2.0 Da	399 ppm	PASS
6000 Da to 7000 Da	6009 Da	6010 Da	1.4 Da	241 ppm	PASS
7000 Da to 8000 Da	7078 Da	7081 Da	3.2 Da	445 ppm	PASS
7000 Da to 8000 Da	7696 Da	7699 Da	3.3 Da	427 ppm	PASS
9000 Da to 10000 Da	9036 Da	9040 Da	4.5 Da	493 ppm	PASS

A.11. *Citrobacter* sp. NLAE-zl-C269 (Hun455)



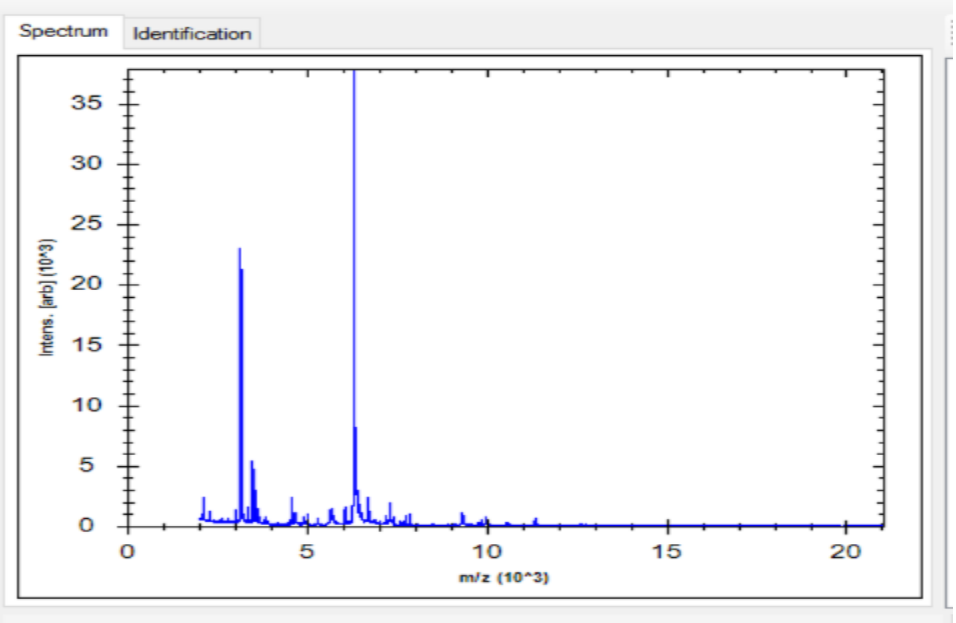
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
4000 Da to 5000 Da	4365 Da	4367 Da	1.9 Da	424 ppm	PASS
5000 Da to 6000 Da	5384 Da	5386 Da	2.1 Da	382 ppm	PASS
6000 Da to 7000 Da	6258 Da	6260 Da	2.2 Da	355 ppm	PASS
7000 Da to 8000 Da	7264 Da	7267 Da	3.2 Da	439 ppm	PASS
8000 Da to 9000 Da	8372 Da	8375 Da	3.4 Da	410 ppm	PASS
9000 Da to 10000 Da	9168 Da	9172 Da	4.5 Da	488 ppm	PASS
Above 10000 Da	10301 Da	10306 Da	4.8 Da	461 ppm	PASS

A.12. *Proteiniclasticum ruminis* DSM 24773 (Hun201)



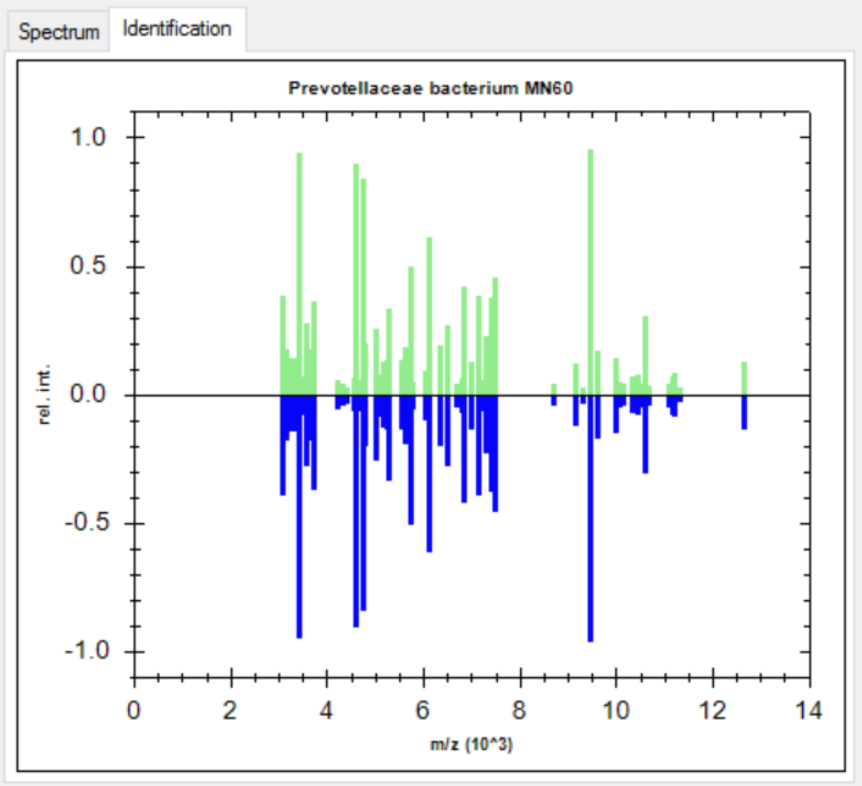
Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
2000 Da to 3000 Da	2997 Da	2999 Da	1.0 Da	344 ppm	PASS
4000 Da to 5000 Da	4087 Da	4088 Da	1.2 Da	291 ppm	PASS
5000 Da to 6000 Da	5329 Da	5331 Da	1.4 Da	255 ppm	PASS
6000 Da to 7000 Da	6460 Da	6463 Da	2.3 Da	350 ppm	PASS
7000 Da to 8000 Da	7309 Da	7311 Da	1.6 Da	219 ppm	PASS
8000 Da to 9000 Da	8173 Da	8175 Da	2.5 Da	311 ppm	PASS
8000 Da to 9000 Da	8992 Da	8993 Da	1.8 Da	198 ppm	PASS

A.13. *Bacteroidales bacterium* WCE2008 (Hun327)



Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
4000 Da to 5000 Da	4564 Da	4566 Da	1.9 Da	416 ppm	PASS
5000 Da to 6000 Da	5304 Da	5306 Da	2.1 Da	386 ppm	PASS
6000 Da to 7000 Da	6054 Da	6057 Da	2.9 Da	482 ppm	PASS
6000 Da to 7000 Da	6315 Da	6318 Da	2.2 Da	353 ppm	PASS
7000 Da to 8000 Da	7288 Da	7291 Da	3.2 Da	439 ppm	PASS
8000 Da to 9000 Da	8501 Da	8505 Da	3.5 Da	407 ppm	PASS
9000 Da to 10000 Da	9307 Da	9311 Da	3.6 Da	389 ppm	PASS

A.14. *Prevotellaceae bacterium* MN60 (Hun374)



Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3061 Da	3062 Da	1.0 Da	340 ppm	PASS
4000 Da to 5000 Da	4579 Da	4580 Da	1.3 Da	277 ppm	PASS
5000 Da to 6000 Da	5292 Da	5294 Da	2.0 Da	385 ppm	PASS
6000 Da to 7000 Da	6127 Da	6129 Da	1.5 Da	238 ppm	PASS
7000 Da to 8000 Da	7136 Da	7138 Da	2.4 Da	332 ppm	PASS
7000 Da to 8000 Da	7475 Da	7478 Da	3.2 Da	433 ppm	PASS
9000 Da to 10000 Da	9452 Da	9456 Da	4.5 Da	481 ppm	PASS

A.15. *Enterobacter* sp. KPR-6 (Hun299)

Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3125 Da	3126 Da	1.0 Da	333 ppm	PASS
4000 Da to 5000 Da	4362 Da	4364 Da	1.2 Da	282 ppm	PASS
5000 Da to 6000 Da	5248 Da	5250 Da	2.0 Da	387 ppm	PASS
6000 Da to 7000 Da	6254 Da	6256 Da	2.2 Da	355 ppm	PASS
7000 Da to 8000 Da	7259 Da	7262 Da	2.4 Da	331 ppm	PASS
7000 Da to 8000 Da	8276 Da	8278 Da	1.7 Da	207 ppm	PASS
9000 Da to 10000 Da	9502 Da	9504 Da	2.7 Da	288 ppm	PASS

A.16. *Bacteroides ovatus* NLAE-zl-C500 (Hun462), spectral picture missing.

Mass range	Lower Mass	Upper Mass	Mass Deviation	Relative Mass Deviation (RMD)	Pass (if RMD <500 ppm)
3000 Da to 4000 Da	3144 Da	3145 Da	1.6 Da	499 ppm	PASS
4000 Da to 5000 Da	4588 Da	4590 Da	1.9 Da	414 ppm	PASS
4000 Da to 5000 Da	5792 Da	5795 Da	2.9 Da	492 ppm	PASS
6000 Da to 7000 Da	6359 Da	6362 Da	3.0 Da	470 ppm	PASS
7000 Da to 8000 Da	7505 Da	7508 Da	3.3 Da	433 ppm	PASS
9000 Da to 10000 Da	9296 Da	9299 Da	3.6 Da	388 ppm	PASS
Above 10000 Da	10429 Da	10433 Da	4.8 Da	458 ppm	PASS

Appendix 2. 16S rRNA sequences used for blast search

Sequence ID	Sequence bases
Rumen_Betaine_RB04_27F	TATTCTTAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTCATACTGGGGGATAACAGCTGGAAACGACTGTTAATACCGCATAAGCGCACRGTGTCGCATGACACAGTGTGAAAAACTCCGGTGGTATGAGATGGGCCCGCGTCAGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGCAACGATCTGTAGCCGGACTGAGAGGTCGGACGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTGGGGGATATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAGGCTCGCAAGAGAGATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGACGG
Rumen_Betaine_RB05_27F	GGGTGAGTAACGCGTGGGTAACCTGCCTCATACTGGGGGATAACAGCTGGAAACGACTGTTAATACCGCATAAGCGCACRGTGTCGCATGACACAGTGTGAAAAACTCCGGTGGTATGAGATGGGCCCGCGTCAGATTAGCCAGTTGGCGGGGTAACGGCCCACCAAAGCAACGATCTGTAGCCGGACTGAGAGGTCGGACGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTGGGGGATATTGCACAATGGAGGAAACTCTGATGCAGCGACGCCGCGTGAGTGAAGAAGTATTTCGGTATGTAAAGCTCTATCAGCAGGGAAGAAAGGCTCGCAAGAGAGATGACGGTACCTGACTAAGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGACGGTTATGCAAGTCTGAA
Rumen_Betaine_RB06_27F	TGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTACTAGCGGGGGATAACTATTGGAAACGATAGCTAATACCGCATAACAGCATTTAACACATGTTAGATGCTTGAAAGGAGCAATTGCTTCACTAGTAGATGGACCTGCGTTGTATTAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGAGAAGAACGTGTGTGAGAGTGGAAAGTTCACACAGTGACGGTAACTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCC

	GAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTAATAAGTCTGAAGTTAAA
Rumen_Betaine_RB47_A_27F	TGCAGTCGACGAAGCACTTATYACGATCCCTTCGGGGTGACGATTTGTTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTACCTTGTACAGGGGGACAACAGTTGGAAACGACTGCTAATACCGCATAAGCGCAC
Rumen_Betaine_RB48_B_27F	TGCAGTCGACGAAGCAACTTATCACGATCCCTTCGGGGTGACGATTTGTTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTTGTACAGGGGGACAACAGTTGGAAACGACTGCTAATACCGCATAAGCGCACAG
Rumen_Betaine_RB49_C_27F	ACGATCCCTTCGGGGTGACGATTTGTTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTACCTTGTACAGGGGGACAACAGTTGGAAACGACTGCTAATACCGCATAAGCGCAC
Rumen_Betaine_RB50_D_27F	AGTGGCGAACGGGTGAGTAACACGTGACCAACCCGCCCCCTCCTCCGGGACAACCTCGGGAAACCGTGGCTAATACCGGATACTCCGGCTGCAGCGCATGCTGCGGCCGGGAAAGCCCAGACGGGAGGGGATGGGGTCGCGGCCCATCAGGTAGACGGCGGGGCAACGGCCCGCCGTGCCGACGACGGGTAGCCGGGCTGAGAGGCTGATCGGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTTGCGCAATGGGGGGAACCCTGACGCAGCGACGCCGCGTGCGGGACGAAGGCCTTCGGGTCGTAAACCGCTTTCAGCAGGGAAGAACAATGACGGTACCTGCAGAAGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGAGCGAGCGTTATCCGGATTCATTGGGCGTAAAGCGCGCGCAGGCGGCCGGCCAAGCGGGCCTCGTCGAAGCCGTGGGCTCAACCCGCGGAAGCGACCCGAACTGGCCGGCTCGAGTGAGGTAGGGGAGGATGGAATTCCCGGTGTAGCGGTGGAATGCGCAGATATC
Rumen_Betaine_RB51_E_27F	TGCAGTCGAACGAAGCAACTTATTACGATCCCTTCGGGGTGACGATTTGTTGACTGAGTGGCGGACGGGTGAGTAACGCGTGGGTAACCTGCCTTGTACAGGGGGACAACAGTTGGAAACGACTGCTAATACCGCATAAGCGCACAGCTTCGCATGAAGCAGTGTGAAAAGTTTTTC
Rumen_Choline_RCH10_27F	ACCTTCCCTATAGTAGAGAATAGCCCGGTGAAAATCGGATTAATGCTCTATGTTATTCAATGCGGACATCTAAGTTGAATCAAAGGTTTACCGCTATAGGATGGGGATGCGTCTGATTAGGTTGTTGGCGGGGTAACGGCCCACCAAGCCYACGATCAGTAGGGGTTCTGAGAGGAAGGTCCCCCACATTGGAACTGAGACACGGTCCAAACTCCTACGGGAGGCAGCAGTGAGGAATATTGGTCAATGGACGAAAGTCTGAACCAGCCAAGTA

	GCGTGCAGGATGACGGCCCTATGGGTTGTAAACTGCTTTTATACAG GGATAAAGTCGGGAACGTGTTCCCGTTTGTAGGTACTGTATGAATAA GGACCGGCTAATTCCGTGCCAGCAGCCGCGGTAATACGGAAGGT CCGGGCGTTATCCGGATTTATTGGGTTTAAAGG
Rumen_Choline_ RCH11_27F	TGCAGTCGAGGGGCAGCAGATCGAAAGCTTGCTTTTGATGCTGGC GACCGGCGCACGGGTGAGTAACGCGTATCCAACCTTCCCTATAGT AGAGAATAGCCCGGCGAAAGTCGRATTAATGCTCTATGTTGTATTTA GAAGACATCTGAAGAATACCAAAGGTTTACCGCTATAGGATGGGGA TGCGTCTGATTAGGTWGTWGGCGGGGTAACGGCCCACCWAGCC SACGATCAGTAGGGGTTCTGAGAGGAAGGTCCCCCACATTGGAAC TGAGACACGGTCCAAACTCCTACGGGAGGCAGCAGTGAGGAATAT TGGTCAATGGACGRAAGTCTGAACCAGCCAAGTAGCGTGCAGGAT GACGGCCCTATGGGTTGTAAACTGCTTTTATATAGGGATAAAGTCGG GGACGTGTCCCCGTTTGTAGGTACTATATGAATAAGGACCGGCTAAT TCCGTGCCAGCAGCCGCGGTAATACGGAAGGTCCGGGCGTTATC CGGATTTATTGGGTTTAAAGGGAGCGCAGGCTGATGATTAAGCGTG ACGTGAAATGTAGCCGCTCAACGGCTGAACTGCGTCGCGAACTG GTTATC
Rumen_Choline_ RCH22b_27F	ACGATCCCTTCGGGGTGACGATTTGTTGACTGAGTGGCGGACGGG TGAGTAACGCGTGGGTAACCTACCTTGTACAGGGGGATAACAGTTG GAAACGACTGCTAATACCGCATRAGCGCAC
Rumen_Choline_ RCH22b_1492R	GACTTCGGGCATTCCCAACTCCCATGGTGTGACGGGCGGTGTGTA CAAGACCCGGGAACGTATTCACCGCGACATTCTGATTCGCGATTA CTAGCGATTCCAGCTTCATGTAGTCGAGTTGCAGACTACAATCCGA ACTGAGACGTTATTTTTGAGATTTGCTCATRGTCACCTACTTGCTTCC CTTTGTTTACGCCATTGTAGCACGTGTGTAGCCCAAATCATAAGGG GCATGATGATTTGACGTCATCCCCACCTTCCTCCCGGTTCACCCG GGCAGTCTCTCTAGAGTGCCCATCCGAAWTGCTGGCTACTAAAGA TAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACA CGAGCTGACGACAACCATGCACCACCTGTCAC

Appendix 3. Meaning of Bruker MALDI Biotyper® identification scores

Meaning of Score Values

Range	Interpretation	Symbols	Color
2.00 - 3.00	High-confidence identification	(+++)	green
1.70 - 1.99	Low-confidence identification	(+)	yellow
0.00 - 1.69	No Organism Identification Possible	(-)	red

Meaning of Consistency Categories (A - C)

Category	Interpretation
(A)	**High consistency:** The best match is a high-confidence identification. The second-best match is (1) a high-confidence identification in which the species is identical to the best match, (2) a low-confidence identification in which the species or genus is identical to the best match, or (3) a non-identification.
(B)	**Low consistency:** The requirements for high consistency are not met. The best match is a high- or low-confidence identification. The second-best match is (1) a high- or low-confidence identification in which the genus is identical to the best match or (2) a non-identification.
(C)	**No consistency:** The requirements for high or low consistency are not met.