Septic Tank Performance in Cold Climates: Challenges, Stakeholder Perceptions, and Microbial Dynamics

Aiman Uteyeva

Environmental and Water Resource Engineering
Department of Civil and Environmental Engineering
Imperial College London

Submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy at Imperial College London

Statement of originality

I hereby declare that the work presented in this thesis is my own. Any contribution made to this thesis by others is explicitly acknowledged.

I acknowledge the use of ChatGPT-5.2 (OpenAI, https://chatgpt.com) to assist with drafting, refining, and modifying R and Python scripts, as well as supporting minor revisions to Figures 2.1 and C.6, which were prepared by the author. All figures and tables were prepared and finalised by the author, and all analyses, interpretations, and written content presented in this thesis constitute the author's original work. No content generated by artificial intelligence has been presented as the author's own without critical review and verification.

Signature:

_____

Aiman Uteyeva

Copyright declaration

Abstract

This research investigated how extreme seasonal temperature fluctuations, ranging from $-27\,^{\circ}\mathrm{C}$ in winter to $31\,^{\circ}\mathrm{C}$ in summer, affect the performance and microbial functioning of septic tank systems in cold-climate settings. Focusing on Kazakhstan, the study presents an integrated assessment combining stakeholder surveys, year-round field monitoring, and shotgun metagenomic analysis to examine how decentralised sanitation systems respond to natural temperature cycles.

Three aims were pursued: (i) examining stakeholder perceptions of sanitation challenges; (ii) assessing year-round septic tank performance under *in situ* thermal conditions; and (iii) characterising microbial and viral communities across seasonal temperature gradients.

A structured survey of 30 stakeholders across Kazakhstan's two coldest cities identified recurring challenges: inappropriate installation depth relative to frost penetration, reduced winter accessibility, irregular desludging, and limited technical oversight, providing new empirical evidence for context-specific management and policy responses.

Field monitoring of ten septic tanks over a complete annual cycle was supported by laboratory-scale, temperature-controlled anaerobic reactor experiments. Metagenomic analysis showed that anaerobic digestion was sustained by complex microbial consortia, including fermentative bacteria, syntrophic acetate-oxidising bacteria, sulphate-reducing bacteria, and hydrogenotrophic methanogens, with acetoclastic methanogens identified in septic tanks for the first time. Seasonal temperature variation was associated with coordinated shifts in microbial community composition and functional gene abundance linked to acetate activation, methanogenesis, and sulphate reduction.

Pathogen-associated bacteria persisted across all seasons and temperature conditions, indicating continuous public health exposure risk. Low temperatures selected for psychrotolerant Moraxellaceae; elevated temperatures enriched Pseudomonadaceae and Arcobacteraceae, taxa with clinical relevance and biofilm-forming capacity. Viral communities dominated by Caudoviricetes closely tracked prokaryotic dynamics, with significant correlations to anaerobic digestion guilds.

These findings provide mechanistic insight into temperature-driven functional

reorganisation in septic tanks, supporting the development of more resilient sanitation strategies with implications for public health, groundwater protection, and climate resilience.

To my late father — your pride still guides my steps.

I honour you with this work.

Acknowledgements

My deepest gratitude goes to my husband, Dr. Rakhim Uteyev, whose unwavering patience, quiet encouragement, and steadfast belief in me carried me through every challenge of this journey. To our three children — Amir, Akhmet, and Amina — you have been my constant source of strength, motivation, and joy, grounding me in ways beyond anything I could have imagined.

I am equally grateful to my parents, whose boundless love, trust, and belief laid the foundations for this academic path long before it formally began. This PhD represents a continuation of an educational journey that started in high school, where I graduated with a gold medal (*summa cum laude*) from the Kazakh Turkish Girls High School, and continued through my Bachelor's degree in Biology with Honours at the George Washington University in the United States and my Master's studies in Economics and Public Policy at Ecole Polytechnique and Sciences Po Paris in France. This path was made possible because my parents placed their confidence in my ambitions without hesitation. At a time when such trust and investment are not always afforded to daughters, they gave me not only their support, but also the freedom to follow my intellectual curiosity, encouraging perseverance, discipline, and a deep love of learning from an early age. Their willingness to trust my capabilities, to value effort as much as achievement, and to stand beside me through every stage of my education has made this work possible. Together with my husband and children, my family created an environment in which pursuing and bringing this work to its current stage was not only possible, but inevitable. This work, in many ways, belongs to all of you as much as it does to me.

I am deeply grateful to my supervisor, Professor Michael R. Templeton, for his continuous guidance, constructive feedback, and unwavering support throughout this doctoral journey. His encouragement, high standards, and trust in my abilities shaped not only the direction of this research but also my development as a researcher and as a person. It has been a privilege to work under his mentorship.

I also extend my sincere thanks to my second supervisor, Professor Sue Grimes, for her support and encouragement throughout the course of this PhD. Her insight and guidance contributed meaningfully to the development of this work.

I am thankful to Professor Woojin Lee, Professor Almagul Kushugulova, Mrs Firyuza Aldyngulova, Dr Snezhanna Merkuryeva, Professor Luis Rojas-Solorzano, Dr Aidar Zhakupov, and Dr Nazira Moldagulova for their professional guidance, collaboration, and generous contributions. I am particularly grateful to Professor Woojin Lee for hosting me as a visiting PhD scholar at Nazarbayev University; to Professor Almagul Kushugulova for facilitating access to the facilities of the Laboratory of Human Microbiome and Longevity, National Laboratory Astana, Nazarbayev University; to Professor Luis Rojas-Solorzano and Dr Aidar Zhakupov for their support in identifying and connecting with appropriate laboratory facilities at Nazarbayev University; to Mrs Firyuza Aldyngulova for facilitating institutional access to the wastewater treatment plant (WWTP) and authorising support for wastewater and inoculum sampling; to Dr Snezhanna Merkuryeva for facilitating access to laboratory facilities and resources for the physicochemical analysis of wastewater samples and inoculum; and to Dr Nazira Moldagulova for providing access to incubation facilities. Their shared commitment to strengthening water and sanitation systems in Kazakhstan has shaped both the direction and the impact of this research.

My special thanks go to Rakhymbek Abildin, technician at the WWTP, for his invaluable practical assistance, including the collection of inoculum from the WWTP and wastewater samples from septic tanks into prepared sampling bottles, as well as hands-on support during challenging fieldwork, particularly in lifting and accessing septic tank manholes.

I am deeply grateful to all stakeholders who participated in the survey and generously shared their time, insights, and experiences. I also extend my sincere appreciation to the households who granted access to their septic tanks and supported year-long monitoring and sampling; their trust and cooperation made this fieldwork possible.

Finally, I wish to extend my heartfelt gratitude to my country, the Republic of Kazakhstan, and its people for sponsoring my doctoral studies and believing in my potential to contribute to our nation's sustainable future. This studentship was fully funded by the Government of the Republic of Kazakhstan through the Centre for International Programmes.

This PhD study was conducted with the approval from the Head of Department, approval from the Research Governance and Integrity Team (RGIT), and favourable opinion from the Science, Engineering and Technology Research Ethics Committee

(SETREC reference 6417731).

Contents

List of Figures

List of Tables

Abbreviations

The abbreviations and acronyms used throughout this thesis are listed below.

Abbreviation	**Definition**
AA	Acetate activation
AD	Anaerobic digestion
AMG	Auxiliary metabolic gene
APHA	American Public Health Association
BOD_5	Biochemical Oxygen Demand (5-day)
bp	Base pair
$CH_3CH_2COO^-$	Propionate
CH_3COO^-	Acetate
CH_3COOH	Acetic acid
CH_4	Methane
CO_2	Carbon dioxide
COD	Chemical Oxygen Demand
Core-M	Core methanogenesis
CST	Conventional septic tank
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
FDR	False Discovery Rate
GC	Gas chromatography
H_2	Hydrogen gas
H_2S	Hydrogen sulphide
HDPE	High-density polyethylene

Continued on next page

Abbreviation	Definition
HICs	High-income countries
HRT	Hydraulic retention time
iPHoP	Integrated Phage–Host Prediction tool
IQR	Interquartile range
JMP	Joint Monitoring Programme
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KO	KEGG Orthology
KoM	King-of-the-Mountain
KtW	Kill-the-Winner
LMICs	Low- and middle-income countries
MCR	Methyl-coenzyme M reductase
NCBI	National Centre for Biotechnology Information
NGO	Non-governmental organisation
N_2O	Nitrous oxide
NMDS	Non-metric Multidimensional Scaling
PCoA	Principal Coordinates Analysis
PERMANOVA	Permutational Multivariate Analysis of Variance
PtW	Piggyback-the-Winner
RefSeq	Reference Sequence
SAO	Syntrophic acetate oxidation
SAOB	Syntrophic acetate-oxidising bacteria
SD	Standard deviation
SDG	Sustainable Development Goal
SR	Sulphate reduction
SRB	Sulphate-reducing bacteria

Continued on next page

Abbreviation	Definition
vOTUs	Viral operational taxonomic units
TS	Total solids
UN	United Nations
UNICEF	United Nations International Children's Emergency Fund
VFA	Volatile fatty acid
VS	Volatile solids
WASH	Water, sanitation and hygiene
WHO	World Health Organisation
WL	Wood-Ljungdahl pathway
WWTP	Wastewater treatment plant

1 Introduction

1.1 Background and context

Access to safe sanitation was formally recognised as a human right by the United Nations General Assembly in 2015 (UN General Assembly, 2015). Although meaningful progress has been achieved over the past decade, current advancements remain far too slow to meet global targets. Projections indicate that by 2030, approximately 2.8 billion people will still lack access to safely managed sanitation services (United Nations, 2015). The World Health Organisation (WHO) and the United Nations International Children's Emergency Fund (UNICEF) Joint Monitoring Programme (JMP) further reports that more than half of the global population is not connected to centralised sewerage networks (WHO/UNICEF JMP, 2022b). To reach the water, sanitation, and hygiene (WASH) targets established under the Sustainable Development Goals (SDGs), the pace of global progress would need to increase fourfold (United Nations, 2021). Achieving this scale of transformation requires approaches that extend beyond conventional systems, with a particular emphasis on technologies that are affordable, scalable, and resilient to external pressures, including climatic variability.

Cold climates present a distinct set of challenges for sanitation infrastructure. Around 2.5 billion people live in regions where air temperature falls below 1 °C for at least one month each year (Leblanc et al., 2019; World Bank, 2022). These conditions exert a significant influence on the functionality, reliability, and cost-effectiveness of on-site sanitation systems (Leblanc et al., 2019). A key constraint is the formation of frozen soil, which becomes impermeable during winter, limiting wastewater infiltration and increasing the risk of soil and groundwater contamination (Buttle and Smith, 2004; Leblanc et al., 2019). Reduced air exchange in clogged or frozen soil can also cause hydraulic failure and wastewater overflow, further exacerbating environmental and public health risks (Leblanc et al., 2019; Pishgar et al., 2021).

At the same time, low temperatures suppress the biological and physical processes essential for faecal sludge stabilisation, slowing anaerobic digestion (AD) and reducing

treatment efficiency. These combined factors make sanitation management in cold, remote, and resource-constrained regions especially challenging (Williamson, 2010; Bian et al., 2024). Addressing these constraints requires targeted, climate-responsive interventions and innovative design strategies tailored to the operational realities of cold environments. Advancements in this area are essential not only for protecting public health but also for supporting broader progress towards sustainable development.

1.2 Research problem and significance

Among improved on-site sanitation technologies, septic tank systems are often regarded as reliable due to their relative effectiveness, simplicity, and adaptability across diverse settings. However, despite their widespread adoption, these systems present substantial financial, operational, and environmental challenges. The costs associated with construction, desludging, and routine maintenance can be prohibitive for many households, limiting accessibility and long-term sustainability. Furthermore, malfunctioning or inadequately maintained septic tanks have been identified as persistent and often under-recognised sources of environmental pollution, contributing to the degradation of soil and groundwater quality (Withers et al., 2014). A key contributor to these failures is the requirement for periodic emptying, which imposes additional service demands and financial burdens on users (Leblanc et al., 2019).

These constraints are exacerbated in cold climates, where low ambient temperatures impose additional limitations on both system performance and affordability. Although temperature is a well-established driver of biological and physicochemical processes within wastewater treatment systems (Koottatep et al., 2014b; Pussayanavin et al., 2015), limited empirical research has directly examined how extremely low temperatures influence the functionality and resilience of septic tanks. This lack of evidence represents a critical knowledge gap with implications for the design, management, and long-term viability of sanitation systems in cold and temperate regions.

Addressing this gap necessitates a deeper understanding of the microbial processes that underpin septic tank operation. Microbial communities mediate the degradation of organic matter and transformation of nutrients, and these processes are highly sensitive to temperature variation (Mara and Horan, 2008). Fluctuations in temperature can alter

microbial activity, growth kinetics, and community structure, with direct consequences for treatment efficiency and system stability (Veeken and Hamelers, 1999; Gerardi, 2003).

This PhD research investigated the response of microbial populations in septic tanks to natural temperature variability, providing new insights into the performance, limitations, and adaptation potential of these systems in cold climates. By interpreting the temperature-dependent dynamics of key microbial groups, the study contributes to the development of more resilient and sustainable on-site sanitation technologies. Ultimately, this work supports broader global efforts to minimise pollution risks, enhance public health protection, and expand equitable access to safely managed sanitation.

1.3 Kazakhstan's sanitation context

Kazakhstan, officially the Republic of Kazakhstan, is a Central Asian nation extending across 2.72 million square kilometres, making it the ninth largest country in the world (Sinor et al., 2023). Despite its considerable landmass, it has one of the lowest population densities globally, with an estimated population of approximately 20 million (OECD, 2016; Bureau of National Statistics, 2023). The country's climate is markedly diverse and extreme, characterised by wide-ranging temperature fluctuations, arid conditions, and pronounced seasonal variability (Sinor et al., 2023). In winter, temperatures can fall as low as $-48\,^{\circ}\mathrm{C}$ to $-42\,^{\circ}\mathrm{C}$ in northern regions and $-5\,^{\circ}\mathrm{C}$ to $-1\,^{\circ}\mathrm{C}$ in the south, while summer temperatures may reach $19\,^{\circ}\mathrm{C}$ to $23\,^{\circ}\mathrm{C}$ in the north and up to $40\,^{\circ}\mathrm{C}$ to $45\,^{\circ}\mathrm{C}$ in the south (Kazhydromet, 2023a; Sinor et al., 2023).

Sanitation coverage across Kazakhstan reflects a combination of centralised and decentralised approaches. According to WHO/UNICEF JMP (2022b), 54% of the population (10.72 million people) rely on pit latrines, 9% (1.72 million people) use septic tanks, and the remaining 37% (7.52 million people) are connected to centralised sewer networks. These figures highlight persistent disparities in access to safely managed sanitation, particularly between urban centres and rural or peri-urban areas. They also emphasise the continued importance of decentralised systems for a substantial share of the population. Given the country's extreme climatic conditions, Kazakhstan represents a particularly relevant context for examining the performance, limitations, and resilience of septic tanks in cold environments.

1.4 Aims and objectives

The overarching aim of this PhD research is to investigate how temperature fluctuations cause problems for households using septic tanks and shape microbial dynamics within septic tanks, with a particular focus on cold-climate environments. By developing a deeper understanding of these temperature-dependent processes, the study seeks to inform practical, sustainable strategies for improving the performance and resilience of on-site sanitation systems operating under extreme environmental conditions.

Research questions

1) What are stakeholder perceptions of the sanitation challenges and feasible solutions related to on-site sanitation in Kazakhstan's coldest cities?

2) Which archaea and bacteria drive anaerobic digestion processes in septic tanks throughout the year in such settings?

3) How do archaea, bacteria, and bacteriophages interact within septic tanks in such settings, and how do these interactions vary with temperature?

Research objectives

1) To examine current sanitation challenges and stakeholder-identified opportunities for improving on-site sanitation practices in Kazakhstan's coldest cities.

2) To evaluate the year-round performance of septic tanks operating under natural field conditions in regions characterised by extreme seasonal temperature fluctuations.

3) To identify the key bacteria and archaea that drive anaerobic digestion processes across varying temperature regimes.

4) To investigate and understand the dynamics of microbial communities and phage–host interactions within septic tanks exposed to a wide range of temperature conditions.

This research highlights the urgent need for sustainable sanitation solutions to extreme climates and provides insights that are relevant both locally and globally. In doing so, the

study contributes to advancing environmental sustainability, safeguarding public health, and supporting progress towards the United Nations Sustainable Development Goals (United Nations, 2015):

- SDG 3: Good Health and Well-being
- SDG 6: Clean Water and Sanitation
- SDG 11: Sustainable Cities and Communities
- SDG 13: Climate Action
- SDG 15: Life on Land

1.5 Scope and limitations

This PhD research investigated the microbial dynamics of septic tanks operating under the extreme seasonal temperature fluctuations characteristic of cold-climate regions. The study employed a dual-method approach that integrated stakeholder perspectives with laboratory-based physicochemical and molecular analyses. Geographically, the research focused on Kazakhstan's coldest cities, where ambient temperatures routinely span from $-40\,^{\circ}\mathrm{C}$ in winter to $40\,^{\circ}\mathrm{C}$ in summer. Within this context, the work combined qualitative and quantitative methods, including stakeholder surveys, systematic wastewater sampling, and advanced molecular characterisation of microbial communities. Collectively, these methods provide a comprehensive understanding of the performance of decentralised sanitation systems subject to severe environmental stressors.

A central component of this research was the long-term investigation of septic tank performance under field conditions. Particular emphasis is placed on the identification of methanogenic archaea, anaerobic bacteria, and bacteriophages, as well as their interactions across a wide temperature spectrum. By analysing microbial dynamics *in situ*, the study aimed to clarify the biological processes underpinning anaerobic digestion in cold climate environments. The knowledge generated offers an empirical foundation for designing and optimising on-site sanitation technologies that remain operationally efficient, environmentally robust, and sustainable despite extreme seasonal variability.

While this research yields insights, a few constraints must be acknowledged. The geographical focus on Kazakhstan means that the findings may not be directly transferable to other cold-climate regions with different socio-economic, infrastructural, or environmental contexts. The microbial analyses are based on samples collected under naturally occurring temperature conditions present at the time of sampling; although ecologically realistic, this may not capture the full spectrum of microbial responses at all seasonal extremes. Additionally, the stakeholder surveys involved a relatively small sample size, enabling in-depth qualitative exploration but potentially limiting broader representativeness. Despite these constraints, the study has established a rigorous foundation for future research and offers practical guidance for improving the design, management, and resilience of sanitation systems in cold environments.

1.6 Thesis outline

This thesis is structured to provide a coherent and comprehensive examination of the research problem, methodology, findings, and broader implications.

- *Chapter 2: Literature Review* surveys the current state of knowledge on the physicochemical characteristics of wastewater, the molecular biology underpinning biological treatment processes, and the specific challenges associated with sanitation management in cold climates. It also identifies key gaps that motivate this research.
- *Chapter 3: Materials and Methods* describes the design and implementation of the stakeholder survey conducted in Kazakhstan's two coldest cities, followed by a detailed account of the physicochemical and molecular laboratory methods used in the study.
- *Chapter 4: Stakeholder Survey Results* presents the findings from the consultations, highlighting the primary constraints and potential opportunities for improving sanitation practices in cold-climate environments.
- *Chapter 5: Results on Physicochemical and Microbial Dynamics in Laboratory-Scale Anaerobic Systems* presents the outcomes of controlled laboratory experiments, examining the physicochemical behaviour and associated microbial dynamics across defined temperature regimes.

- *Chapter 6: Results on Physicochemical and Microbial Dynamics in Household Septic Tanks* reports findings from year-round field monitoring, examining seasonal temperature effects on wastewater physicochemical characteristics and microbial community structure under real operating conditions.

- *Chapter 7: Discussion* synthesises the qualitative and quantitative findings, interpreting their implications for improving the design, operation, and performance of on-site sanitation systems in cold regions. The chapter contextualises the results within the existing literature and highlights how temperature-driven insights from this study can inform more resilient septic tank management strategies.

- *Chapter 8: Conclusions* directly addresses the research questions by summarising the principal findings and contributions of the study, and articulates their relevance for sustainable sanitation management, microbial ecology, and evidence-based policy development.

1.7 List of publications

The author's journal publications are listed below. Parts of Chapters 1, 3, 4, and 7 incorporate text and findings previously published in Uteyeva *et al.* (2024). The material has been expanded and refined for this thesis and forms the basis of two additional academic papers that are currently under review or in preparation:

Publication

- Uteyeva, A., Lee, W., Templeton, M.R. (2024) Stakeholder perceptions, challenges, and sustainable solutions for achieving safely managed sanitation in Kazakhstan: A case study from a cold region. *Journal of Water, Sanitation and Hygiene for Development*, 14(10): 938-949. doi: `https://doi.org/10.2166/washdev.2024.067`

Manuscripts currently in preparation or under review

- *Seasonal temperature control of anaerobic digestion pathways in cold climate septic tanks revealed by shotgun metagenomics.* Target journal: *Water Research.*

- *Temperature-dependent phage–host interactions and the regulation of anaerobic microbial communities in cold climate septic tanks.* Target journal: *Environmental Science: Water Research & Technology.*

2 Literature review

2.1 Introduction

Improving on-site sanitation in regions characterised by strong seasonal extremes remains a global challenge. While decentralised systems such as septic tanks serve approximately 2 billion people (WHO/UNICEF JMP, 2025), their treatment performance depends heavily on anaerobic microbial processes that are highly temperature sensitive. In cold climates, where winter air temperatures routinely fall below $-30\,^{\circ}\mathrm{C}$, reduced biological kinetics can slow organic matter degradation, promote solids accumulation, and compromise treatment stability, leading to increased risks of odour, overflow, and impaired pathogen attenuation. Despite these challenges, empirical evidence describing the year-round biological functioning of septic tanks under extreme seasonal temperature variability remains limited.

This chapter reviews the literature across three interconnected domains central to this thesis:

1. Sanitation systems in cold climates, including technical performance constraints and governance challenges.
2. Stakeholder engagement and institutional dynamics, which shape sanitation design, maintenance, and sustainability.
3. Microbial and viral processes underpinning anaerobic digestion, with emphasis on temperature-dependent functional pathways relevant to septic tanks.

The chapter identifies knowledge gaps surrounding cold-climate operation, microbial community structure, phage–host dynamics, and user/institutional barriers. These gaps provide the rationale for the integrative research design adopted in this thesis, which combines stakeholder surveys, field-scale physicochemical and metagenomic monitoring, and controlled laboratory experiments.

2.2 Sanitation systems and challenges in cold climates

2.2.1 Global sanitation context and persistent inequalities

Sanitation is a fundamental determinant of public health and environmental protection, recognised under SDG 6 (United Nations, 2015). Despite substantial global progress, the WHO/UNICEF JMP (2022b) reports that nearly half of the world's population still lacks access to safely managed sanitation. In many rural, peri-urban, and climatically extreme settings, centralised sewerage systems remain economically prohibitive or technically unfeasible (Leblanc et al., 2019). As a result, decentralised sanitation systems – predominantly septic tanks – play a critical role in bridging service gaps and extending basic sanitation coverage.

However, the majority of technological innovations, performance standards, and monitoring frameworks for on-site sanitation have been developed under tropical or temperate conditions, where seasonal temperature variability is limited. The geographic and climatic bias constrains the applicability of existing design assumptions in regions experiencing prolonged and severe cold seasons. Consequently, sanitation infrastructure in cold climates frequently underperforms, with adverse implications for treatment efficiency, sludge accumulation, groundwater quality, and public health protection.

2.2.2 Septic tanks: configuration and typical performance constraints

Conventional septic tanks (CSTs) comprise two chambers separated by a dividing wall and equipped with manholes and ventilation pipes to facilitate maintenance and limited aeration (Figure 2.1) (Nadi et al., 2017). They are commonly constructed from reinforced concrete, polyethylene, or fibreglass (Bedinger et al., 1997). Wastewater enters the first chamber, where solids settle as sludge through sedimentation, while fats and oils rise to form a scum layer at the surface (Sperling, 2007). This stratification promotes anaerobic conditions essential for microbial degradation (Sperling, 2007; Shaw, 2021). The effluent then flows into the second chamber for further treatment before being discharged into a drain field or infiltration system (Sperling, 2007; Pishgar et al., 2021).

Despite its apparent simplicity, the performance of CSTs is influenced by numerous operational and contextual factors, including adequate hydraulic retention time (HRT),

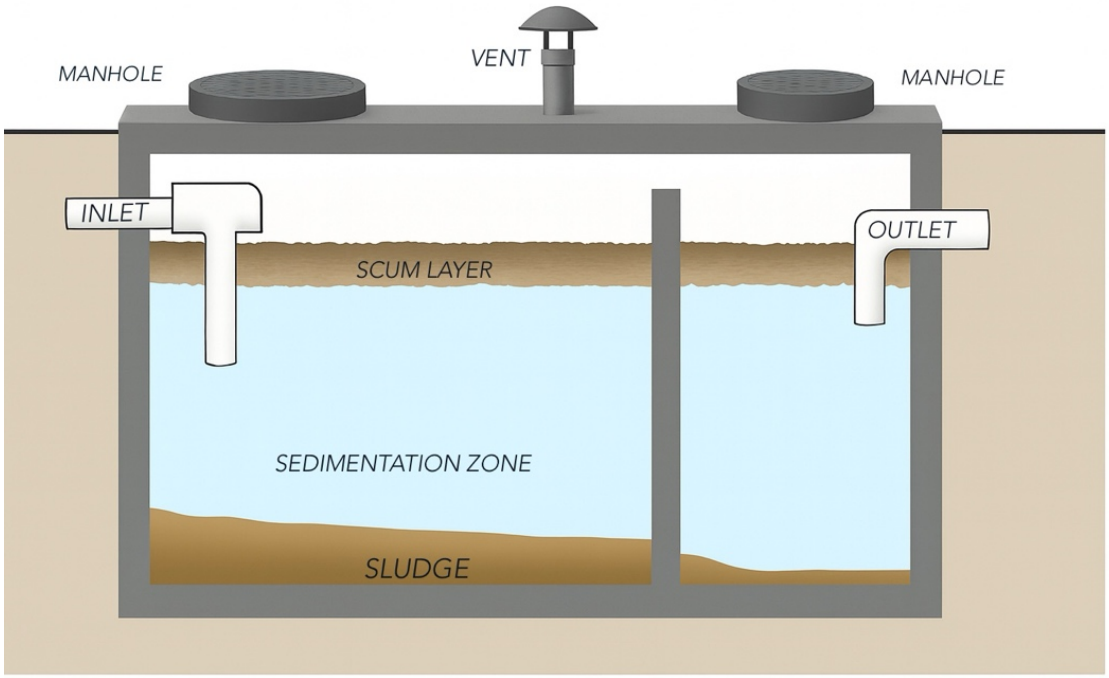

Figure 2.1: A two-chamber conventional septic tank, adapted from Adegoke and Stenstrom (2019).

active microbial consortia, regular desludging, and proper installation depth relative to frost line (Richards et al., 2016). Performance declines when these conditions are not met. Studies from high-income countries (HICs) highlight failures even under strong regulatory frameworks, suggesting that robust governance is essential but not sufficient without good operational practice (Withers et al., 2014).

2.2.3 Cold-climate stresses on septic tank performance

Cold climates impose compound stresses on septic tank systems, simultaneously constraining biological treatment processes, compromising physical infrastructure, and limiting the practical applicability of engineering interventions. An estimated 2.5 billion people live in regions where sub-zero temperatures supress microbial activity, freeze sanitation infrastructure, and reduce soil permeability (Leblanc et al., 2019; World Bank, 2022, 2023), yet the evidence base for effective cold-climate sanitation solutions remains remarkably thin.

At the biological level, anaerobic digestion rates decline sharply under low-temperature conditions, resulting in excessive sludge accumulation and deteriorated effluent quality

(Leblanc et al., 2019). Laboratory-scale investigations have demonstrated that inoculating reactors with naturally cold-adapted microbial communities derived from Arctic soils can sustain key anaerobic digestion processes – including hydrolysis and methanogenesis – at temperatures as low as 4 °C (Bowen et al., 2014; Petropoulos et al., 2017). However, these findings are largely derived from short-term, controlled experiments, and their applicability to household-scale systems operating under variable environmental conditions remains uncertain. The lack of long-term (i.e. across an entire year) field research therefore represents a significant gap in current literature.

Physical stresses further compound these biological limitations. Frozen soils can obstruct infiltration trenches, while ice formation may damage tanks, ventilation systems, and pipework (Buttle and Smith, 2004; Bian et al., 2024). Engineering adaptations have been proposed to mitigate the impacts of low temperatures, including deep burial, double-layer insulation, heat-storage coils, and integration with solar-assisted systems (Polprasert et al., 2018; Koottatep et al., 2020a,c; Bian et al., 2024). Constructed wetlands have also demonstrated potential as post-treatment in cold regions such as Norway and northern China, maintaining treatment performance even under ice-covered conditions (Giaever, 2000; Jenssen, 2002; Zhang et al., 2009). However, these technological adaptations often entail high installation and maintenance costs and require technical expertise seldom available in low- and middle-income countries (LMICs), making them largely unsuitable for rural or resource-constrained settings.

Governance frameworks play an equally critical role in determining system performance alongside environmental conditions. In HICs, regulatory standards, such as EN 12566 in Europe and ANSI/NSF 40 in the United States are supported by municipal inspection regimes and financial mechanisms (Nonet, 2005). Yet even under strong governance structures, monitoring lapses and financial constraints can impede compliance, and effluents with elevated nutrient and pathogen loads continue to be reported (Richards et al., 2016, 2017). In Scandinavian countries, on-site sanitation systems are widely used in sparsely populated areas, but their performance remains variable despite relatively well-developed regulatory frameworks, with effectiveness depending not only on regulatory design but also on enforcement capacity, stakeholder coordination, and user awareness (Laukka et al., 2022; Kinnunen et al., 2023). Empirical evidence from Finland and Sweden further demonstrates that inconsistent treatment performance and non-compliance –

particularly for nutrient removal – are more strongly linked to construction quality, maintenance practices, and operational conditions than to climate alone (Kinnunen et al., 2023). In LMICs, the absence of robust regulatory frameworks is compounded by limited technical capacity and scarce financial resources. In Russia's northern Tyumen region, for example, septic tank effluents have been reported with chemical oxygen demand (COD) concentrations as high as 10,000 mg/L, far exceeding permissible limits, while untreated wastewater is frequently discharged into frozen river systems (Vialkova et al., 2020). In Mongolia, sanitation systems are similarly constrained by freezing temperatures, which inhibit biological treatment, prevent infiltration, and lead to physical failure of infrastructure, significantly increasing operational complexity and costs (Leblanc et al., 2017). More broadly, technical performance across decentralised sanitation systems in cold climates is closely linked to local environmental conditions, settlement patterns, and system design, particularly in peri-urban and low-density settings (Laukka et al., 2022).

Overall, the literature reveals a persistent gap: although the challenges of sanitation in cold climates are widely recognised, affordable, scalable, and field-validated solutions remain scarce. Existing knowledge of septic tank performance is largely derived from temperate or tropical regions, or from short-term laboratory studies that fail to capture the operational variability of real systems. Empirical evidence on the year-round performance of septic tanks under actual cold-climate conditions in LMICs remains extremely limited, where socio-economic and institutional barriers further exacerbate environmental stressors. Addressing this critical gap is central to the present research.

2.3 Stakeholder engagement in sanitation research

Stakeholder engagement has become a central focus in sanitation research, particularly in decentralised systems where responsibilities span government agencies, utilities, non-governmental organisations (NGOs), private contractors, and local communities (Herbst et al., 2009; Acey, 2019; WHO/UNICEF JMP, 2022a). Meaningful engagement is essential not only for identifying institutional and operational barriers but also for ensuring that proposed interventions are socially acceptable, contextually appropriate, and aligned with community needs (Acey, 2019; WHO/UNICEF JMP, 2022a).

Empirical studies demonstrate that structured engagement can strengthen governance,

improve coordination, and support evidence-based decision-making (Acey, 2019; WHO/UNICEF JMP, 2022a). The JMP pilot studies on safely managed on-site sanitation services highlighted the importance of early and continuous engagement through workshops, stakeholder mapping, and collaborative validation exercises (WHO/UNICEF JMP, 2022a). These activities improved institutional alignment and enhanced the comparability of monitoring data across countries such as Indonesia, Ecuador, and Serbia (WHO/UNICEF JMP, 2022a).

In Nigeria, Acey (2019) adopted a mixed-methods approach that combined surveys and interviews to capture both institutional and user perspectives. Interviews with utilities, government agencies, NGOs, landlords, and community-based organisations provided insights into system governance and operational challenges, while focus groups offered collective reflection on service experiences and coping strategies (Acey, 2019). Snowball sampling – a recruitment method in which initial participants recommend additional individuals – was used to identify informal and peripheral actors who might otherwise be absent from formal stakeholder lists, revealing how fragmented responsibilities and informal arrangements shaped service outcomes and opportunities for community participation (Acey, 2019).

Herbst et al. (2009) likewise employed a mixed-methods design integrating household surveys, semi-structured interviews, and focus groups. Surveys established a quantitative baseline for sanitation access and use, whereas interviews provided richer insights into individual experiences and behavioural drivers (Herbst et al., 2009). Focus groups provided a platform for dialogue, validation of findings, and triangulation across data sources, thereby strengthening both the reliability and interpretive depth of the study (Herbst et al., 2009).

Recognising limitations in traditional participatory tools, Davis et al. (2018) compared interviews, focus groups, and photovoice across twenty Indian communities with community-based sanitation systems. Interviews were the most effective in eliciting comprehensive sanitation priorities, including abstract dimensions such as dignity, affordability, and social risk (Davis et al., 2018). In contrast, focus groups and photovoice were sometimes constrained by local power dynamics and social hierarchies, which could inhibit open expression, particularly from women or marginalised participants (Davis

et al., 2018).

Collectively, these studies illustrate the value of methodological pluralism in stakeholder engagement. Structured tools such as surveys and interviews are effective for systematically capturing individual views but often overlook tacit knowledge or collective norms. Focus groups offer insight into community-level dynamics yet may reproduce existing social hierarchies. Visual or participatory approaches such as photovoice can amplify lived experience and enhance participant ownership but require careful facilitation to ensure that outputs are meaningfully translated into planning or policy. For sanitation research, the implication is clear: no single method is sufficient. Effective stakeholder engagement requires a deliberate combination of approaches that balance depth with breadth and prioritise inclusivity while maintaining analytical rigour.

2.4 Microbial processes in anaerobic systems

2.4.1 Anaerobic digestion and functional microbial groups

Anaerobic digestion is a multi-stage microbial process in which complex organic matter is progressively degraded into methane (CH_4), carbon dioxide (CO_2), and other metabolic by-products (Figure 2.2) (Luostarinen et al., 2007; McLeod et al., 2015). The process comprises four interdependent stages – hydrolysis, acidogenesis, acetogenesis, and methanogenesis – each mediated by specialised microbial guilds whose activity is strongly influenced by environmental conditions (Angelidaki and Ahring, 1993; Luostarinen et al., 2007; Almansa et al., 2023).

Hydrolysis is the initial and often rate-limiting step of AD, during which hydrolytic bacteria depolymerise complex organic substrates. Facultative and obligate anaerobes secrete extracellular enzymes, including proteases and lipases, that convert carbohydrates, fats, and proteins into soluble monomers such as sugars, fatty acids, and amino acids (Gerardi, 2003; Metcalf and Eddy, 2013; Zinder and Dworkin, 2013; Franke-Whittle et al., 2014). Hydrolytic–fermentative bacteria, primarily members of Firmicutes (e.g. *Clostridium* spp.) and Bacteroidetes, initiate this breakdown and supply substrates for downstream metabolic stages (Gerardi, 2003).

During acidogenesis (fermentation), fermentative bacteria convert hydrolysis products

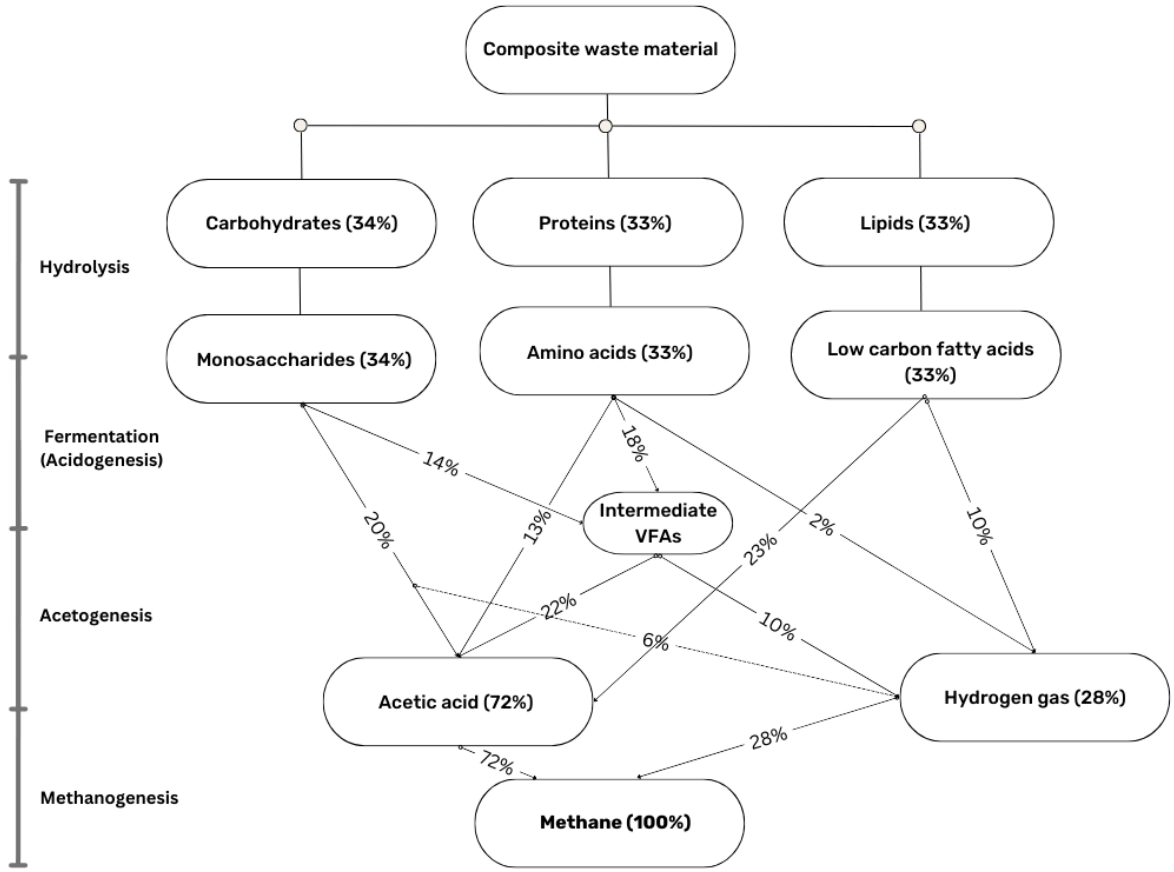

Figure 2.2: Conceptual schematic of anaerobic oxidation of composite waste material and key AD pathways, adapted from Metcalf and Eddy (2013). Percentages indicate typical literature-based distributions and may vary with waste composition and operating conditions.

into volatile fatty acids (VFAs), CO_2, hydrogen gas (H_2), and other reduced intermediates including ethanol and lactate (Ahring et al., 1995; Hori et al., 2006; Metcalf and Eddy, 2013; Franke-Whittle et al., 2014). Representative acidogenic taxa include *Lactobacillus*, *Enterobacter*, and *Clostridium* species (Gerardi, 2003).

Acetogenesis further transforms these intermediates under thermodynamically constrained conditions by regulating the flow of carbon toward acetate. In environments characterised by low-hydrogen partial pressures, thermodynamically constrained conditions, syntrophic acetogenic bacteria oxidise VFAs – particularly propionate ($CH_3CH_2COO^-$) and butyrate – into acetate (CH_3COO^-), CO_2, and H_2 (Metcalf and Eddy, 2013; Franke-Whittle et al., 2014). As emphasised by Conrad (2023), low temperatures can favour acetate formation (e.g. via homoacetogenesis), contributing to acetate accumulation and reinforcing acetate-centred methane formation under strong kinetic constraints, rather than implying obligatory routing through syntrophic acetate oxidation.

Syntrophic cooperation becomes critical where the removal of hydrogen is necessary to render otherwise energetically unfavourable reactions thermodynamically feasible. Under these conditions, continuous consumption of H_2 by hydrogenotrophic methanogens enables acetogenic VFA oxidation and, in specific stressed systems, may also support syntrophic acetate oxidation (SAO) (Stams and Plugge, 2009; Metcalf and Eddy, 2013). Accordingly, key syntrophic genera, such as *Syntrophomonas* and *Syntrophobacter* (family Syntrophaceae) form a critical metabolic link between VFA degradation and methane production. However, the dominant acetate conversion pathway – direct acetoclastic methanogenesis or SAO – remains system-dependent and cannot be attributed to temperature alone (Conrad, 2023).

Methanogenesis is the terminal stage of AD. Methanogenic archaea convert acetate, CO_2, and H_2 into biogas typically composed of approximately 60% CH_4, 38% CO_2, and trace hydrogen sulphide (H_2S) (Metcalf and Eddy, 2013; Franke-Whittle et al., 2014). Two principal pathways dominate: hydrogenotrophic methanogenesis and acetoclastic methanogenesis. Hydrogenotrophic methanogens, including members of the Methanobacteriales and Methanomicrobiales, reduce CO_2 using H_2 as an electron donor to produce CH_4 (Equation 2.1) (Metcalf and Eddy, 2013; Franke-Whittle et al., 2014):

$$4\mathrm{H}_2 + CO_2 \rightarrow CH_4 + 2H_2O \tag{2.1}$$

Acetoclastic methanogens, including members of the Methanotrichaceae (obligate acetoclasts) and Methanosarcinaceae (facultative acetoclasts), cleave acetate to produce CH_4 and CO_2 (Equation 2.2) (Metcalf and Eddy, 2013):

$$\mathrm{CH_3COOH} \rightarrow \mathrm{CH_4} + \mathrm{CO_2} \tag{2.2}$$

Beyond these principal routes, methanogens can also utilise one-carbon compounds via methylotrophic pathways. Formate ($HCOO^-$) is utilised by hydrogenotrophic methanogens (Equation 2.3); methanol (CH_3OH) is metabolised by methylotrophic Methanosarcinaceae (Equation 2.4); and methylamines (e.g. $(CH_3)_3N$) are degraded by members of Methanomassiliicoccales and Methanosarcinaceae (Equation 2.5) (Mara and Horan, 2008; Metcalf and Eddy, 2013):

$$4\mathrm{HCOO^-} + 4H^+ \rightarrow \ CH_4 + 3CO_2 + 2H_2O \tag{2.3}$$

$$4\mathrm{CH_3}OH \rightarrow 3CH_4 + CO_2 + 2H_2O \tag{2.4}$$

$$4(\mathrm{CH_3})_3N + 6H_2O \rightarrow \ 9CH_4 + 3CO_2 + 4NH_3 \tag{2.5}$$

Although these pathways are well characterised in engineered digesters, their relevance in cold-climate household septic tanks remains poorly understood. Several studies report limited or absent acetoclastic methanogens in cold-climate septic tanks, suggesting that methane formation may rely more heavily on hydrogenotrophic methanogenesis and/or system-dependent alternative acetate turnover routes (Shaw and Dorea, 2021). The episodic loading, low temperatures, and variable influent composition characteristic of household systems may further favour stress-tolerant and slow-growing taxa, although long-term, *in situ* field evidence remains scarce (Metcalf and Eddy, 2013; Shaw and Dorea, 2021).

All stages of AD are thermodynamically interconnected through interspecies electron transfer, such that disruption of any single stage can propagate through the system (Lee et al., 2009; Stams and Plugge, 2009). Under sub-optimal conditions, this can lead to inefficient fermentation, accumulation of intermediate products, odour generation, and excessive sludge build-up, ultimately compromising overall septic tank performance (Gray, 2004).

2.4.2 Syntrophic interactions and microbial competition

Syntrophic interactions are central to the stability of anaerobic digestion because they enable the degradation of substrates that are thermodynamically unfavourable under standard anaerobic conditions (Stams and Plugge, 2009; McKay et al., 2019). Syntrophy describes a cooperative metabolic relationship in which two or more microorganisms jointly catabolise a compound that neither partner can degrade in isolation, thereby supporting energy conservation in oxygen-depleted environments (Gerardi, 2003; Stams and Plugge, 2009). These interactions depend on interspecies electron transfer—primarily via H_2, formate, or acetate—which couples the oxidation of reduced intermediates to terminal processes such as methanogenesis or sulphate reduction (Stams and Plugge, 2009).

A key requirement for syntrophic metabolism is the maintenance of extremely low hydrogen partial pressure. Without rapid H_2 consumption by hydrogenotrophic methanogens, the oxidation of energy-poor substrates becomes thermodynamically unfavourable, particularly for VFAs such as propionate (Stams and Plugge, 2009; McKay et al., 2019; Dyksma et al., 2020). This dependency is illustrated by syntrophic propionate oxidation, in which propionate is converted to acetate, CO_2, and H_2 (Reaction 1 in Table 2.1) (Stams and Plugge, 2009; McKay et al., 2019; Dyksma et al., 2020). The H_2 released is immediately consumed by hydrogenotrophic methanogens, sustaining the low hydrogen partial pressure required for syntrophy (Reaction 2 in Table 2.1) (Stams and Plugge, 2009; McKay et al., 2019; Dyksma et al., 2020). Under stable mesophilic conditions, acetoclastic methanogens such as *Methanosaeta* and *Methanosarcina* directly cleave acetate to CH_4 and CO_2 (Reaction 3 in Table 2.1) (Gerardi, 2003).

Table 2.1: Energetics of syntrophic growth on propionate, adapted from Stams and Plugge (2009).

#	**Reactions**	$\Delta G^{\circ\prime *}$	$\Delta G^{\prime *}$
1.	$CH_3CH_2COO^- + 2H_2O \rightarrow CH_3COO^- + CO_2 + 3H_2$	72 kJ	-21 kJ
2.	$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O$	-131 kJ	-15 kJ
3.	$CH_3COO^- + H^+ \rightarrow CH_4 + CO_2$	-36 kJ	-36 kJ

* *$\Delta G^{\circ\prime}$ values are reported under biochemical standard conditions (pH 7), while $\Delta G^{\prime}$ reflects in situ conditions influenced by hydrogen partial pressure.*

Acetoclastic methanogenesis is kinetically sensitive to low temperatures and can be strongly inhibited by elevated ammonia concentrations (Guo et al., 2014). While low temperature alone typically constrains reaction rates rather than fully suppressing acetoclastic pathways, the presence of additional stressors, such as ammonia inhibition, high organic loading, or process instability, may shift acetate conversion towards syntrophic acetate-oxidising bacteria (SAOB) as an alternative metabolic route (Franke-Whittle et al., 2014; Dyksma et al., 2020). Under such conditions, SAOB oxidise acetate to CO_2 and H_2, with subsequent hydrogen consumption by partner methanogens required to maintain thermodynamic feasibility, as described in Equation 2.6 (Gerardi, 2003; Guo et al., 2014):

$$CH_3COO^- + H^+ + 2H_2O \rightarrow 2CO_2 + 4H_2 \tag{2.6}$$

This reaction is thermodynamically unfavourable under standard conditions and proceeds only when tightly coupled to hydrogenotrophic methanogenesis, which rapidly consumes H_2 and maintains the low hydrogen partial pressure required for syntrophic growth. Under such constrained conditions, methane production can proceed via the combined SAO–hydrogenotrophic pathway (Equation 2.7) (Gerardi, 2003; Guo et al., 2014):

$$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O \tag{2.7}$$

In addition to syntrophic acetate oxidation, the Wood–Ljungdahl (WL) pathway also

underpins *homoacetogenesis*, a competing acetate-centred process in which CO_2 is reduced to acetate using H_2 as the electron donor. Homoacetogenic bacteria utilise the same core WL pathway enzymes as syntrophic-oxidising bacteria but operate the pathway in the reductive direction, producing acetate rather than oxidising it (Karekar et al., 2022). This metabolic overlap means that commonly used WL marker genes (e.g. *fhs*, *fdh* and related C_1-metabolism genes) cannot discriminate between homoacetogenesis and syntrophic acetate oxidation when considered in isolation.

The performance of SAO and other syntrophic processes is influenced by temperature, pH, and organic loading rate (Stams and Plugge, 2009; Guo et al., 2014; Wu et al., 2016; Zhou et al., 2018). Disturbances to these interactions commonly result in VFA accumulation and reduced methane yields, reflecting breakdown of interspecies electron transfer (Guo et al., 2014; Zhou et al., 2018; Dyksma et al., 2020).

In addition to these syntrophic relationships, sulphur cycling introduces a distinct form of microbial competition within anaerobic systems. Sulphate-reducing bacteria (SRB) and methanogens utilise overlapping electron donors – primarily H_2, acetate, and formate – but sulphate reduction yields substantially more energy than methanogenesis (Stams and Plugge, 2009). Consequently, when sulphate is available, SRB (e.g. members of Desulfovibrionaceae, Desulfohalobiaceae, and Desulfonatronaceae) can outcompete both hydrogenotrophic and acetoclastic methanogens, diverting electrons away from methane formation and increasing H_2S production (Mara and Horan, 2008; Naphtali et al., 2022).

Table 2.2: Energetics of methanogenesis and sulphate reduction, adapted from Ozuolmez et al. (2015).

#	**Reactions***	$\mathbf{\Delta G_r^{\circ **}}$ **(kJ mol^{-1})**
1.	$CH_3COO^- + 4H_2O \rightarrow 4H_2 + 2HCO_3^- + H^+$	214.70
2.	$CH_3COO^- + H_2O \rightarrow CH_4 + HCO_3^-$	–14.74
3.	$4H_2 + SO_4^{2-} + H^+ \rightarrow HS^- + 4H_2O$	–262.06
4.	$CH_3COO^- + SO_4^{2-} \rightarrow HS^- + 2HCO_3^-$	–47.36
5.	$HCO_3^- + 4H_2 + H^+ \rightarrow CH_4 + 3H_2O$	–229.44

* *(1) Acetate oxidation; (2) acetoclastic methanogenesis; (3) hydrogenotrophic sulphate reduction; (4) acetate-oxidising sulphate reduction; (5) hydrogenotrophic methanogenesis.*

** ΔG_r° *values are standard Gibbs free energy changes calculated considering aqueous inorganic carbon and sulphur speciation.*

In domestic wastewater, sulphate concentrations are typically low, constraining SRB activity and allowing methanogenesis to dominate (Talavera-Cortés et al., 2025). However, SRB may still participate in VFA transformation, modulate redox conditions, and influence syntrophic pathways during periods of fluctuating loading or electron-acceptor availability (Mara and Horan, 2008; Naphtali et al., 2022). Thus, the balance between cooperative syntrophic metabolism (e.g., SAOB–methanogen interactions) and competitive processes (SRB–methanogen substrate rivalry) is governed by substrate availability, redox potential, and environmental stresses (McKay et al., 2019). In cold-climate sanitation systems such as septic tanks, preserving stable syntrophic interactions is critical for resilient organic matter degradation and sustained methane cycling (Sieber et al., 2012; Yun et al., 2023; Rama et al., 2024).

2.4.3 Functional gene markers relevant to this thesis

Functional marker genes provide a mechanistic link between microbial community composition and process-level performance by enabling inference of key anaerobic metabolic pathways operating under variable environmental conditions (Mei and Liu, 2022). Marker genes therefore offer a robust means of tracking functional responses to temperature stress and shifting conditions in cold-climate sanitation systems.

In this study, marker genes were selected to represent the principal anaerobic pathways governing acetate turnover, electron flow, and methane production (Mei and Liu, 2022). Acetate activation was represented by the phosphotransacetylase-acetate kinase pathway genes (*pta* and *ackA*), which catalyse the conversion of acetate to acetyl-CoA and constitute a central entry point into downstream anaerobic metabolism (Mei and Liu, 2022). Methanogenic potential was assessed using *mcrA*, encoding methyl-coenzyme M reductase, the terminal enzyme in methane formation across acetoclastic, hydrogenotrophic, and methylotrophic methanogens (Mei and Liu, 2022). Dissimilatory sulphate reduction was represented by *dsrA* and *dsrB*, encoding the alpha and beta subunits of dissimilatory sulphite reductase, a key competing anaerobic electron-accepting pathway that can influence hydrogen and acetate availability (Mei and Liu, 2022).

Genes associated with the WL pathway were included to capture acetate-centred C_1 metabolism potentially linked to syntrophic acetate oxidation (Conrad, 2023). Specifically, WL-associated markers (e.g. *fhs*, *fdh* and related CODH/ACS components) reflect the capacity for acetyl-CoA cleavage and C_1 carbon transformations (Conrad, 2023). However, as the WL pathway is reversible, these genes do not distinguish between homoacetogenesis and SAO, which operate in opposite metabolic directions and are governed by system-specific thermodynamic and environmental constraints. Accordingly, WL-associated genes are interpreted in this thesis as indicators of functional potential, rather than definitive evidence of active SAO, and are evaluated in conjunction with methanogenic and sulphate-reducing markers and environmental context.

2.4.4 Temperature sensitivity of functional guilds

Temperature is a primary control on microbial kinetics, methane production, and pollutant removal efficiency in anaerobic systems, exerting both direct effects on enzymatic activity and indirect effects through shifts in community structure (Metcalf and Eddy, 2013; Pussayanavin et al., 2015; Polprasert et al., 2018). Even modest temperature fluctuations of 1 °C to 2 °C can disrupt acetate conversion and promote VFA accumulation, while elevated temperatures may exacerbate ammonia-related inhibition, particularly under high organic loading (Appels et al., 2008; Velkushanova et al., 2021). Stable performance is most commonly associated with mesophilic conditions (30 °C to 40 °C), whereas operation at psychrophilic temperatures (< 15 °C) generally leads to reduced reaction rates unless compensated by extended hydraulic and solids retention times (Velkushanova et al., 2021).

Temperature strongly constraints the early stages of anaerobic digestion by regulating extracellular enzyme kinetics and rates of substrate conversion (Gerardi, 2003; Zinder and Dworkin, 2013; Nie et al., 2021). As temperature decreases, reaction rates decline due to reduced enzyme efficiency and slower transport across microbial membranes, resulting in diminished hydrolysis and acidogenesis (Veeken and Hamelers, 1999; Nie et al., 2021). Experimental evidence demonstrates that extracellular enzyme activity under thermophilic conditions (55 °C) can be nearly twice that observed under mesophilic operation (35 °C), underscoring the extent to which low temperatures in cold climates limit hydrolysis capacity (Veeken and Hamelers, 1999). Below approximately 20 °C,

both hydrolysis and acidogenesis become markedly slower, restricting the supply of intermediates required to sustain downstream acetogenic and methanogenic processes (Nie et al., 2021).

Temperature also exerts a strong control on methanogenic pathway expression by regulating reaction kinetics, growth rates, and substrate turnover. At lower temperatures, overall methanogenic rates decline sharply; however, acetate-based methanogenesis often mediated by acetoclastic methanogens is often favoured due to higher growth efficiencies and more rapid acetate conversion (Conrad, 2023). At psychrophilic temperatures, methane production may persist through the activity of cold-tolerant methanogens capable of maintaining metabolic function below 10 °C, although absolute reaction rates remain substantially reduced (Table 2.3) (Nie et al., 2021). Importantly, much of this current understanding of temperature-driven methanogenic shifts is derived from laboratory-scale reactors operated under controlled conditions. Evidence from household cold-climate septic tanks remains sparse and heterogeneous, with field studies reporting both the persistence of taxa typically associated with mesophilic systems and pronounced community instability, highlighting the need for *in situ* investigations under realistic operating conditions (Shaw and Dorea, 2021; Naphtali et al., 2022).

Table 2.3: Methanogens capable of growing in cold habitats, adapted from Nie et al. (2021).

Species	$T_{\min}$ (°C)	$T_{\max}$ (°C)	T_{opt} (°C)
Methanolobus psychrophilus	0	25	18
Methanobacterium sp. MB	5	30	28
Methanogenium frigidum	-2	28	23
Methanogenium boonei	5	30	19
Methanosarcina baltica	5	28	21
Methanosarcina lacustris	1	35	25
Methanococcoides alaskense	5	28	24–26
Methanococcoides burtonii	-2	28	23

Collectively, the literature demonstrates that temperature exerts a strong control over hydrolytic, fermentative, and methanogenic processes in anaerobic systems, yet it also reveals substantial uncertainty regarding how these temperature effects manifest

in cold-climate household septic tanks. The majority of existing studies are based on controlled laboratory reactors treating agricultural substrates or wastewater treatment plant (WWTP) digesters, which differ fundamentally from CSTs in influent composition, loading regimes, hydraulic and solids retention times, and exposure to seasonal temperature fluctuations (Mara and Horan, 2008; Fan et al., 2023). As a result, mechanistic insights derived from these systems may not be directly transferrable to decentralised sanitation contexts.

Field-based observations further complicate this picture. Reports of absent, reduced, or unstable methanogenic populations in CSTs (Viraraghavan and Dickenson, 1991; Pussayanavin et al., 2015; Shaw and Dorea, 2021) suggest that temperature alone may not fully explain observed microbial shifts. Instead, interacting factors, including influent characteristics, episodic loading, system configuration, and maintenance practices, are likely to modulate temperature effects and influence overall process stability. Long-term *in situ* monitoring across natural seasonal cycles is therefore essential to disentangle the relative contributions of temperature and system-specific drivers, and to clarify how microbial community dynamics ultimately govern septic tank performance in cold-climate settings.

2.4.5 Effects of pH and wastewater characteristics on treatment efficiency

pH is a critical parameter governing the stability of anaerobic reactors (Yenigün and Demirel, 2013). Methanogenic archaea operate optimally within a narrow pH range of 6.8–7.5; deviations toward acidity suppress methanogenesis and allow acidogenesis to dominate, leading to VFA accumulation and reduced methane production (Yenigün and Demirel, 2013; Franke-Whittle et al., 2014; Chen et al., 2023; Yang et al., 2024). Shifts outside this optimal range can also alter microbial competition. Under acidic (pH < 6.0) or alkaline (pH > 8.0) conditions, SRB gain a competitive advantage over pH-sensitive methanogens (Mara and Horan, 2008).

Physicochemical parameters such as total solids (TS), volatile solids (VS), COD, and biochemical oxygen demand (BOD) provide insight into the organic strength of wastewater and the extent of anaerobic degradation (Metcalf and Eddy, 2013). Elevated VS and TS concentrations indicate limited hydrolysis, excessive sludge accumulation, or inadequate desludging frequency, whereas high COD and BOD levels reflect incomplete degradation

and poor reactor performance (Veeken and Hamelers, 1999; Chen et al., 2023; Yang et al., 2024). Such conditions may arise from short HRTs, organic overloading, or reduced microbial activity under cold temperatures. Collectively, these factors influence both the operational stability and long-term treatment performance of septic tanks in cold-climate environments.

2.5 Pathogenic microorganisms in on-site sanitation systems

In addition to the functional microbial guilds responsible for anaerobic digestion, on-site sanitation systems such as septic tanks harbour a diverse range of microorganisms originating from human excreta, including bacteria with recognised pathogenic potential. These communities comprise taxa associated with enteric infections, waterborne transmission, and opportunistic environmental pathogens, reflecting the complex microbial inputs characteristic of domestic wastewater (Numberger et al., 2019; Wang et al., 2021). Studies of wastewater systems have demonstrated that microbial assemblages typically include members of families such as Enterobacteriaceae, Campylobacteraceae, Clostridiaceae, Aeromodanaceae, and Pseudomonadaceae, many of which are lined to human and animal health risks (Numberger et al., 2019; Wang et al., 2021).

Pathogenic microorganisms in wastewater are commonly classified according to their transmission pathways and ecological characteristics into three broad groups: (i) enteric bacteria, which originate from the gastrointestinal tract and are transmitted via the faecal–oral route; (ii) waterborne bacteria, which are capable of surviving and persisting in aquatic environments; and (iii) environmental or opportunistic pathogens, which are often associated with biofilms and may exhibit resistance to environmental stressors and antimicrobial agents (Numberger et al., 2019).

The persistence and survival of pathogenic microorganisms in on-site systems are governed by a combination of environmental and operational factors, including temperature, hydraulic retention time, substrate availability, and system configuration (Wang et al., 2021). Temperature, in particular, plays a critical role in regulating both microbial activity and pathogen survival. Although lower temperatures are generally associated with reduced metabolic rates, they do not necessarily result in effective pathogen inactivation (Wang et al., 2021). Evidence from wastewater treatment studies

indicates that certain pathogenic taxa may persist or even increase in relative abundance under specific conditions (Numberger et al., 2019; Wang et al., 2021). In addition, conventional on-site sanitation systems are primarily designed for the removal of solids, organic matter, and nutrients, rather than for the systematic removal of pathogens, resulting in variable and often limited pathogen attenuation (Akunna et al., 2017; Wang et al., 2021). In cold-climate septic tanks, reduced biological activity combined with limited treatment efficiency may therefore allow viable pathogenic populations to persist for extended periods.

The environmental implications of pathogen persistence are particularly significant in decentralised sanitation systems, where treated or partially treated effluent is frequently discharged to subsurface infiltration systems. In such settings, pathogens may be transported through soil and into groundwater, posing risks to drinking water supplies and public health (Chuah and Ziegler, 2018). Field-based studies have demonstrated that groundwater in areas relying on on-site sanitation systems is highly vulnerable to faecal contamination, with shallow wells frequently exhibiting persistent contamination and strong temporal variability linked to hydroclimatic conditions such as rainfall and water table fluctuations (Chuah and Ziegler, 2018).

Despite the recognised importance of pathogen removal for safeguarding environmental and human health, the dynamics of pathogenic microorganisms in on-site sanitation systems remain insufficiently understood, particularly under conditions of strong seasonal temperature variability. Existing literature indicates that the microbiological quality of effluents from on-site wastewater treatment systems and the factors influencing pathogen removal, including environmental conditions, remain insufficiently characterised (Wang et al., 2021). In particular, there is a lack of studies examining how temperature-driven changes in microbial community structure influence the persistence, relative abundance, and potential risks associated with pathogenic microorganisms in cold-climate sanitation systems.

This gap highlights the need for integrated investigations that link microbial ecology with environmental and public health outcomes, particularly in regions where decentralised sanitation systems operate under extreme climatic conditions.

2.6 Viral ecology and phage–host interactions in anaerobic systems

2.6.1 Phages in engineered anaerobic environments

Bacteriophages (phages), viruses that infect prokaryotes, are abundant in anaerobic digesters and wastewater treatment reactors, where they influence microbial population dynamics, community structure, and functional stability (Figure 2.3) (Tamaki et al., 2012; Dyksma et al., 2020). Viral communities in these systems are dominated by double-stranded deoxyribonucleic acid (DNA; dsDNA) phages belonging to the class Caudoviricetes, including members of the former families Siphoviridae, Myoviridae, and Podoviridae (Tamaki et al., 2012; Heyer et al., 2019). Collectively, these phages infect a wide range of anaerobic bacteria and methanogenic archaea (Tamaki et al., 2012; Heyer et al., 2019). For example, *Desulfovibrio* species are frequently targeted by Myoviridae, while Siphoviridae infect fermentative bacteria such as *Clostridium* spp. as well as methanogenic genera, including *Methanococcus* and *Methanobacterium* (Harirchi et al., 2022). Viral community composition has been shown to explain over 40% of the variance in prokaryotic community structure in anaerobic reactors (Zhang et al., 2017), highlighting the substantial ecological influence of phages and their potential effects on treatment performance.

Recent studies further emphasise the active regulatory role of phages in anaerobic digestion. Fan et al. (2023) showed that phage abundance often tracks host dynamics, suggesting that viruses act as responsive ecological participants rather than passive indicators of system disturbance. Their analysis revealed putative phage interactions with key functional guilds – including SRB, SAOB, and methanogens – and identified auxiliary metabolic genes (AMGs) associated with sulphur cycling, long-chain fatty acid degradation, and methanogenesis within putative viral genomes (Fan et al., 2023). These findings suggest that phages may influence both host population regulation and functional capacity in anaerobic systems.

Metagenomic surveys targeting the *mcrA* gene – encoding the α-subunit of methyl-coenzyme M reductase (MCR), the enzyme catalysing the terminal step of methane formation – have detected numerous viral operational taxonomic units (vOTUs)

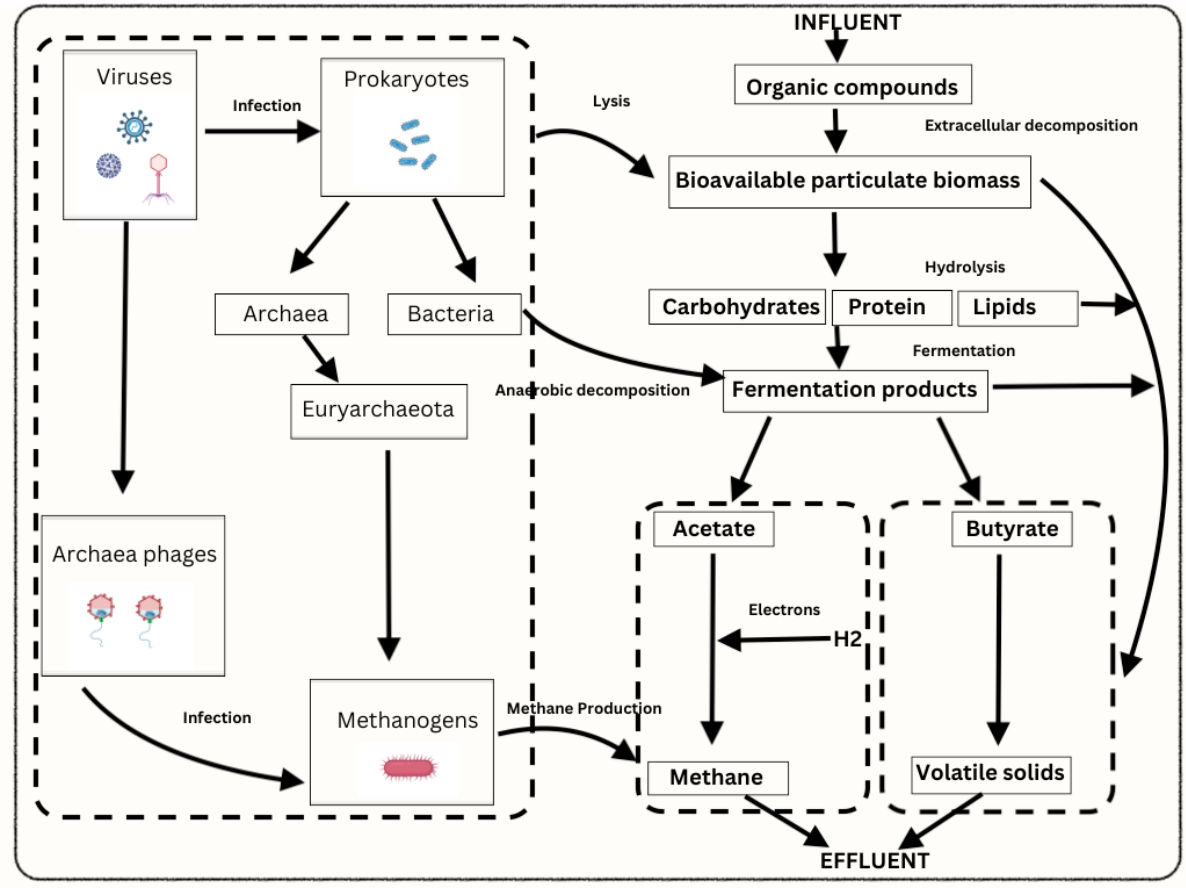

Figure 2.3: Dynamics of microorganisms and biochemical pathway analysis linking microorganisms to anaerobic reactor performance, adapted from Zhang et al. (2017).

associated with methylotrophic methanogens across diverse anoxic environments (Wang et al., 2022). Given that MCR is essential to all methanogenic pathways, viruses infecting methanogenic hosts may directly or indirectly modulate methane production and influence interactions between methane and sulphur cycling within anaerobic systems (Wang et al., 2022).

2.6.2 Ecological models of viral regulation

Phages regulate microbial populations through two primary infection pathways (Tamaki et al., 2012; Dyksma et al., 2020):

1. Lytic infection, in which host cells are lysed, releasing viral progeny and cellular constituents into the environment; and

2. Lysogeny, where phage DNA integrates into the host genome, enabling vertical transmission and potentially modifying host physiology and ecological function.

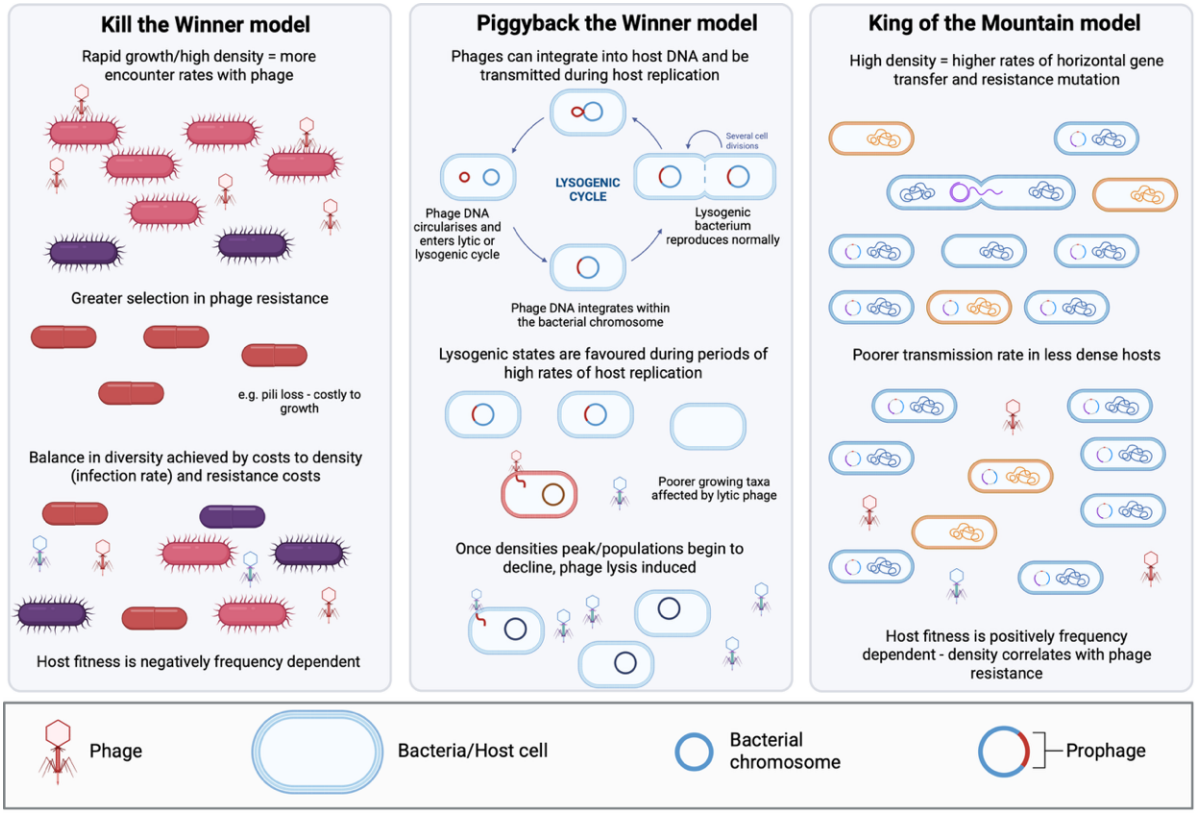

Figure 2.4: Ecological models explaining how bacteriophages regulate microbial growth, adapted from Castledine and Buckling (2024).

The ecological consequences of phage–host interactions are commonly interpreted through several conceptual models (Figure 2.4). The "Kill-the-Winner" (KtW) model proposes that phages preferentially infect dominant, fast-growing populations via lytic infection, thereby preventing competitive exclusion and promoting community diversity (Zhang et al., 2017; Fan et al., 2023). In contrast, the "Piggyback-the-Winner" (PtW) model suggests that under favourable growth conditions, temperate phages preferentially adopt a lysogenic lifestyle, integrating into host genomes and propagating alongside expanding host populations rather than suppressing them (Zhang et al., 2017; Fan et al., 2023; Castledine and Buckling, 2024). The "King-of-the-Mountain" (KoM) model further extends this framework by emphasising the capacity of dominant hosts to maintain ecological dominance through rapid resistance evolution and horizontal gene transfer of antiviral defence mechanisms, thereby reducing the regulatory impact of phages (Castledine and Buckling, 2024).

Collectively, these models highlight that viral influences in anaerobic systems are strongly context dependent. Depending on environmental conditions and microbial

life-history strategies, phages may function as top-down regulators through host lysis, as symbiotic partners enhancing host resilience via lysogeny, or as passive followers of host population dynamics (Zhang et al., 2017; Castledine and Buckling, 2024).

2.6.3 Knowledge gaps in decentralised sanitation systems

Although considerable insights have been gained from full-scale anaerobic digesters and WWTP bioreactors, phage ecology in decentralised sanitation systems such as septic tanks remains poorly characterised. Septic tanks differ markedly from engineered digesters: they operate at lower biomass densities, experience intermittent and highly variable loading, and are subject to large seasonal temperature fluctuations. These conditions can alter phage–host interactions and reduce the predictability of biological treatment processes.

Furthermore, while bioinformatic predictions suggest that phages may influence carbon and sulphur cycling (Zhang et al., 2017; Wang et al., 2022; Fan et al., 2023), the *in situ* activity of these viruses in field systems has yet to be experimentally validated. Methodological limitations – including the lack of standardised viral extraction protocols for sludge matrices and the underrepresentation of archaeal viruses – further constrain the interpretation of phage dynamics in septic tank environments (Quemin et al., 2016).

Overall, current evidence indicates that phages play dual and context-dependent roles in anaerobic digestion: destabilising microbial assemblages through lytic activity while potentially enhancing metabolic performance via lysogeny and auxiliary metabolic genes (Zhang et al., 2017; Wang et al., 2022; Fan et al., 2023; Castledine and Buckling, 2024). In cold-climate septic tanks – where low temperatures already constrain microbial growth – viral regulation of SRB, SAOB, and methanogens may critically influence both methane generation and effluent quality. However, no field studies have yet linked viral community dynamics to septic tank performance across natural seasonal cycles. Addressing this gap through integrated viromic and metagenomic analyses, as undertaken in this thesis, is essential for advancing the understanding and optimisation of anaerobic treatment processes in decentralised, cold climate sanitation systems.

2.7 Controlled laboratory-scale reactors as tools for understanding septic tank processes

Laboratory-scale septic tank reactors provide controlled environments for isolating the effects of temperature, HRT, organic loading rate, and microbial inoculum on anaerobic treatment performance. Under such controlled conditions, numerous studies have demonstrated that mesophilic operation (30 °C to 40 °C) enhances COD and BOD removal and stimulates biogas production, whereas psychrophilic operation (5 °C to 10 °C) leads to pronounced reductions in treatment efficiency (Viraraghavan and Dickenson, 1991; Koottatep et al., 2014a; Pussayanavin et al., 2015). These findings highlight the strong kinetic dependence of anaerobic processes on temperature under steady-state laboratory conditions.

Thermophilic laboratory configurations have also demonstrated high treatment efficiency. Koottatep et al. (2020a) reported that an up-flow septic tank reactor maintained at 50 °C and operated at a 23-hour HRT achieved high pollutant removal and supported the enrichment of thermotolerant methanogenic taxa, including *Methanosaeta* and *Methanothermobacter*. Complementary modelling work by Koottatep et al. (2014a) further suggested that operation at 60 °C could reduce *Escherichia coli* concentrations by up to six orders of magnitude, underscoring the sensitivity of pathogen inactivation to elevated temperatures in controlled reactor systems.

Beyond temperature manipulation, laboratory studies have explored alternative reactor configurations and operational strategies. Zheng et al. (2023) demonstrated that both increasing temperature and HRT in packed anaerobic baffled reactors significantly improved COD and NH_3-N removal, indicating that reactor design can modulate temperature effects on treatment performance. Similarly, McLeod et al. (2015) employed a factorial experimental design combined with multivariate statistical analysis to semi-continuous laboratory reactors and showed that HRT was the dominant control on solids removal, accounting for 79% of the variation in TS reduction and 59% of the variation in VS reduction. Importantly, their regression models closely reproduced performance trends observed in full-scale anaerobic digesters, suggesting that laboratory reactors can reliably capture key operational sensitivities (McLeod et al., 2015).

Overall, laboratory-scale reactors have proven indispensable for mechanistic investigations of anaerobic treatment processes, enabling systematic evaluation of temperature effects, reactor design, and operational controls. While such systems necessarily simplify real septic tank operation, they provide critical insights into process kinetics, microbial responses, and performance limits that would be difficult to resolve under uncontrolled conditions. However, the wide range of temperature variability characteristic of cold-climate environments has not yet been systematically explored in laboratory-scale septic tank reactors, limiting current understanding of process responses across realistic cold–warm temperature gradients.

2.8 Summary of knowledge gaps

The review identifies several key gaps that this thesis is designed to address:

1) Lack of long-term field evidence: Absence of year-round, field-based monitoring of septic tanks operating in cold-climate settings characterised by extreme seasonal temperature variability.

2) Limited mechanistic understanding: Insufficient understanding of how methanogenic and SAO pathways respond to freeze-thaw dynamics, constraining prediction of septic tank performance in cold climates.

3) Unexplored viral ecology: Absence of studies examining phage ecology in septic tanks and its seasonal interactions with bacterial and archaeal communities.

4) Fragmented stakeholder integration: Limited integration of stakeholder perspectives with physicochemical, technical, and biological evidence, constraining holistic understanding of practical sanitation challenges.

5) Restricted real-world applicability: Limited transferability of laboratory-based findings to household septic systems operating in cold-climate environments.

Collectively, these gaps justify the development of integrated, multidisciplinary methodological framework adopted in this thesis.

3 Materials and methods

3.1 Research design and approach

This study employed a mixed-methods research design integrating stakeholder perspectives, field-based household septic tank monitoring, laboratory-scale experiments, and metagenomic analyses to investigate how temperature variation influences the performance and microbiology of on-site sanitation systems in cold climates. The chapter outlines the ethical procedures, stakeholder engagement approach, field sampling strategy, reactor setup, physicochemical analyses, DNA extraction and sequencing workflow, and statistical methods used to evaluate microbial and viral dynamics across seasonal thermal regimes.

Following an overview of ethical approval (Section 3.2), the chapter describes the regional climatic setting that shapes the operation of sanitation systems in northern Kazakhstan (Section 3.3). Stakeholder engagement procedures are detailed in Section 3.4, followed by field sampling of domestic septic tanks in Section 3.5 and laboratory-scale reactor experiments designed to gain an understanding of the potential scale of the temperature effects on microbial performance under controlled conditions in Section 3.6. Physicochemical and molecular analyses are outlined in Sections 3.7 and 3.8, respectively, while statistical and comparative methods are detailed in Section 3.9. Section 3.10 describes the analytical framework applied to characterise phage–host relationships. Section 3.11 synthesises these components by describing how field observations, laboratory reactor experiments, and molecular analyses were integrated within a unified methodological framework. Together, these approaches provide a comprehensive basis for the year-round investigation of septic tank performance in cold-climate settings. Extracts from Sections 3.3-3.10 have been published or are currently in preparation for submission to peer-reviewed journals (see *Section 1.7*).

3.2 Ethical considerations

This study has obtained the approval from the Research Governance and Integrity Team (RGIT) and a favourable opinion from the Science, Engineering and Technology Research Ethics Committee of Imperial College London (SETREC reference #6417731). The approved protocol is provided in Appendix A.

All fieldwork procedures complied with Imperial College's Research Ethics Code and Fieldwork Health and Safety Regulations. Household owners participating in the field sampling provided informed consent for access to septic tanks and for anonymised publication of the results. No personally identifiable data were collected.

3.3 Regional context and study location

The research was undertaken in the Akmola region of northern Kazakhstan, an area characterised by pronounced annual thermal variability. Ambient temperatures range from approximately $-49\,^{\circ}\mathrm{C}$ in winter to $42\,^{\circ}\mathrm{C}$ in summer (Kazhydromet, 2023b). Mean annual precipitation exceeds 300 mm, and soils typically freeze from October to April, with frost penetration depths averaging 180 cm and reaching 274 cm in the coldest years (Severskiy, 2011; Baisholanov et al., 2017). These conditions impose substantial constraints on subsurface infrastructure and biological treatment processes, making the region a suitable setting for investigating temperature-dependent sanitation performance.

The soils of the Akmola region are dominated by ordinary and southern chernozems, classified as Mollisols under the United States Department of Agriculture Soil Taxonomy, and are characterised by predominantly heavy loam textures (Baisholanov et al., 2017; Weil and Brady, 2017). Although these soils can exhibit moderate permeability under unfrozen conditions, their high silt and clay content, combined with seasonal freezing, substantially reduce infiltration capacity during winter and early spring. These soil–climate interactions limit effluent exfiltration from subsurface drainage systems and impose additional seasonal constraints on on-site system performance in the region (Severskiy, 2011; Baisholanov et al., 2017; Weil and Brady, 2017; Ghangrekar, 2022).

Kazakhstan is classified by the World Bank as an upper-middle-income country (World Bank, 2025); however, marked socioeconomic disparities persist between major urban

centres and peri-urban and rural settlements. While cities such as Astana benefit from relatively high household incomes and substantial public investment, smaller towns and rural communities exhibit lower average incomes and greater reliance on self-managed infrastructure, including on-site sanitation systems (Bureau of National Statistics, 2025; WHO/UNICEF JMP, 2025).

The study encompassed three locations representing contrasting settlement typologies within the Akmola region: urban Astana, peri-urban Atbasar, and rural Zhibek Zholy. Astana, the national capital with a population of approximately 1.3 million, is widely recognised as the world's second-coldest capital city (Grjibovski et al., 2012; Bureau of National Statistics, 2020; Sinor et al., 2023). Atbasar, a peri-urban town with a population of approximately 34,000, holds Kazakhstan's national record for the lowest observed air temperature of $-52\,^{\circ}\mathrm{C}$ (Britannica, 2009; 2GIS Mapping System, 2023; Kazhydromet, 2023b). These two cities, located approximately 243 km apart, served as the sites for the stakeholder engagement component of the study, capturing diverse institutional and service delivery contexts.

Field-based microbial and physicochemical monitoring of wastewater collected from household septic tanks was conducted in Zhibek Zholy, a rural settlement located approximately 40 km west of Astana ($51^{\circ}04'09''$ N, $71^{\circ}46'38''$ E) (2GIS Mapping System, 2023). With an estimated population of 4,500 (Bureau of National Statistics, 2023), the settlement relies predominantly on on-site sanitation systems. Its climatic conditions, settlement layout, and reliance on septic tanks provided a representative field setting for examining temperature-driven variation in septic tank performance and microbial community structure.

3.4 Stakeholder engagement

3.4.1 Study design and participant recruitment

Stakeholder engagement was conducted in Astana and Atbasar, the two urban and peri-urban sites described in Section 3.3. These locations were selected to represent contrasting governance arrangements and service delivery contexts relevant to on-site sanitation systems in cold-climate regions.

A total of 30 stakeholders were recruited in 2023, representing key actors across the sanitation service chain, including national and municipal authorities, environmental regulators, utilities, NGOs, private desludging operators, political parties, professional associations, and households (Table B.1 in Appendix B). This multi-stakeholder design was adopted to capture perspectives from institutional, operational, and user levels and to enable comparison of roles, responsibilities, and constraints within the sanitation system.

A combined sampling strategy was applied. Institutional stakeholders were selected using purposive sampling to ensure inclusion of participants with direct knowledge and experience of sanitation governance, service provision, and regulation. Participants were identified through municipal contacts, professional networks, and NGO partnerships, allowing representation of diverse institutional perspectives. Households were selected using a random sampling approach to reduce selection bias and capture a broader range of user experiences. Five households in Astana and ten in Atbasar were identified using a random number generator and assigned unique identifiers to ensure traceability while maintaining anonymity. This approach enabled inclusion of diverse sanitation system types and user conditions across both study locations.

3.4.2 Representation and mitigation of sampling bias

While the sample was not intended to be statistically representative of all regions of Kazakhstan, it was designed to capture a diverse and analytically relevant range of stakeholder perspective across the sanitation service chain in cold-climate urban settings.

Several measures were taken to enhance representativeness and minimise potential bias. First, inclusion of multiple stakeholder groups ensured coverage of different roles within the sanitation system, including policy, service provision, and end users. Second, recruitment across two cities with differing infrastructural and governance characteristics enabled inclusion of varied operational contexts. Third, the use of random sampling for households reduced the likelihood of selection bias and supported the inclusion of diverse system types, installation conditions, and user experiences.

Together, these measures provided robust coverage of key stakeholder perspectives and sanitation system conditions relevant to the study objectives.

3.4.3 Data collection

Participants were offered the option to complete the survey independently or to participate in a structured face-to-face interview. This flexible approach was adopted to maximise participation across stakeholders with differing institutional roles, time availability, and preferences.

3.4.4 Survey instrument

The survey instrument included a combination of closed and open-ended questions designed to capture both quantitative responses and qualitative insights. The questionnaire addressed the following themes:

- system design and installation practices
- desludging and maintenance practices
- winter accessibility and frost-related failures
- user experiences of odour, backflow, overflow, and effluent management
- institutional responsibilities, regulatory gaps, and perceived solutions

The full questionnaire is provided in Appendix B (Figure B.1).

3.4.5 Data processing and analysis

All responses were anonymised and, where necessary, translated prior to analysis. Qualitative responses were analysed using thematic coding, with findings organised by stakeholder group to enable comparison of perceptions, responsibilities, and constraints across the sanitation service chain.

This analytical approach allowed identification of both shared and stakeholder-specific challenges, reflecting the multi-level nature of sanitation system performance in cold-climate environments.

3.5 Field sampling of household septic tanks

3.5.1 Household selection and system characteristics

Ten households in Zhibek Zholy, Kazakhstan, were selected for monthly sampling over the period January–December 2024. Households were included based on the following eligibility criteria:

1. Exclusive reliance on an on-site septic tank for the management of domestic wastewater;

2. Year-round physical accessibility, including during winter months under conditions of seasonal snow cover and frozen ground; and

3. Willingness of householders to permit repeated site access for monitoring and sampling activities throughout the full study period.

The selected systems were representative of typical rural on-site sanitation practices in northern Kazakhstan. The systems were selected to minimise structural and operational heterogeneity, thereby enabling seasonal effects to be examined within a common system type rather than across fundamentally different configurations. All septic tanks received combined household wastewater, comprising both blackwater (toilet-derived wastewater containing human excreta) and greywater generated from routine domestic activities (Velkushanova et al., 2021). Greywater inputs typically originated from kitchen sinks, bathroom hand basins, showers, and laundry appliances. Toilets were conventional gravity-flush systems, with no evidence of low-flush or water-saving designs, and showers were standard domestic installations rather than water-conserving units. Dishwashers were uncommon, whereas washing machines were present in all ten households.

Septic tanks were predominantly single- or two-chamber systems, constructed *in situ* using reinforced concrete rings or prefabricated concrete units. Tank geometries were primarily cylindrical, although a small number exhibited rectangular concrete designs, reflecting household construction practices rather than standardised engineering specifications.

Septic tanks were generally buried at depths of approximately 2.5 m below ground level,

with burial depth constrained by regional frost penetration, groundwater conditions, and spatial limitations of the household plot. In most cases, the upper portion of the tanks was located within the seasonally frozen soil layer, while the lower section remained below the maximum frost depth, as inferred from measured burial depths and direct site observations conducted during winter sampling campaigns. Effluent from the tanks was discharged to simple subsurface infiltration systems (drainage fields), typically consisting of gravel-filled trenches.

Key septic tank characteristics were recorded to support contextual interpretation of physicochemical and biological data and are summarised in Table 3.1.

Table 3.1: Characteristics of study households and septic tank systems.

Parameters/ Households	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Mean**
Volume of a septic tank (m^3)	6	4	4	4	6	4	4	4	4	4	4.4
Emptying frequency of a septic tank (times per year)	1	3	3	3	2	4	3	3	2	1	2.5
Number of occupants in a household	5	4	4	4	5	5	4	4	4	4	4.3

3.5.2 Sampling schedule and field measurements

Sampling was conducted monthly from January to December 2024. To ensure temporal consistency and comparability between campaigns, all sampling events were carried out on the second Monday of each month. For analytical purposes, seasons were defined as winter (December–February), spring (March–May), summer (June–August), and autumn (September–November), reflecting regional climatic conditions.

At each site, sampling and field measurements followed a standardised protocol:

- Prior to sampling, informed consent was obtained from the homeowner to access the property and conduct septic tank sampling, with all activities carried out in

accordance with agreed safety and privacy considerations.

- A 2 L grab sample was collected from the first chamber of the septic tank using a stainless-steel core sampler (Figure 3.1).
- Samples were collected at a standardised depth of approximately 1.5 m below the liquid surface to ensure consistency across systems and minimise disturbance of settled solids.
- Ambient air temperature and wastewater temperature were recorded *in situ* at the time of sampling using a calibrated digital thermometer.
- Site conditions were documented, including odour presence, snow/ice cover, visible signs of overflow or surface pooling, and any access constraints encountered during sampling.
- Upon completion of sampling, the septic tank access lid was securely reseated and sealed to its original condition to maintain system integrity and household safety.

3.5.3 Sample preservation and transport

Immediately following collection, samples were placed in insulated containers with ice packs and transported to the Astana WWTP laboratory. All samples were processed within 4 hours of collection to minimise physicochemical and biological alterations during transport.

For molecular analyses, subsamples were aliquoted into sterile 50 mL centrifuge tubes under aseptic conditions and stored at $-80\,^{\circ}\mathrm{C}$ until DNA extraction (Figure 3.2). This preservation protocol was adopted to ensure long-term stability of nucleic acids to maintain comparability across sampling events and seasons.

Figure 3.1: A stainless-steel core sampler.

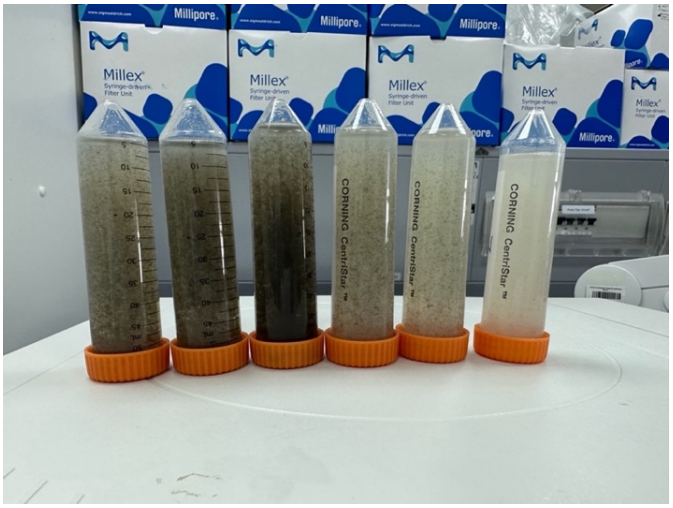

Figure 3.2: 50 mL wastewater aliquots.

Figure 3.3: Laboratory-scale reactors filled with inoculum.

3.6 Laboratory-scale reactor experiments

3.6.1 Experimental setup

Laboratory-scale anaerobic reactors were constructed following the design described by McLeod et al. (2015). Each reactor consisted of a laboratory-grade polyethylene bottle (ISOLAB, Germany) with a total working capacity exceeding 2 L (Figure 3.3). Reactors were configured to minimise variability in operational conditions and to isolate temperature as the primary factor influencing anaerobic performance.

Four temperature regimes were investigated to represent the range of conditions relevant to cold-climate septic tank operation:

- 4 °C (psychrophilic)
- 12 °C (low-temperature transitional range)
- 24 °C (mesophilic)
- 36 °C (upper mesophilic to thermotolerant range relevant to septic systems)

Each temperature condition was operated in triplicate, and the entire experimental sequence was repeated across three independent cycles to ensure reproducibility and to assess the consistency of observed trends.

3.6.2 Operation and sampling

Reactors were inoculated with homogenised sludge collected from the anaerobic chamber of the Astana WWTP. This inoculum was selected because anaerobic digester sludge contains a diverse and well-established microbial consortium, including fermentative bacteria, SAOB, SRB, and methanogenic archaea (Luostarinen et al., 2007; Mara and Horan, 2008; Bowen et al., 2014). These functional groups mediate the core anaerobic metabolic pathways responsible for organic matter degradation and methane production in septic tank systems, notwithstanding differences in scale and hydraulic and thermal variability between septic tank systems and municipal anaerobic digesters (Luostarinen et al., 2007; Mara and Horan, 2008; Bowen et al., 2014). Anaerobic digester sludge is therefore widely regarded as a functionally representative and appropriate inoculum for laboratory studies investigating decentralised sanitation processes and the effects of temperature on anaerobic digestion (Gerardi, 2003; Luostarinen et al., 2007; Naphtali et al., 2022).

Following inoculation, reactors were operated in batch mode. Reactor headspace was vented every three days to prevent excessive pressure accumulation and to maintain safe operating conditions, in line with established practice for anaerobic laboratory systems (Koottatep et al., 2014a). At two-day intervals, 1 L of reactor contents was withdrawn for physicochemical analysis. Withdrawn volumes were not replaced, allowing progressive assessment of anaerobic degradation under controlled temperature conditions without introducing additional substrate variability.

All sampling procedures were conducted using sterile techniques to minimise contamination. Reactors were gently mixed manually prior to sampling to ensure representative subsamples were obtained, while avoiding disturbance of settled biomass.

3.7 Physicochemical analyses

All physicochemical analyses were conducted at the Astana WWTP laboratory in accordance with the American Public Health Association (APHA) Standard Methods for the Examination of Water and Wastewater (American Public Health Association, 2017). The parameters analysed included pH, total solids, volatile solids, chemical oxygen demand, 5-day biochemical oxygen demand (BOD_5), and volatile fatty acids. VFAs were

analysed during an initial calibration phase; however, routine VFA measurements were discontinued following gas chromatography (GC) column malfunction.

Duplicate samples and analytical blanks ensured quality control; batches were re-run if quality control criteria were not met. Analytical methods and instruments are detailed in Table 3.2, and definitions of the measured parameters are provided in Appendix C (C.1).

Table 3.2: Analytical methods and instruments used for physicochemical analysis (American Public Health Association, 2017).

Parameters	**Method**	**Instruments**
pH	Electrometric method	pH meter
TS	Gravimetric method	Drying oven, muffle furnace
VS	Gravimetric method	Muffle furnace
COD	Closed reflux, colorimetric method	Spectrophotometer
BOD_5	5-day BOD Test	BOD incubator
VFA	GC with flame ionisation detection	TRACE 1300 GC (Thermo Fisher Scientific, Waltham, MA, USA)

3.8 Molecular and metagenomic analyses of household septic tank samples

3.8.1 Household septic tank sample concentration and DNA extraction

DNA extraction was conducted at the Laboratory of Human Microbiome and Longevity, National Laboratory Astana, Nazarbayev University. Microbial and viral particles were concentrated using a pellet-based direct extraction approach (Solids Method), as described by Jiang et al. (2024). This approach was selected for its demonstrated efficiency in recovering both prokaryotic and viral fractions from relatively small wastewater volumes, making it well suited for low-biomass environmental samples.

For each sample, a 50 mL aliquot of wastewater was centrifuged at 13,000 ×g for 10 minutes, and the resulting pellet was retained for DNA extraction and downstream molecular analysis. This concentration step minimised biomass loss and enhanced recovery

of viral and microbial nucleic acids, which is critical for shotgun metagenomic sequencing of septic tank wastewater (Qiagen, 2018; Jiang et al., 2024).

Table 3.3: Key laboratory equipment, kits, and software used for wastewater DNA extraction, sequencing, and analysis of household septic tank samples.

Method	**Materials**
Sample preparation	Centrifuge (Eppendorf 5430 R, Germany)
DNA extraction	Qiagen AllPrep PowerViral DNA/RNA Kit (Qiagen, CA, USA); Microcentrifuge (Hanil M15S, South Korea); TissueLyser III (Qiagen, CA, USA); Vortex mixer (IKA® Vortex Genius 3, IKA-Werke GmbH & Co. KG, Germany)
DNA quality control	Thermo Scientific NanoDrop™ 8000 spectrophotometer
Shotgun metagenomic sequencing	Illumina NovaSeq platform with paired-end 150 bp sequencing
Bioinformatic processing	Custom pipeline implemented in Python and R

DNA was extracted using the Qiagen AllPrep PowerViral DNA/RNA Kit (Cat No. 28000-50; Qiagen, CA, United States). This kit incorporates Inhibitor Removal Technology, which is particularly effective at removing humic substances, polysaccharides, and other inhibitory substances commonly present in domestic wastewater (Qiagen, 2018; Jiang et al., 2024).

Each concentrated 50 mL sample was processed in accordance with the manufacturer's protocol, with minor modifications (Figure C.6 in Appendix C). For each extraction, 200 μL of sample was combined with 600 μL of Solution PM1 (pre-warmed to 55 °C) in a PowerBead Tube. Mechanical homogenisation was performed using a Qiagen TissueLyser III at 25 Hz for 10 minutes to ensure effective disruption of both prokaryotic and viral particles (Qiagen, 2018). This step replaced vortex mixing with a vortex adapter, which was unavailable during laboratory processing (see Table 3.3).

Following lysis, Solution IRS was added to remove residual inhibitors. Nucleic acids were subsequently bound to a silica-based MB Spin Column using Solution PM3 containing high concentrations of binding salts. The column was sequentially washed with Solution PM4 (100% ethanol) and Solution PM5 (isopropanol) to remove residual

contaminants and salts. Purified DNA was eluted in 50 μL of RNase-free water applied directly to the membrane centre, yielding material suitable for shotgun metagenomic sequencing and downstream analyses.

3.8.2 DNA quantification and quality control

The concentration and purity of dsDNA extracted from household septic tank samples were assessed using a Thermo Scientific NanoDrop^{TM} 8000 spectrophotometer (Table 3.2). This high-throughput instrument enabled simultaneous measurement of up to eight samples, ensuring analytical efficiency and consistency across the dataset.

DNA concentrations were quantified in ng µL^{-1}, while A_{260}/A_{280} absorbance ratios were used as the indicators of nucleic acid purity (Thermo Scientific, 2008). Ratios between 1.8 and 2.0 were considered indicative of high-quality DNA with minimal contamination from proteins, phenolic compounds, or other organic inhibitors (Thermo Scientific, 2008). In addition, full absorbance spectra were visually inspected for each sample, with particular attention to the height and shape of the characteristic 260 nm peak associated with pure nucleic acids (Thermo Scientific, 2008).

Only samples meeting all quality thresholds – DNA concentration $\geq$ 20 ng µL^{-1}, $A_{260/}A_{280}$ ratios within the range 1.8–2.0, and appropriate spectral profiles – were advanced to shotgun metagenomic sequencing. Samples failing to meet these criteria were re-analysed and, where necessary, re-extracted. This stringent quality control ensured high analytical reliability and minimised downstream sequencing artefacts.

3.8.3 Shotgun metagenomic sequencing

High-quality DNA samples were sequenced by HaploX Biotechnology (Hong Kong) using the Illumina NovaSeq platform with a paired end 150 base pair (PE150) sequencing strategy. Sequencing yielded approximately 6 Gb of data per sample, with Q30 scores exceeding 95%, indicating that over 95% of bases were sequenced with an error probability below 0.1%, consistent with Illumina high-accuracy performance standards (HaploX, 2024).

Shotgun metagenomics was selected in preference to amplicon-based (e.g.,16S rRNA gene) or culture-dependent methods due to its ability to provide an unbiased and

comprehensive characterisation of microbial communities. While 16S rRNA sequencing enables broad bacterial and archaeal taxonomic profiling, it is limited by primer bias and does not capture viral diversity or functional gene content (Durazzi et al., 2021). In contrast, shotgun metagenomics enables simultaneous detection of bacteria, archaea, and viruses, provides higher taxonomic resolution, and allows direct interrogation of functional genes (Durazzi et al., 2021; HaploX, 2024). This approach is therefore particularly suited to linking microbial structure with metabolic potential, including methanogenesis, syntrophic pathways, phage–host interactions, and microbial responses to temperature variation (Durazzi et al., 2021).

3.8.4 Bioinformatic workflow

Raw sequencing reads were quality-filtered, screened to remove host-derived sequences, and assembled into contigs – derived as continuous DNA sequences reconstructed from overlapping short reads (Istrail et al., 2016) – by HaploX Biotechnology using established assembly pipelines. Contigs shorter than 1,000 bp were excluded to reduce assembly artefacts and improve the reliability of downstream taxonomic and functional annotation (HaploX, 2024).

Taxonomic annotation of bacterial, archaeal, and viral sequences was performed using the National Centre for Biotechnology Information (NCBI) Reference Sequence (RefSeq) database (HaploX, 2024). Functional annotation was conducted using Kyoto Encyclopaedia of Genes and Genomes (KEGG) Orthology (KO) identifiers (HaploX, 2024). Marker genes were selected to represent key anaerobic metabolic pathways relevant to this study, including acetate activation (*ackA, pta*), syntrophic acetate oxidation inferred from Wood-Ljungdahl pathway genes, methanogenesis (*mcrA)*, and dissimilatory sulphate reduction (*dsrA, dsrB*). Annotated outputs were subsequently integrated into downstream statistical and multivariate analyses.

3.9 Statistical analysis

All statistical and multivariate analyses were conducted in R (version 4.5.1), with custom scripts provided in Appendix E to ensure reproducibility. Given the compositional nature of shotgun metagenomic data, relative abundance matrices were transformed using centred

log-ratio (CLR) transformation, and dissimilarities were quantified using Aitchison distances, calculated as Euclidean distances in CLR-transformed space (Subharuppguhaa, 2022). This approach mitigates compositional bias and enables valid comparisons of microbial and viral community structures across temperature regimes and sampling periods (Subharuppguhaa, 2022).

3.9.1 Handling of uneven dataset structure

The dataset was unbalanced due to variability in DNA yield and purity across field samples, which resulted in incomplete metagenomic coverage for some septic tanks at certain sampling time points. Consequently, not all systems were represented consistently across the full annual cycle. To accommodate missing observations and unequal group sizes, analyses were restricted to statistical and multivariate methods robust to unbalanced designs. These included distance-based ordination techniques – Principal Coordinates Analysis (PCoA) and non-metric multidimensional scaling (NMDS) – as well as permutation-based hypothesis testing using permutational multivariate analysis of variance (PERMANOVA) (Subharuppguhaa, 2022). Relative abundance-based community profiling was also employed to support descriptive interpretation of seasonal patterns (Subharuppguhaa, 2022).

3.9.2 Assessment of sample similarity and treatment of replicate samples

To assess compositional similarity among samples collected from the same septic tank within the sampling month, Bray–Curtis dissimilarity was calculated based on relative abundance profiles. This analysis was performed prior to multivariate and diversity analyses to define appropriate analytical units and minimise pseudo-replication. Pairwise Bray–Curtis distances were calculated for all within-tank, within-month sample pairs. Samples were considered compositionally distinct when Bray–Curtis dissimilarity exceeded a threshold of 0.30, indicating substantial differences in relative abundance structure (Subharuppguhaa, 2022), whereas samples with dissimilarity values below this threshold were considered compositionally similar.

Where replicate samples exhibited Bray–Curtis dissimilarity values below 0.30, their relative abundance profiles were averaged to generate a single representative profile for that tank and month. Replicate samples exceeding this threshold were retained as

independent observations and treated as separate samples in downstream analyses. All retained and merged samples are documented in Table C.1 (Appendix C).

This procedure preserved biologically meaningful within-tank variability while avoiding artificial inflation of sample size arising from compositionally redundant replicates.

3.9.3 Multivariate comparisons

PCoA based on Aitchison distance matrices was used to assess multivariate dissimilarities in microbial and viral community composition across temperature gradients and seasonal categories (Istrail et al., 2016; Quince et al., 2017; Subharuppguhaa, 2022). This ordination approach facilitated visual evaluation of clustering patterns associated with seasonal variation and environmental conditions. In parallel, NMDS was applied to the same distance matrices as a complementary, rank-based ordination method. NMDS provided a robust representation of ecological structure under non-linear relationships, zero-inflation, and high-dimensional data typical of metagenomic datasets, thereby supporting interpretation of community restructuring across seasons and temperature regimes (Istrail et al., 2016; Quince et al., 2017; Subharuppguhaa, 2022).

3.9.4 Diversity metrics

Microbial and viral community diversity was quantified using complementary within-sample (α-diversity) and between-sample (β-diversity) metrics. Alpha-diversity was assessed using Shannon diversity, which integrates richness and evenness, and inverse Simpson diversity, which emphasises dominance patterns and effective diversity (Subharuppguhaa, 2022; Sam et al., 2025). These indices were used to characterise within-sample community structure and to evaluate seasonal stability in microbial and viral assemblages (Subharuppguhaa, 2022; Fan et al., 2023).

Beta-diversity metrics were used to quantify compositional dissimilarities between samples, enabling assessment of temperature- and season-driven restructuring of microbial and viral communities (Subharuppguhaa, 2022; Fan et al., 2023). Together, these complementary measures supported interpretation of how seasonal temperature variability influences both within-system diversity and broader shifts in community composition relevant to septic tank performance and management.

3.9.5 Community-level differences

Community-level differences in composition were evaluated using PERMANOVA to test whether observed variation in β-diversity was significantly associated with seasonal temperature regimes beyond that expected by random permutation (Subharuppguhaa, 2022). Samples were aggregated into four seasonal categories (winter, spring, summer, autumn) to reduce short-term temporal variability and to enhance statistical power. Permutations were constrained to septic tank identity to account for repeated sampling within individual systems.

3.9.6 Statistical significance and multiple-testing correction

Statistical significance was evaluated using conventional thresholds ($p \leq 0.05$, $p \leq 0.01$, and $p \leq 0.001$), reflecting increasing confidence that observed patterns did not arise by chance (Subharuppguhaa, 2022). To account for multiple hypothesis testing across taxa, functional genes, and environmental variables, p-values were adjusted using the Benjamini-Hochberg False Discovery Rate (FDR) correction (Istrail et al., 2016; Subharuppguhaa, 2022). Only associations with FDR-adjusted $p - values(q - values) \leq 0.05$ were considered statistically significant and interpreted in the context of microbial and viral community dynamics.

In graphical outputs, statistical significance was denoted using asterisks: $q \leq 0.05$ (*), $q \leq 0.01$ (**), and $q \leq 0.001$ (***). This convention ensured that only robust and reproducible relationships were highlighted in the interpretation of ecological patterns and correlations.

3.9.7 Univariate statistical analysis

Differences in microbial relative abundance across temperature conditions (4 °C, 12 °C, 24 °C, and 36 °C) were assessed using the Kruskal–Wallis rank-sum test due to the non-normal distribution of the data, the presence of outliers, and heterogeneity of variances among groups. Data are reported as medians and interquartile ranges (IQRs), representing the spread between the 25th and 75th percentiles. Where significant overall differences were detected, pairwise comparisons were conducted using the Wilcoxon rank-sum test. To account for multiple comparisons, p-values were adjusted using the

Benjamini-Hochberg FDR procedure, with statistical significance defined at $q \leq 0.05$.

3.9.8 Correlation and environmental association analysis

Ecological relationships among microbial functional guilds, viral families, and physicochemical parameters were examined using pairwise Spearman rank correlation analysis. This non-parametric approach is appropriate for non-normally distributed, zero-inflated datasets and is robust to missing observations (Subharuppguhaa, 2022). Correlations were calculated between the relative abundances of key microbial guilds (methanogens, SAOB, SRB, and fermenters), viral families, and physicochemical variables including wastewater temperature, pH, COD, BOD_5, TS, and VS. Statistical significance was assessed using FDR correction, and only correlations with $q \leq 0.05$ were retained for interpretation.

Environmental drivers of community composition of laboratory-scale reactors were evaluated using environmental fitting (ENVFIT) implemented in the *vegan* package in R. Environmental variables were fitted onto ordination space derived from Aitchison distance, with statistical significance assessed using 999 permutations. The strength of association was expressed as the squared correlation coefficient (R^2).

Correlations were visualised as correlation heatmaps, with colour gradients indicating both the direction and strength of associations, where red denotes positive correlations and blue denotes negative correlations.

3.10 Phage–host association analysis of household septic tank samples

Phage–host associations were inferred from viral contigs using Integrated Phage–Host Prediction tool (iPHoP) version 1.0.0, a genome-based framework that integrates multiple complementary signals to enable robust host assignment (Roux, 2022). The iPHoP algorithm combines sequence similarity to reference genomes, shared protein domain profiles, and genomic feature-based metrics to probabilistically link viral contigs to prokaryotic hosts (Roux, 2022).

Viral contigs derived from metagenomic assemblies and quality-filtered using VirSorter2 and CheckV (Subharuppguhaa, 2022; Fan et al., 2023)(scripts provided in

Appendix E), were provided as input to iPHoP. Host prediction was performed using default parameters, and only predictions exceeding the recommended confidence threshold of 90% were retained for downstream analysis (Roux, 2022). Host assignments were resolved at genus and species levels where possible. Predicted phage–host associations were subsequently grouped according to functional microbial guilds relevant to anaerobic digestion, including fermentative bacteria, SRB, SAOB, and methanogenic archaea.

Seasonal phage–host association analysis was restricted to Septic Tank 2 due to the substantial computational demands of genome-based host prediction. The iPHoP framework requires large reference databases (approximately 315 Gb) and integrates multiple programming languages and machine-learning components (Roux, 2022), resulting in high memory and processing requirements. Septic Tank 2 was therefore selected to allow consistent application of the complete iPHoP pipeline across all four seasons. In addition, the household characteristics of this system – including tank volume, desludging frequency, and number of occupants – were close to the mean values observed across the ten monitored systems (Table 3.1), supporting its use as a representative case study.

All scripts used for phage–host association and downstream analyses are provided in Appendix E. To facilitate comparison of predicted host composition across seasons and functional guilds, phage–host associations were normalised within each season and visualised using 100% stacked bar plots, representing the relative distribution of predicted host families within each functional guild and seasonal category.

3.11 Integration of field, laboratory, and molecular approaches

This study combined field-scale monitoring, laboratory-controlled experiments, molecular characterisation, and statistical analysis to address the research objectives in a complementary manner. Field data captured real-world system behaviour under natural temperature fluctuations, while laboratory reactors enabled isolation of temperature effects under controlled conditions. Molecular and metagenomic analyses provided mechanistic insight into the microbial and viral processes underpinning observed physicochemical trends. Together, these approaches enabled robust interpretation of seasonal performance dynamics in cold-climate septic tanks.

4 Results: Stakeholder survey

Extracts from this chapter have been published (see Paper 1 in the Section 1.7).

4.1 Overview of survey participants

The stakeholder survey comprised thirty participants representing four key groups involved in sanitation service provision and use: government officials, NGOs, private desludging operators, and households.

Findings are presented thematically, with emphasis on differences across stakeholder groups reflecting their distinct roles within the sanitation service chain. This grouping enabled a structured comparison of reported experiences, operational responsibilities, and constraints across the sanitation system.

4.2 Stakeholder perspectives on sanitation systems

4.2.1 Government stakeholders

Government stakeholders demonstrated a high level of awareness of sanitation challenges associated with cold-climate conditions. However, responses indicated the presence of systemic constraints affecting implementation capacity across the relevant institutions, including national ministries, municipal authorities, and water utilities. These constraints were primarily associated with limited financial resources and competing policy priorities, and challenges in the consistent enforcement of construction and maintenance standards for on-site sanitation systems.

Stakeholders further identified gaps in environmental monitoring and inter-institutional coordination. A notable proportion of respondents highlighted difficulties in aligning responsibilities and actions across agencies, reflecting the fragmented nature of governance arrangements within the sanitation sector. These limitations were observed to contribute to persistent operational issues, including the improper installation of septic tanks (e.g., above the frost penetration depth), inadequate

maintenance practices, and limited oversight of potential impacts on groundwater quality.

Respondents also emphasised the interdependence between water supply and sanitation infrastructure, noting that failures in water systems, such as pipe bursts during winter, can directly affect the functioning of sanitation services. Overall, whilst institutional awareness appears to be well established, strengthening enforcement mechanisms, improving coordination, and enhancing resource allocation remain key priorities.

4.2.2 Non-governmental organisations

Representatives of NGOs demonstrated a strong awareness of sanitation challenges and were particularly active in identifying practical, context-specific solutions. NGOs were reported to play an important intermediary role between households and formal institution, supporting both the implementation and maintenance of sanitation facilities in cold-climate conditions.

Their contributions included the construction and management of public sanitation facilities, as well as the provision of technical guidance to households. This guidance focused on practical measures such as appropriate insulation, installation depth relative to frost penetration, and general system maintenance.

However, NGO stakeholders also identified constraints affecting the scale and sustainability of their interventions. These included limited financial resources, restricted technical capacity, and insufficient coordination with government authorities. Respondents indicated that enhanced collaboration with public sector institutions, together with increased funding and capacity-building support, would be required to expand the effectiveness of NGO-led initiatives.

4.2.3 Private desludging operators

Private desludging operators reported a range of operational challenges associated with winter conditions. These included the formation of frozen sludge layers, restricted access to sanitation facilities, and the freezing of infrastructure components such as manholes and pipes.

Operators also highlighted logistical constraints affecting service delivery, including

limitations in equipment suitability for cold conditions and the absence of consistent regulatory oversight. In addition, the prevalence of improperly installed on-site systems was identified as a significant challenge, as such systems complicate desludging operations and increase the frequency of service interventions.

These findings underscore the importance of appropriate system design, installation standards, and regulatory enforcement in supporting effective service provision under cold-climate conditions.

4.2.4 Households

Household respondents reported a range of operational challenges associated with on-site sanitation systems. Approximately one-third of participants indicated limited access to sanitation facilities during the cold season. Reported issues included septic tank overflows, sewage backups into indoor fixtures, and restricted access to pit latrines due to snow and ice accumulation.

Pit latrines were particularly affected by deep frost penetration (1.5–2 m or more), which impeded biological decomposition and infiltration processes. Households also reported the need for more frequent desludging during winter due to reduced system performance.

A prominent finding concerned the widespread use of single-compartment septic tanks constructed without professional guidance, particularly in low-income neighbourhoods. Three households in Astana and seven in Atbasar reported installing systems using concrete rings without consideration of frost depth or subsoil permeability. These systems functioned largely as holding tanks, provided limited solids–liquid separation, and required frequent emptying.

In contrast, professionally installed two-compartment systems demonstrated more reliable performance and reduced maintenance requirements. Overall, household experiences reflected variation in system type, installation quality, and local environmental conditions across the two study locations.

4.3 Cross-stakeholder comparison

Differences in perspectives were observed across stakeholder groups. Government stakeholders primarily emphasised institutional and regulatory constraints, whilst households focused on operational and accessibility challenges. Private operators highlighted logistical and technical barriers to service provision, whereas NGOs emphasised community-level support and practical interventions.

Despite these differences, all stakeholder groups reported recurring challenges associated with freezing conditions, infrastructure performance, and maintenance requirements, indicating shared system-level constraints across the sanitation service chain.

4.4 Awareness and system-level challenges

The survey results indicate a high level of awareness of sanitation challenges associated with cold climates, with all participants reporting operational difficulties during extreme cold winter conditions.

Twenty out of thirty respondents indicated familiarity with SDG 6 and Kazakhstan's national commitments to achieving universal access to safely managed sanitation by 2030. However, concerns were raised regarding the effectiveness of current implementation efforts.

Respondents also highlighted the interdependence between water supply and sanitation systems. Seventy-seven percent of participants expressed concerns about unreliable water supply and the risk of pipe freezing and bursting during winter. Such failures were reported to directly affect sanitation system functionality, particularly in systems dependent on water-based flushing.

In Atbasar, respondents noted that the existing wastewater treatment plant struggles to operate under winter conditions and requires rehabilitation. Stakeholders further highlighted the vulnerability of water supply infrastructure to freezing. As one public sector representative stated, "We just hope the pipes will not burst or freeze during this year's winter season."

4.5 Performance and accessibility of on-site sanitation systems

Significant limitations in the performance and accessibility of on-site sanitation systems were reported. Although only a small proportion of respondents described their facilities as unsafe year-round, a larger proportion indicated that cold weather frequently restricted access to functional sanitation.

Pit latrines were particularly vulnerable to freezing conditions, while septic tanks experienced reduced performance associated with solids accumulation and decreased biological activity. Households reported adopting temporary coping strategies, including snow removal, sand spreading, and the construction of wooden walkways to improve access. However, these measures were often insufficient during periods of heavy snowfall or rapid freeze-thaw cycles.

System performance was influenced by both installation quality and maintenance practices (Table 4.1). Regular inspection and desludging were identified as essential for preventing overflows (wastewater escaping from the system into the environment) and backups (wastewater returning to indoor fixtures).

NGOs were reported to support sanitation access through the construction and maintenance of public sanitation facilities and by providing technical advice to households. In the surveyed cities, NGOs supported the construction and upkeep of ten public toilets in high-footfall areas and provided guidance on insulation, frost protection, and basic design improvements.

Table 4.1: Stakeholder perceptions regarding the sanitation situation in the northern region of Kazakhstan.

Stakeholder perceptions	**Yes (%)**
Are you aware of sanitation problems related to cold weather conditions?	100
Have you heard about SDG 6?	67
Do you think your sanitation problems are addressed at the government level?	63
Is your sanitation facility connected to centralised wastewater treatment facilities?	50
Does your household have any problems with the operation and maintenance of existing sanitation infrastructure?	33

Stakeholder perceptions	Yes (%)
Do you have limited access to safe sanitation during the cold season?*	33
Do you have limited access to safe sanitation during all seasons?*	7

* *The term limited access is used to describe how people found it harder to physically access the latrines in the cold season than in the warm season (e.g. because of ice, snow, etc.).*

4.6 Sanitation challenges in cold-climate conditions

The most frequently reported sanitation problems during winter were foul odours and sewage overflows, cited by 93% of respondents. In contrast, only a small proportion identified pathogen contamination as a primary concern. These findings align with previous studies, such as Beisenova et al. (2021), which linked inadequate septic system design to coliform and fungal contamination in local groundwater.

Stakeholders identified five main categories of challenges affecting sanitation system performance under cold-climate conditions:

1. **Unsafe access to sanitation facilities**

   Slippery and icy access paths were widely reported, particularly where facilities were located at a distance from dwellings. Temporary mitigation measures were often insufficient during severe conditions (Figure 4.1).

2. **Formation of frozen sludge layers**

   Frozen sludge crusts impeded desludging and routine maintenance (Figure 4.2), particularly during prolonged sub-zero temperatures.

3. **Freezing of structural components**

   Manholes and pit latrine slabs frequently froze shut, preventing inspection and maintenance (Figure 4.3).

4. **Pipe freezing and hydraulic failure**

   Pipe freezing was consistently identified as a major cause of system malfunction, often linked to insufficient insulation and inadequate burial depth.

5. **Overflow and backup events**

   Septic tank overflows and sewage backups were commonly reported, particularly in systems with limited capacity and irregular maintenance (Figure 4.4).

These challenges were observed across both study locations and were strongly associated with sustained low temperatures. Improperly constructed systems lacking adequate insulation, burial depth, and design considerations were particularly vulnerable (Table 4.2).

Table 4.2: Survey results on sanitation challenges and barriers in cold regions of Kazakhstan.

Barriers and challenges	**Agree (n = 30)**	**%**
Understanding the importance of sanitation issues and their impact on drinking water quality and the environment	30	100
Limited political support	15	50
Limited public awareness	15	50
Limited financial support	15	50
Limited availability of sanitation experts	14	47
Limited coordination of sanitation issues among stakeholders	13	43

Figure 4.1: A slippery path leading to the on-site sanitation facility in Astana.

Figure 4.2: Crust formed inside the septic tank.

Figure 4.3: A common pit latrine scene in wintertime.

Figure 4.4: Sewage overflow incident from a septic tank (February 2024).

4.7 Reported sanitation practices and coping strategies

Stakeholders reported a range of practices aimed at maintaining sanitation system functionality during cold conditions. These included pre-winter desludging, insulation of infrastructure components, installation below frost depth where feasible, maintenance of access pathways, and the use of backup sanitation options.

Implementation varied across stakeholder groups. Public sector representatives referred to state-led interventions, such as heated public toilets and insulated water and wastewater pipelines. In contrast, households relied primarily on locally developed coping strategies, often supported by NGOs.

NGOs played a facilitating role by providing technical advice and supporting community-level implementation. As one NGO representative stated, "If we want safe sanitation, we need to look for sustainable solutions on our own."

The stakeholders highlighted the importance of insulating infrastructure components using materials such as felt, fiberglass, or polyurethane foam, and the use of durable, cold-resistant materials such as reinforced concrete or high-density polyethylene (HDPE), which was perceived as more resistant to freezing (Agarwal and Gupta, 2017; Nguyen et al., 2021). Pre-winter desludging was also frequently identified as a measure to reduce overflow risks during periods of reduced biological activity, consistent with practices reported in other cold-climate contexts adopted in Finland, Sweden, and Norway (Laukka et al., 2022). Installation of systems below the frost line was considered effective, although associated costs limited its feasibility for low-income households. Maintaining safe access pathways and ensuring the availability of backup sanitation options were also commonly reported.

Despite the range of practices identified, respondents did not report modifications to septic tank design, such as the adoption of multi-compartment systems to enhance solids separation and digestion. This suggests limited awareness of fundamental design principles for effective wastewater treatment.

4.8 Summary of findings

The stakeholder survey demonstrates that sanitation system performance in cold-climate regions is influenced by a combination of operational challenges, infrastructure constraints, and seasonal environmental conditions.

Recurring issues included septic tank overflows, sewage backups, and freezing-related failures, often associated with inadequate system design, insufficient burial depth, and inappropriate material selection. Installation quality and maintenance practices were identified as key determinants of system performance. Similar observations have been reported in previous studies (Withers et al., 2014).

Stakeholders also highlighted gaps in environmental monitoring and institutional capacity, particularly in relation to groundwater quality and inspection of on-site systems. Limited institutional capacity, including staff shortages, was identified as a constraint affecting monitoring and maintenance activities.

NGOs were identified as important actors in supporting sanitation access through community engagement and technical guidance, although their impact was constrained by resource and coordination limitations. Evidence from prior studies supports the effectiveness of such participatory approaches. Close cooperation between local authorities, NGOs, and community stakeholders has been shown to improve sanitation outcomes and promote equitable service delivery (Pathak, 1996; Thye et al., 2015). Implementing a collaborative governance framework could therefore help address systemic barriers and advance progress towards safe, climate-resilient sanitation across Kazakhstan.

Overall, these findings provide an empirical basis for understanding sanitation system performance in cold-climate settings and inform subsequent analysis of treatment processes under both field and laboratory conditions.

5 Results: Physicochemical and microbial dynamics in laboratory-scale anaerobic systems

Extracts of this chapter have appeared in Paper 2 and Paper 3 (see Section 1.7).

5.1 Overview of laboratory-scale results and analytical scope

This chapter presents physicochemical and microbiological results from laboratory-scale anaerobic reactors operated under controlled temperature regimes representative of seasonal conditions relevant to septic tank operation in cold-climate regions. The laboratory experiment was conducted prior to the initiation of field sampling and was specifically designed to isolate the influence of temperature on anaerobic treatment performance under simplified, reproducible conditions.

The experimental design deliberately encompassed a broad thermal range, spanning psychrophilic, transitional, and upper mesophilic reference conditions. This approach addresses a key limitation of existing laboratory studies, which have typically focused on narrow temperature intervals or single operating regimes. By decoupling temperature effects from the structural, operational, and behavioural variability inherent to household systems, the laboratory reactors establish a controlled reference framework for interpreting temperature-driven patterns later observed in full-scale septic tanks.

Molecular analysis was conducted to characterise temperature-dependent variation in microbial and viral communities and their functional potential. Shotgun metagenomic sequencing was used to quantify prokaryotic taxa, classify anaerobic digestion functional guilds and pathogen-associated bacteria, and assess viral community composition, while functional gene profiling was applied to examine key metabolic pathways of AD guilds.

This chapter examines temporal variation in key physicochemical indicators of anaerobic treatment performance, together with associated changes in microbial and viral community structure and functional potential under defined thermal conditions. These controlled observations provide a mechanistic reference for temperature-dependent processes and establish a baseline against which the more complex and heterogeneous field-scale dynamics are evaluated.

5.2 Physicochemical characteristics

Batch anaerobic reactors were operated for 21 days at fixed temperatures of 4 °C, 12 °C, 24 °C and 36 °C to quantify temperature-dependent variation in wastewater physicochemical parameters under controlled conditions. All reactors were run in triplicate.

Across all temperature regimes, effluent pH remained within the range of 5–7 throughout the incubation period, indicating stable anaerobic conditions and remaining consistent with the near-neutral pH values measured in the field-scale septic tank samples (Chen et al., 2020).

Temporal decreases in COD, BOD_5, TS, and VS were observed under all temperature conditions (Figures 5.1-5.4). However, both the rate and the extent of removal varied markedly with temperature. Reactors operated at 24 °C and 36 °C showed the greatest reductions across all parameters, with the most pronounced responses occurring at 36 °C. At this temperature, BOD_5 declined rapidly during the early incubation period and the highest overall VS removal was achieved, reflecting enhanced biological activity under mesophilic conditions.

In contrast, reactors incubated at 4 °C and 12 °C showed slower, more gradual declines in COD, BOD_5, TS, and VS. Removal efficiencies under these conditions were reduced relative to warmer reactors, with modest decreases in COD and BOD_5 and limited, though detectable, solids reduction. These patterns indicate that anaerobic processes persisted at low temperatures, albeit with substantially constrained reaction kinetics.

Overall, the laboratory results demonstrate a clear and consistent temperature dependence of physicochemical treatment performance. By isolating temperature as the primary experimental variable, these findings provide a controlled baseline for interpreting

seasonal variability in organic matter and solids removal observed in household septic tanks under field conditions.

The parallel trends observed across TS, VS, COD, and BOD_5 removal suggest strong coupling between solids degradation and organic matter conversion, indicating that a substantial fraction of total solids is biodegradable. This relationship is explored further in the Discussion.

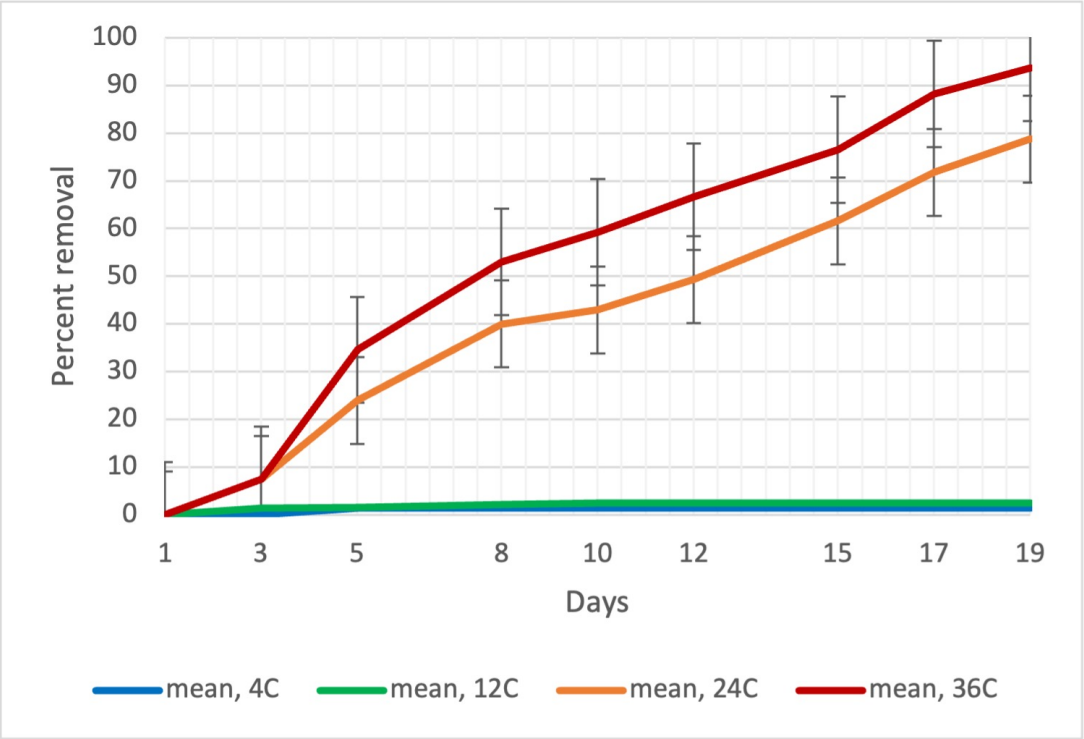

Figure 5.1: Average COD concentration in lab-scale reactors incubated at 4 °C, 12 °C, 24 °C, and 36 °C.

5.3 Implications of laboratory-scale temperature effects

The laboratory-scale results demonstrate a strong and systematic influence of temperature on anaerobic physicochemical treatment performance when other operational variables are held constant. Increasing incubation temperature within the mesophilic range was associated with faster reaction rates and greater overall reductions in COD, BOD_5, TS, and VS, reflecting enhanced microbial activity and organic matter conversion under warmer conditions. These patterns are consistent with established anaerobic digestion kinetics, in which enzymatic activity and microbial growth are positively correlated with temperature (Gerardi, 2003; Metcalf and Eddy, 2013).

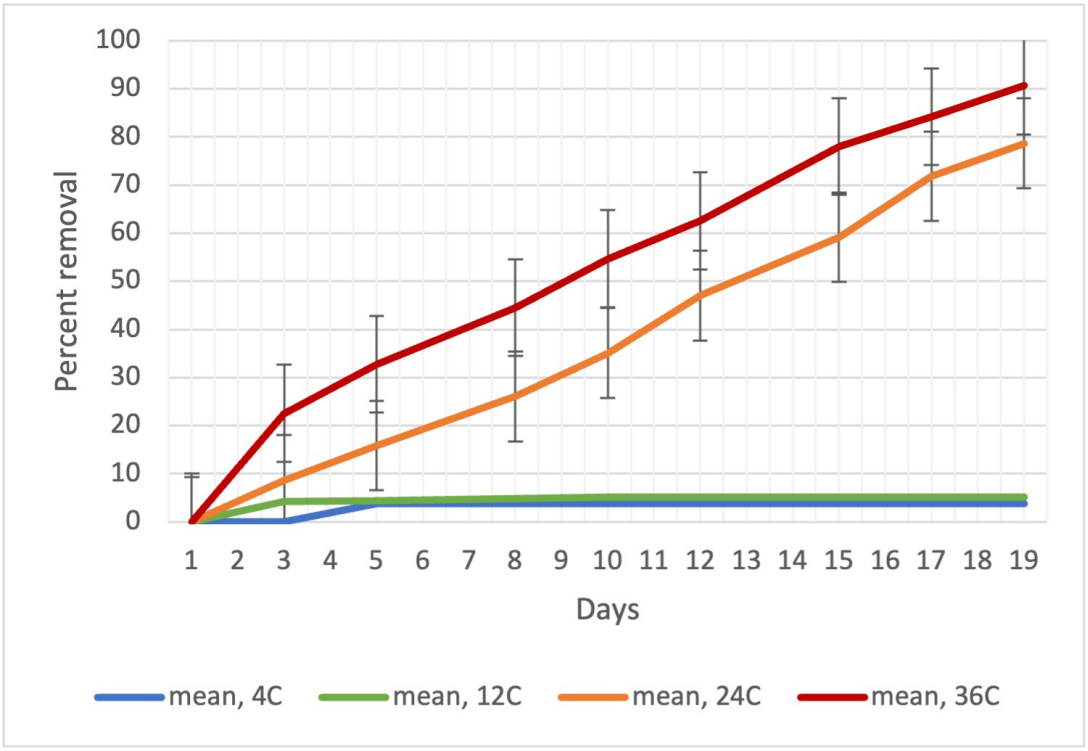

Figure 5.2: Average BOD_5 concentration in lab-scale reactors incubated at 4 °C, 12 °C, 24 °C, and 36 °C.

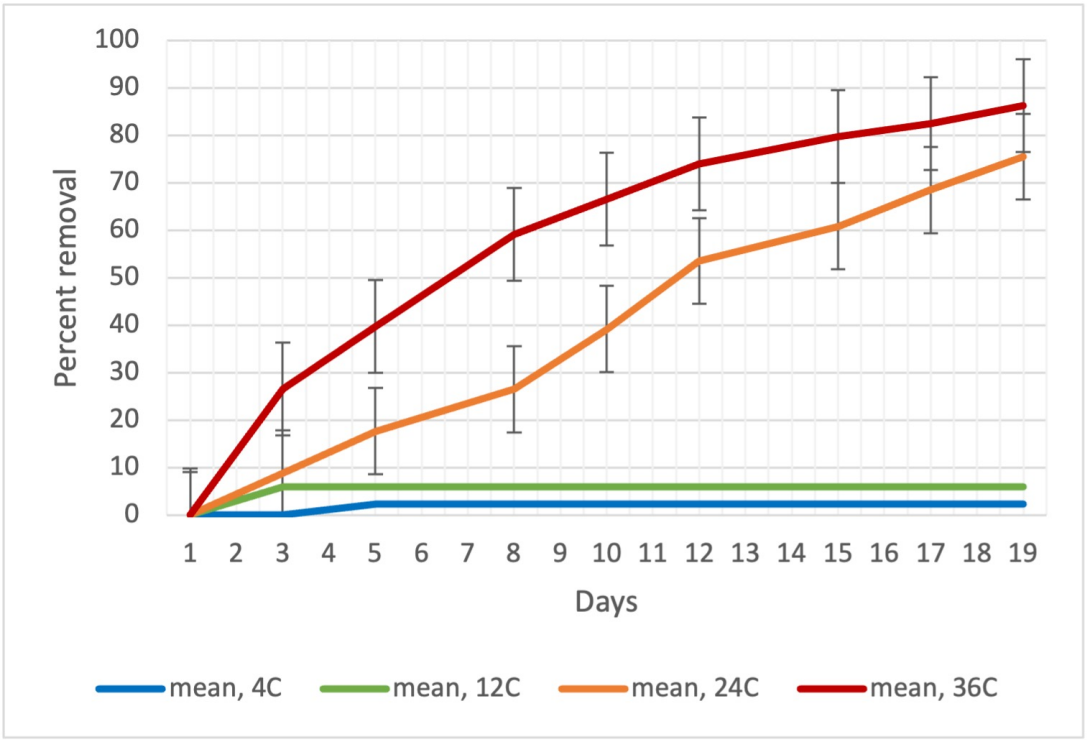

Figure 5.3: Average TS concentration in lab-scale reactors incubated at 4 °C, 12 °C, 24 °C, and 36 °C.

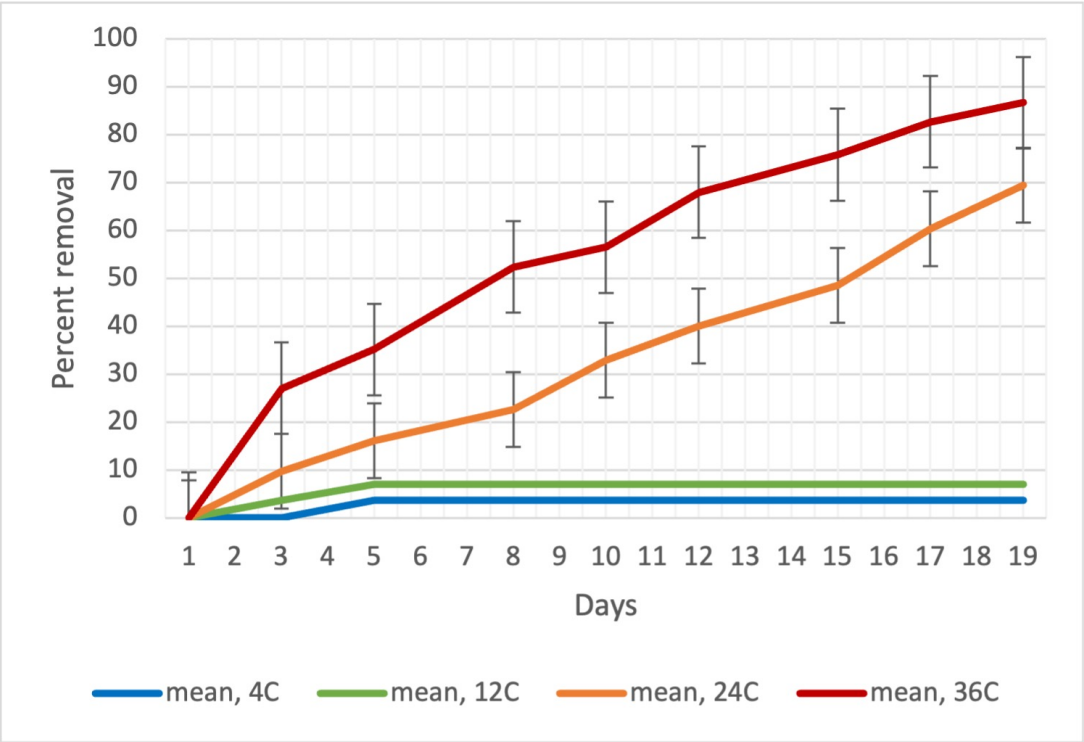

Figure 5.4: Average VS concentration in lab-scale reactors incubated at 4 °C, 12 °C, 24 °C, and 36 °C.

Notably, anaerobic activity was sustained across all temperature regimes, including at 4 °C and 12 °C. Although removal rates were substantially lower at these temperatures, measurable declines in COD, BOD_5, TS, and VS indicate that biological processes persisted under cold conditions. This finding aligns with previous observations of psychrophilic and low-temperature anaerobic systems, where treatment performance is constrained primarily by reaction kinetics rather than by complete biological inhibition (Luostarinen, 2005; Luostarinen et al., 2007).

The contrast between warm- and cold-reactor performance provides a mechanistic basis for understanding seasonal variability commonly reported in full-scale septic tank systems. Under warmer conditions, enhanced microbial metabolism supports more effective degradation of organic matter and solids, whereas colder conditions impose kinetic limitations that reduce treatment processes efficiency (Luostarinen, 2005; Luostarinen et al., 2007; Pussayanavin et al., 2015). By isolating temperature effects from structural and operational variability, the laboratory reactors demonstrate that temperature exerts a strong and direct influence on anaerobic physicochemical performance.

These controlled observations establish a reference context for subsequent field-scale

analyses. While household septic tanks operate under more complex and heterogeneous conditions – including variable influent composition, episodic loading, and infrastructure constraints – the laboratory results provide a benchmark against which seasonal trends observed *in situ* can be interpreted.

5.4 Microbial and viral community structure

5.4.1 Community dissimilarity across temperatures

The distribution of Aitchison distances revealed temperature-dependent differences in community structure that varied in character across the experimental gradient (Figure 5.5). Within-temperature variability was lowest at 4 °C, where samples clustered tightly with small inter-sample distances and a low median, indicating high compositional similarity and community stability under psychrophilic conditions. At 24 °C, within-temperature variability was similarly low and tight, with narrow IQR and low median comparable to 4 °C, indicating that communities under these mesophilic conditions were also compositionally consistent across sampling points.

In contrast, 12 °C and 36 °C exhibited markedly different patterns, though for distinct reasons. At 12 °C, the median within-temperature distance was low, similar to 4 °C and 24 °C, but the IQR was extremely large, extending from near zero to approximately 160. This indicates that while many samples were compositionally similar to one another, a small subset diverged dramatically, producing a highly heterogeneous distribution characteristic of a transitional or unstable community state. At 36 °C, both the median and the spread of within-temperature distances were high, with the IQR spanning approximately 15 to 165 and the median approaching 160, indicating that compositional divergence among samples was consistently pronounced rather than driven by isolated outliers.

Between-temperature distances were consistently higher than within-temperature distances at 4 °C and 24 °C, confirming that temperature acts as a meaningful structuring factor across the experimental gradient. However, between-temperature distances were comparable to or exceeded by the upper range of within-temperature distances at 12 °C and 36 °C, reflecting the elevated compositional heterogeneity characteristic of these conditions.

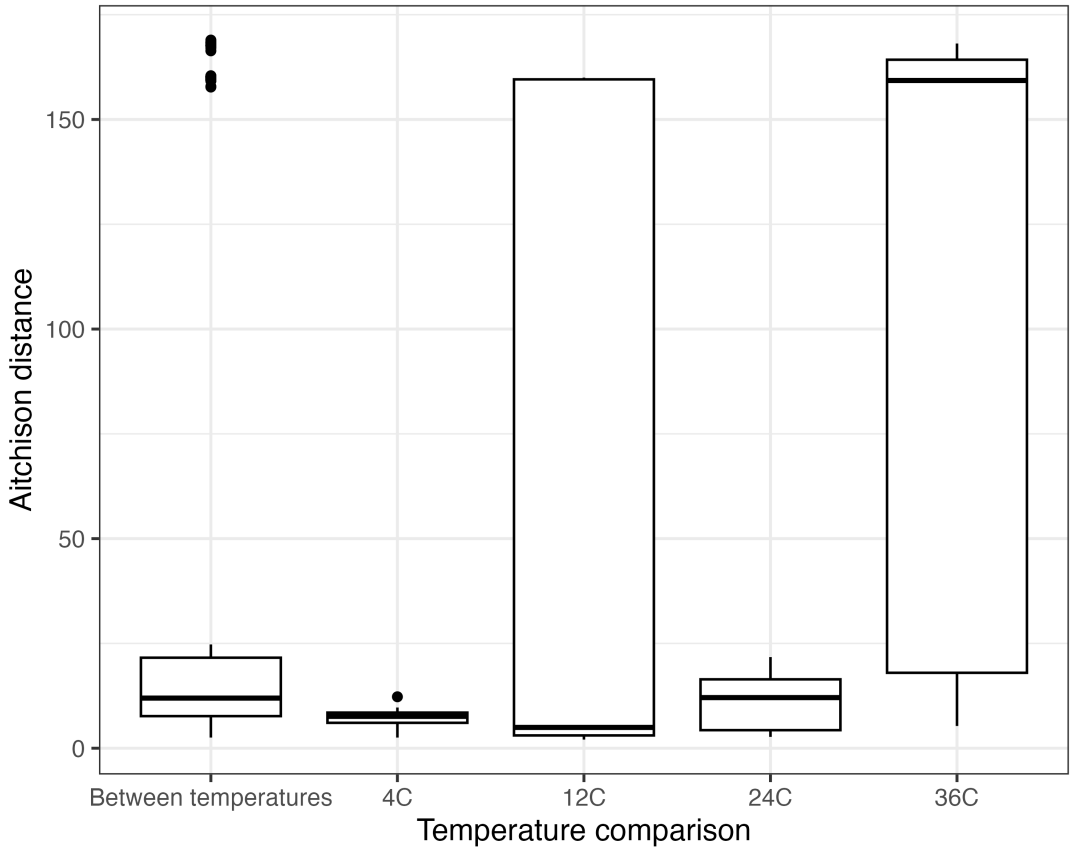

Figure 5.5: Aitchison distance within and between temperature conditions. Boxplots showing the distribution of Aitchison distances for prokaryotic and viral communities within each temperature condition (4 °C, 12 °C, 24 °C, 36 °C) and between temperature conditions. The central line represents the median; boxes indicate interquartile ranges; whiskers extend to 1.5 x IQR; points denote outliers.

5.4.2 Alpha-diversity patterns

A strong inverse relationship was observed between prokaryotic and viral diversity when assessed using the inverse Simpson index (Spearman $\rho = -0.84$, p = 3.7 x 10^{-7}; Figure 5.6). Samples with higher prokaryotic diversity consistently corresponded to lower viral diversity, and vice versa, with 36 °C samples exhibiting the highest prokaryotic dominance and among the lowest viral diversity values across the dataset.

A near-perfect negative correlation was observed between prokaryotic and viral richness (Spearman $\rho = -0.99$, p = 2.5 x 10^{-21}; Figure 5.7). Examination of the scatterplot reveals that samples from 4 °C, 12 °C, and 24 °C overlapped almost entirely in their richness on both axes, forming a tight cluster with minimal variation. The correlation is therefore driven primarily by the 36 °C temperature group, which exhibited distinctly higher prokaryotic richness and lower viral richness relative to all other temperatures. The near-perfect coefficient thus reflects a strong between-group contrast rather than a continuous gradient across the full temperature range.

In contrast, no statistically significant relationship was observed between prokaryotic and viral Shannon diversity (Spearman $\rho = -0.12$, p = 0.58; Figure 5.8). The wide scatter of points across temperature groups, with no consistent directional trend, indicates that evenness-weighted diversity does not exhibit a systematic association between domains across the experimental conditions. Taken together, the alpha-diversity results indicate that prokaryotic and viral communities are inversely coupled in terms of richness and dominance structure, but that this coupling is most strongly expressed at the extremes of the temperature gradient rather than uniformly across it.

5.4.3 Coupling of prokaryotic and viral beta diversity

A strong and statistically significant correlation was observed between prokaryotic and viral beta diversity (Mantel r = 0.86, p = 0.001; Figure 5.9), indicating that pairwise compositional differences among prokaryotic communities were closely mirrored by corresponding differences in viral communities. The scatterplot reveals a strongly bimodal distribution: the vast majority of pairwise sample comparisons clustered at near-zero Aitchison distances on both axes, indicating high compositional similarity among samples within the same temperature condition, while a small number of

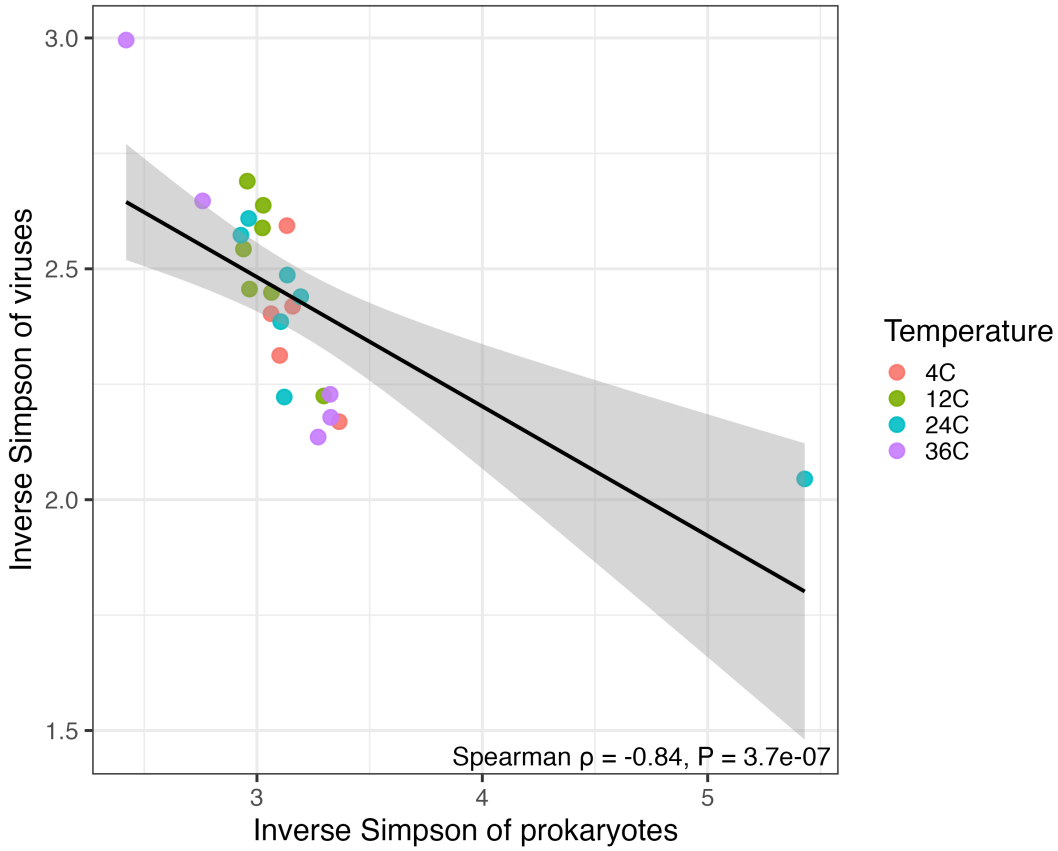

Figure 5.6: Relationship between prokaryotic and viral diversity (inverse Simpson index). Scatterplot showing the relationship between inverse Simpson diversity of prokaryotic and viral communities across all samples. Points are coloured by incubation temperature. The fitted regression line with 95% confidence interval is shown. Spearman correlation coefficient and associated p-value are indicated.

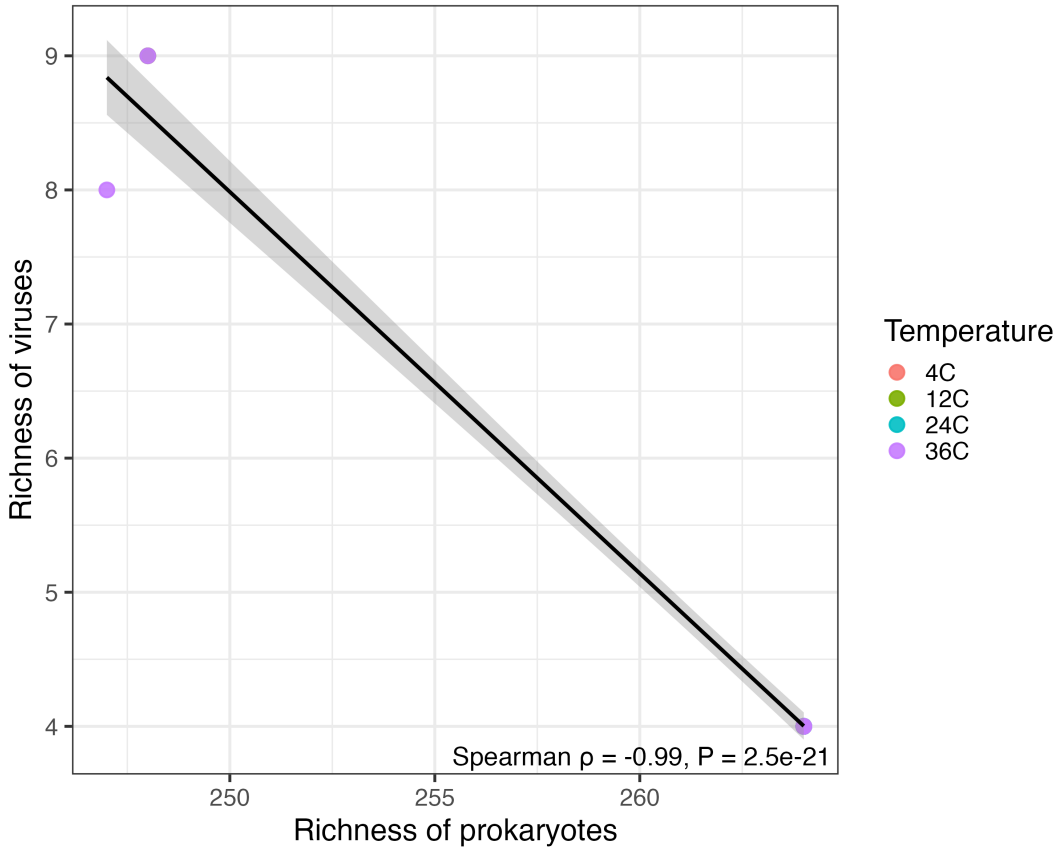

Figure 5.7: Relationship between prokaryotic and viral richness. Scatterplot showing the association between prokaryotic and viral richness across all samples. Points are coloured by temperature. The fitted regression line with 95% confidence interval is shown, together with Spearman correlation statistics.

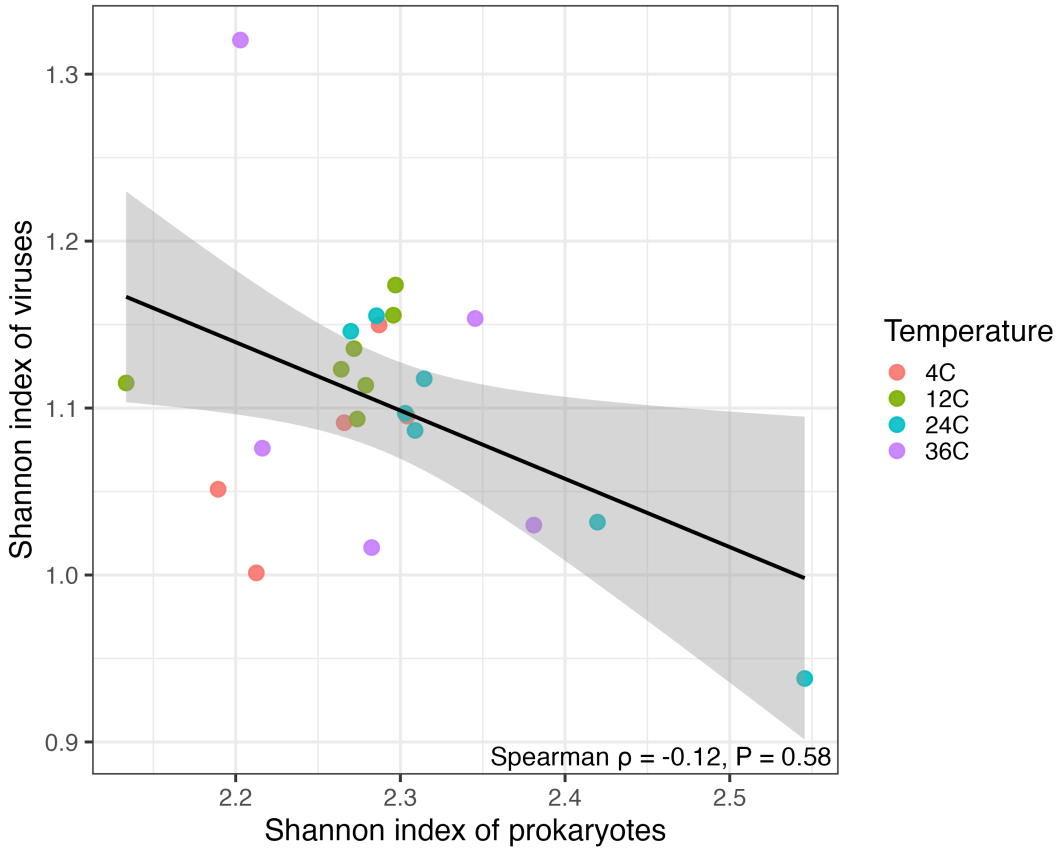

Figure 5.8: Relationship between prokaryotic and viral Shannon diversity. Scatterplot illustrating the relationship between Shannon diversity indices of prokaryotic and viral communities. Points are coloured by temperature. The fitted regression line with 95% confidence interval is shown, along with the Spearman correlation coefficient and p-value.

comparisons – corresponding to between-temperature contrasts involving the outlier samples identified at 12 °C and 36 °C – produced very high distances on both axes (prokaryotic distances exceeding 150, viral distances exceeding 20). The high Mantel coefficient is therefore primarily driven by concordance at the extremes of compositional difference rather than by a uniformly continuous relationship across the full range of pairwise comparisons. Within the main cluster of low-distance comparisons, prokaryotic and viral compositional differences track closely, supporting the interpretation of tightly coupled community structuring across domains under most conditions.

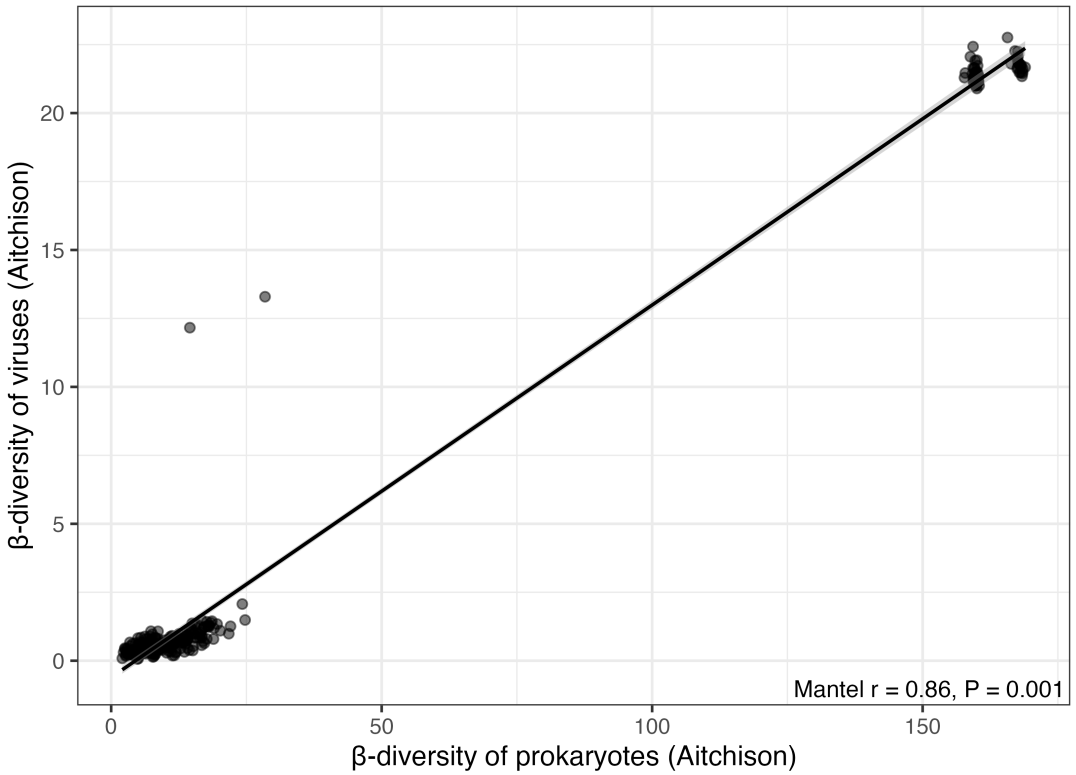

Figure 5.9: Concordance between prokaryotic and viral beta diversity. Scatterplot showing the relationship between pairwise Aitchison distances of prokaryotic and viral communities. Each point represents a pairwise comparison between samples. The fitted regression line is shown. Mantel correlation coefficient and associated p-value are reported.

5.4.4 Ordination of community composition

Principal Coordinates Analysis based on Aitchison distances further characterised the influence of temperature on community structuring. For prokaryotic communities, the first principal coordinate explained 97.6% of the variance and separated a subset of samples with markedly divergent composition from the main cluster (Figure 5.10). The majority of samples – including most samples from 4 °C and 24 °C, and several

from 12 °C and 36 °C – clustered tightly near the origin along PCoA1, indicating high compositional similarity. The samples that separated to the right along PCoA1, with scores of approximately 140–150, included both 12 °C and 36 °C samples, consistent with the extreme within-temperature variance observed for both conditions in the Aitchison distance analysis. Thus, separation along PCoA1 is not exclusively associated with higher temperatures but reflects the emergence of distinct compositional states at both the transitional (12 °C) and upper mesophilic (36 °C) temperature conditions. Along the second principal coordinate (1.1% of variance), additional spread was evident, particularly among 24 °C and 36 °C samples, including two samples with markedly negative PCoA2 scores, indicating secondary axes of compositional variation not captured by the primary gradient.

A similar but more compressed pattern was observed for viral communities, where PCoA1 explained 91% of variance (Figure 5.11). The majority of samples formed a tight cluster near the origin, including all 4 °C and 24 °C samples, while one 12 °C sample and two 36 °C samples separated along PCoA1 at scores of approximately 16–18. One 36 °C sample additionally showed a strongly negative PCoA2 score, indicating an isolated divergent viral community state at upper mesophilic conditions. As with the prokaryotic ordination, separation along the primary axis was not restricted to higher temperatures – the divergent 12 °C viral sample mirrors the pattern observed in prokaryotic community structure, confirming that compositional instability at this transitional temperature is expressed across both microbial and viral domains. The second axis (8.6% of variance) captured a small but meaningful proportion of additional community variation.

5.4.5 Implications

Together, these results demonstrate that temperature exerts a strong but heterogeneous influence on both prokaryotic and viral community structure across the experimental gradient. Rather than a simple linear relationship between temperature and community dissimilarity, the data reveal three distinct structural regimes. At 4 °C and 24 °C, communities were compositionally stable and homogeneous within each temperature condition. At 12 °C, communities exhibited low median dissimilarity but extreme variance, indicating a transitional state in which most samples maintained similar composition while a subset underwent marked divergence – a pattern consistent with unstable or

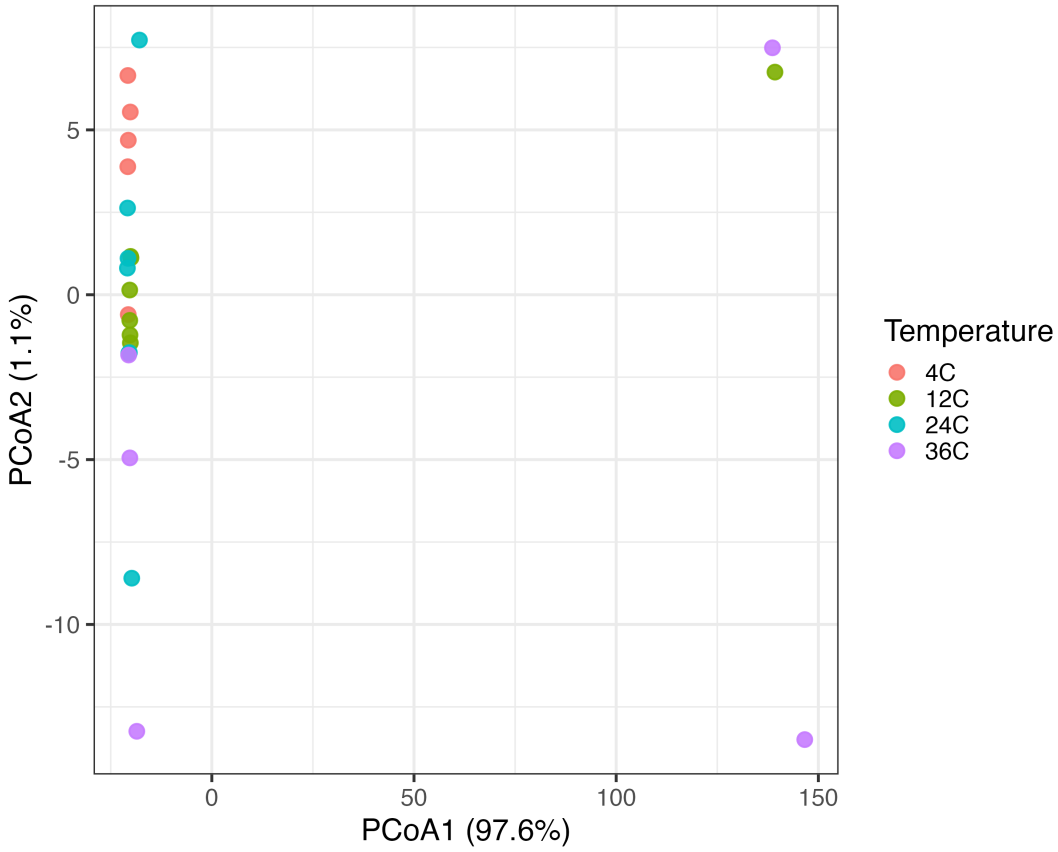

Figure 5.10: Principal Coordinates Analysis of prokaryotic community composition. Ordination plot based on Aitchison distances showing the distribution of prokaryotic communities across samples. Points are coloured by incubation temperature. The percentage of variance explained by each principal coordinate axis is indicated.

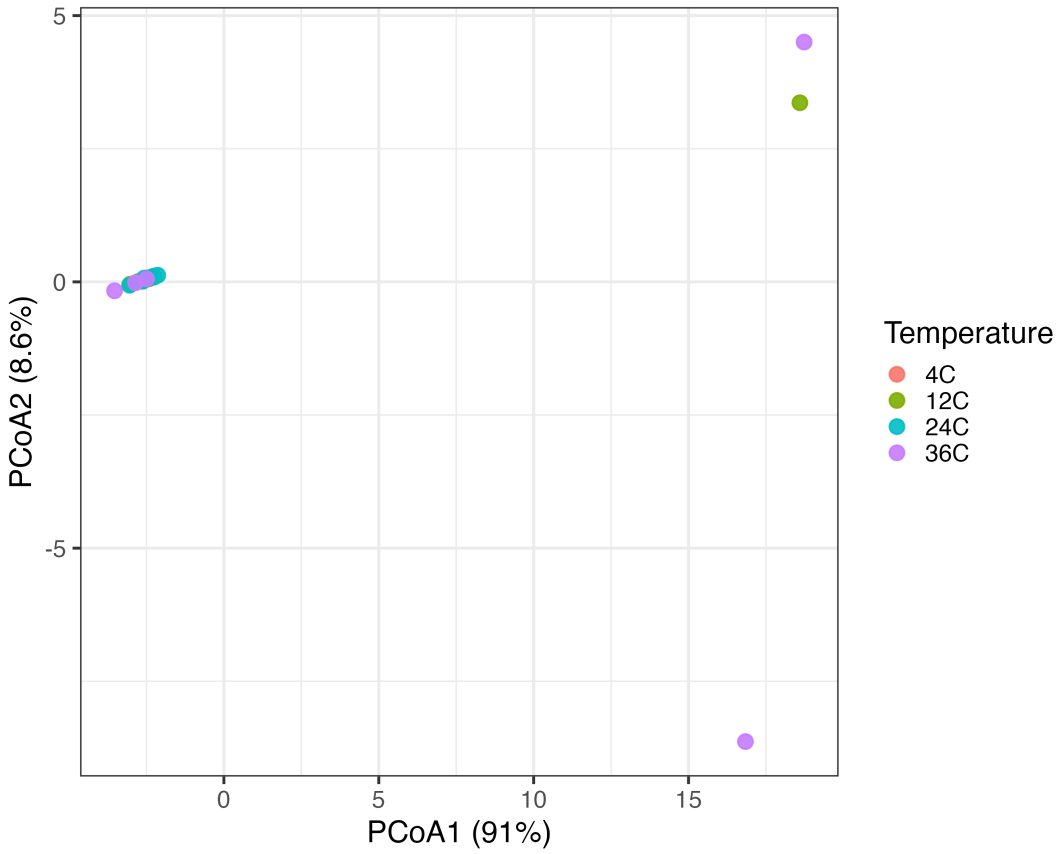

Figure 5.11: Principal Coordinates Analysis of viral community composition. Ordination plot based on Aitchison distances showing the distribution of viral communities across samples. Points are coloured by incubation temperature. The percentage of variance explained by each principal coordinate axis is indicated.

fluctuating ecological dynamics under moderately cold conditions. At 36 °C, communities were consistently and highly dissimilar from one another, reflecting pronounced temporal compositional turnover under upper mesophilic conditions. The strong concordance between prokaryotic and viral beta diversity, driven primarily by between-temperature contrasts, indicates that changes in one domain are closely aligned with changes in the other, but that this alignment is most clearly expressed at the compositional extremes defined by the outlier temperature conditions.

5.5 Relative abundance dynamics of AD guilds and viral communities across temperature and incubation time

This section examines the functional composition of AD-associated microbial guilds and viral communities in controlled laboratory-scale anaerobic reactors incubated under four temperature conditions (4 °C, 12 °C, 24 °C, and 36 °C). Prokaryotic taxa were classified into four functional guilds central to anaerobic digestion: fermentative bacteria, SAOB, SRB, and methanogenic archaea. The relative abundance and taxonomic composition of these guilds, together with associated viral communities, were analysed to characterise temperature-dependent microbial organisation under controlled experimental conditions.

5.5.1 Relative abundance of AD functional guilds

The relative abundance of AD functional guilds varied systematically across temperature and incubation time (Figure 5.12). Fermentative bacteria dominated across all temperature conditions and sampling points, consistently comprising the majority of the microbial community irrespective of temperature. This pattern confirms the central and temperature-resilient role of fermentation as the primary upstream process in the anaerobic degradation cascade under all conditions tested.

At 4 °C, fermenters accounted for nearly the entirety of the community at most sampling points, with minor and temporally stable contributions from SRB and SAOB. Methanogens remained at very low relative abundance throughout, with limited temporal variation across the incubation period, indicating strongly suppressed methanogenic activity under psychrophilic conditions.

At 12 °C, fermenters remained dominant but a gradual and progressive increase in

the relative abundance of both SRB and SAOB was observed over the incubation period. SRB, in particular, showed a clear temporal increase from early to later sampling points, while methanogens remained consistently low. Community structure was relatively stable overall, with incremental rather than abrupt compositional shifts.

At 24 °C, a marked redistribution of guild composition was observed relative to lower temperatures. While fermenters still represented a substantial proportion of the community, their relative abundance decreased over time, accompanied by a pronounced increase in SRB and SAOB. SRB showed a clear and progressive temporal increase, becoming a dominant secondary guild by the later sampling points. Methanogens remained at low relative abundance but exhibited slightly increased contributions compared to 4 °C and 12 °C.

At 36 °C, the community structure differed substantially from that observed at lower temperatures. Fermenters remained dominant in some samples but showed greater variability. SRB exhibited a strong increase in relative abundance, becoming one of the dominant guilds at multiple sampling points and rivalling fermenters in relative contribution at the later incubation stages. SAOB also contributed more substantially compared to lower temperatures. Methanogens remained a minor component overall, though their relative abundance was slightly higher than at 4 °C and 12 °C, and temporal variation was more pronounced at this temperature, with clear compositional shifts between successive sampling points.

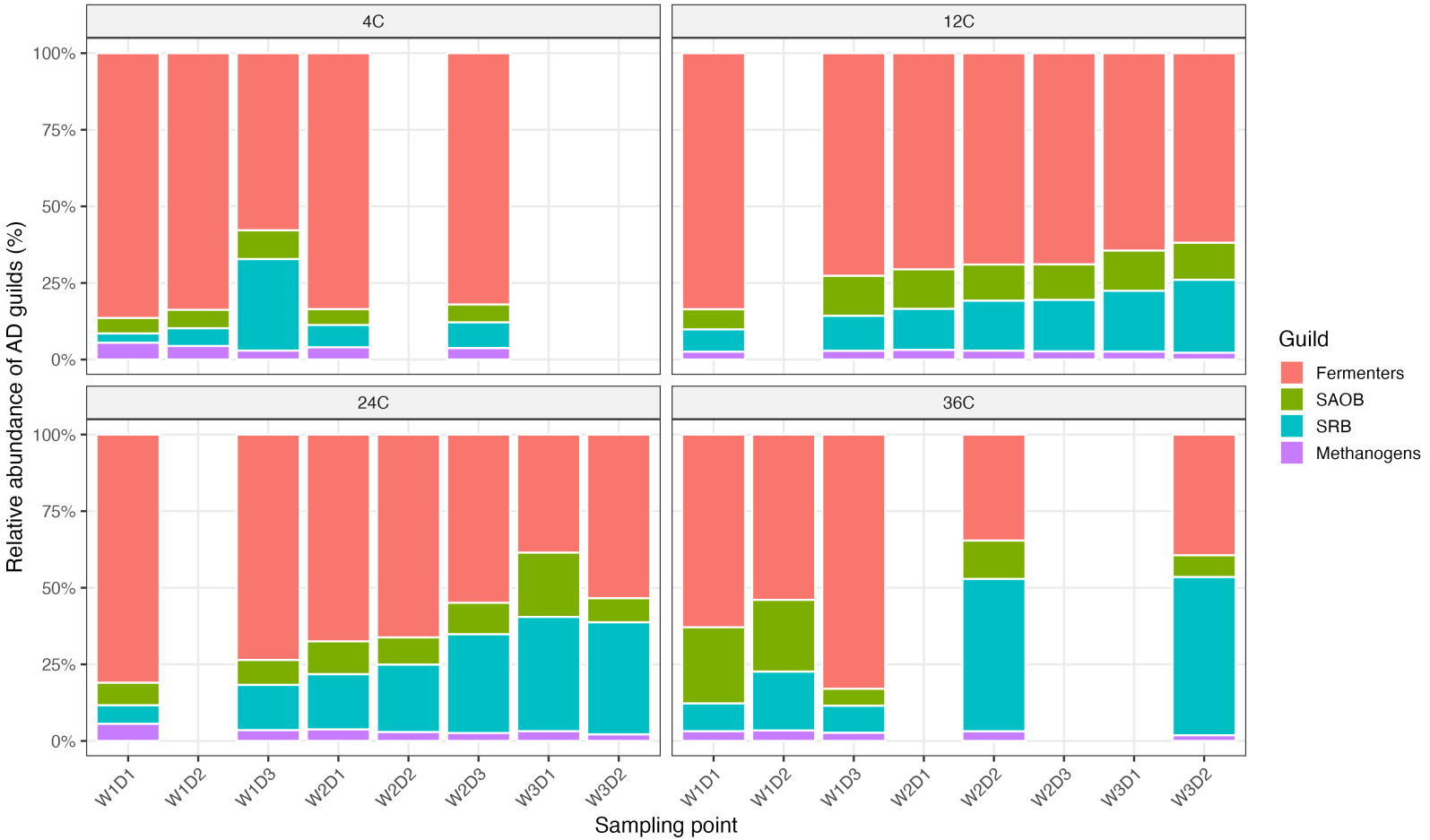

Figure 5.12: Relative abundance of AD guilds across sampling points at four incubation temperatures (4 °C, 12 °C, 24 °C, and 36 °C). Stacked bar plots show the relative abundance (%) of major AD functional guilds. Each bar represents a single sampling point.

5.5.2 Relative abundance of fermentative bacterial families

The composition of fermentative bacterial families showed clear temperature- and time-dependent variation (Figure 5.13). Across all temperatures, Anaerolineaceae and Carnobacteriaceae were among the most consistently represented families within the fermentative guild.

At 4 °C, the fermentative community was characterised by a limited number of dominant families with relatively stable composition across sampling points. Temporal variation was minimal, consistent with the low compositional turnover observed for the community as a whole under psychrophilic conditions.

At 12 °C, fermentative diversity increased, with a broader distribution of families contributing to the total abundance. Several families showed gradual changes in relative abundance over time, though no single family dominated consistently across all sampling points.

At 24 °C, the fermentative community became more heterogeneous, with multiple families contributing substantially, and temporal shifts evident in the relative dominance of different taxa. Some families increased while others declined between sampling points, reflecting greater compositional dynamism under mesophilic conditions.

At 36 °C, fermentative family composition showed strong temporal variability. While Anaerolineaceae remained consistently present, its relative contribution fluctuated considerably, and the distribution of relative abundances was more even across multiple families compared to lower temperatures, particularly in mid-incubation sampling points where compositional evenness was greatest.

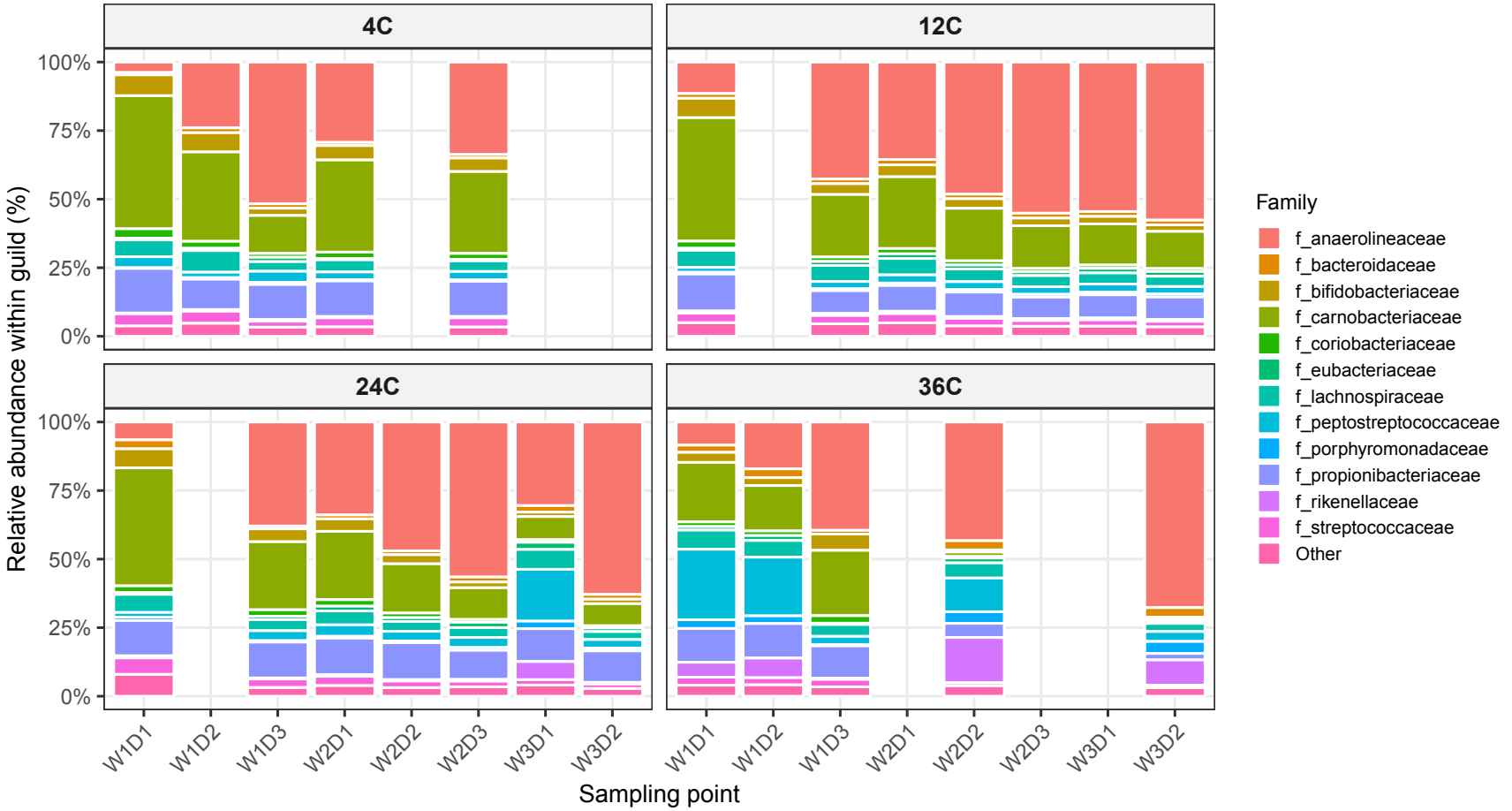

Figure 5.13: Relative abundance of fermentative bacterial families across sampling points at four incubation temperatures (4 °C, 12 °C, 24 °C, and 36 °C). Stacked bar plots showing the relative abundance (%) of dominant fermentative bacterial families. Each bar represents a single sampling point.

5.5.3 Relative abundance of methanogenic archaeal families

Methanogenic community composition varied across temperatures and incubation time, with Methanobacteriaceae and Methanotrichaceae constituting the primary families across all conditions (Figure 5.14).

At 4 °C, methanogenic community was dominated almost entirely by Methanobacteriaceae, with Methanotrichaceae as a secondary contributor. Temporal variation was limited and overall methanogenic relative abundance remained low.

At 12 °C, methanogenic diversity increased, with Methanobacteriaceae remaining prominent but Methanotrichaceae and Methanosarcinaceae contributing more substantially. Temporal changes were gradual, consistent with the incremental community shifts observed at this temperature.

At 24 °C, the composition shifted toward a more balanced distribution among multiple families. Methanobacteriaceae remained a major component, while contributions from Methanotrichaceae and Methanosarcinaceae increased and varied across sampling points, reflecting greater compositional dynamism under mesophilic conditions.

At 36 °C, methanogenic composition exhibited pronounced temporal variability. Methanobacteriaceae remained dominant in several samples, but Methanotrichaceae became markedly more prominent relative to lower temperatures, and Methanosarcinaceae also increased in relative contribution. The relative proportions of these families fluctuated considerably across the incubation period, indicating active restructuring of the methanogenic guild under upper mesophilic conditions.

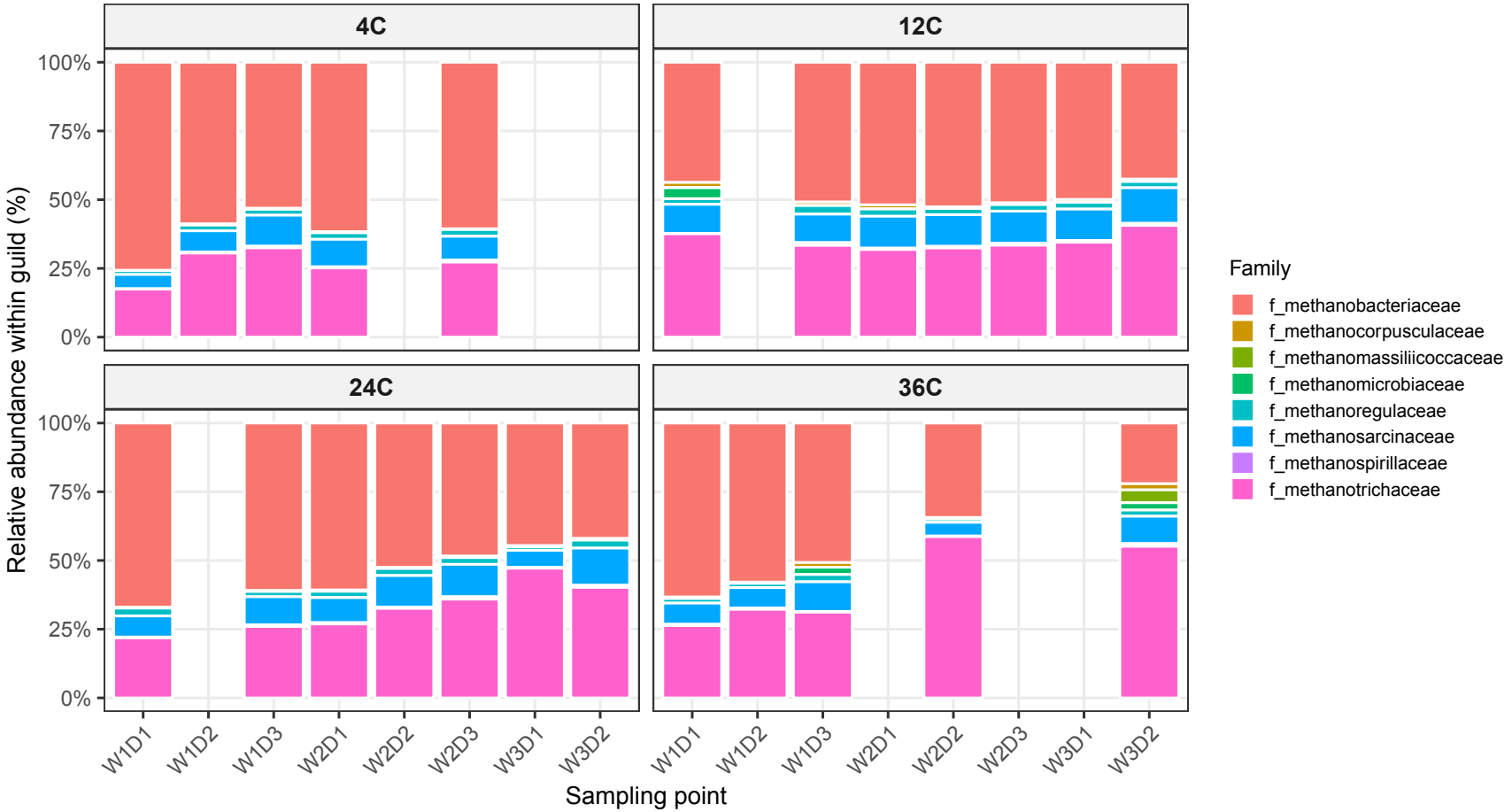

Figure 5.14: Relative abundance of methanogenic archaeal families across sampling points at four incubation temperatures (4 °C, 12 °C, 24 °C, and 36 °C). Stacked bar plots showing the relative abundance (%) of methanogenic families. Each bar represents a single sampling point.

5.5.4 Relative abundance of SAOB families

The composition of SAOB communities showed clear variation with temperature, with Clostridiaceae consistently representing the dominant family across all conditions (Figure 5.15).

At 4 °C, the SAOB community was dominated almost entirely by Clostridiaceae, with minor contributions from other taxa. Temporal variation was limited.

At 12 °C, additional SAOB families including Sporomusaceae contributed more substantially, resulting in a more distributed composition. The relative abundance of Clostridiaceae decreased compared to 4 °C, while Peptococcaceae and Sporomusaceae increased modestly over time.

At 24 °C, SAOB composition became more heterogeneous, with several families contributing substantially. Temporal variation was evident, with shifts in the relative dominance of different taxa across sampling points.

At 36 °C, SAOB communities showed increased temporal variability, with multiple families contributing at varying proportions. Importantly, while Clostridiaceae remained a consistent component of the SAOB guild, its relative dominance was not maintained uniformly across all sampling points. In several mid-incubation samples, Clostridiaceae relative abundance declined substantially, with Sporomusaceae and Peptococcaceae becoming co-dominant. This pattern of intermittent Clostridiaceae dominance, interspersed with periods of more even family distribution, represents a key distinguishing feature of SAOB dynamics at 36 °C relative to lower temperatures.

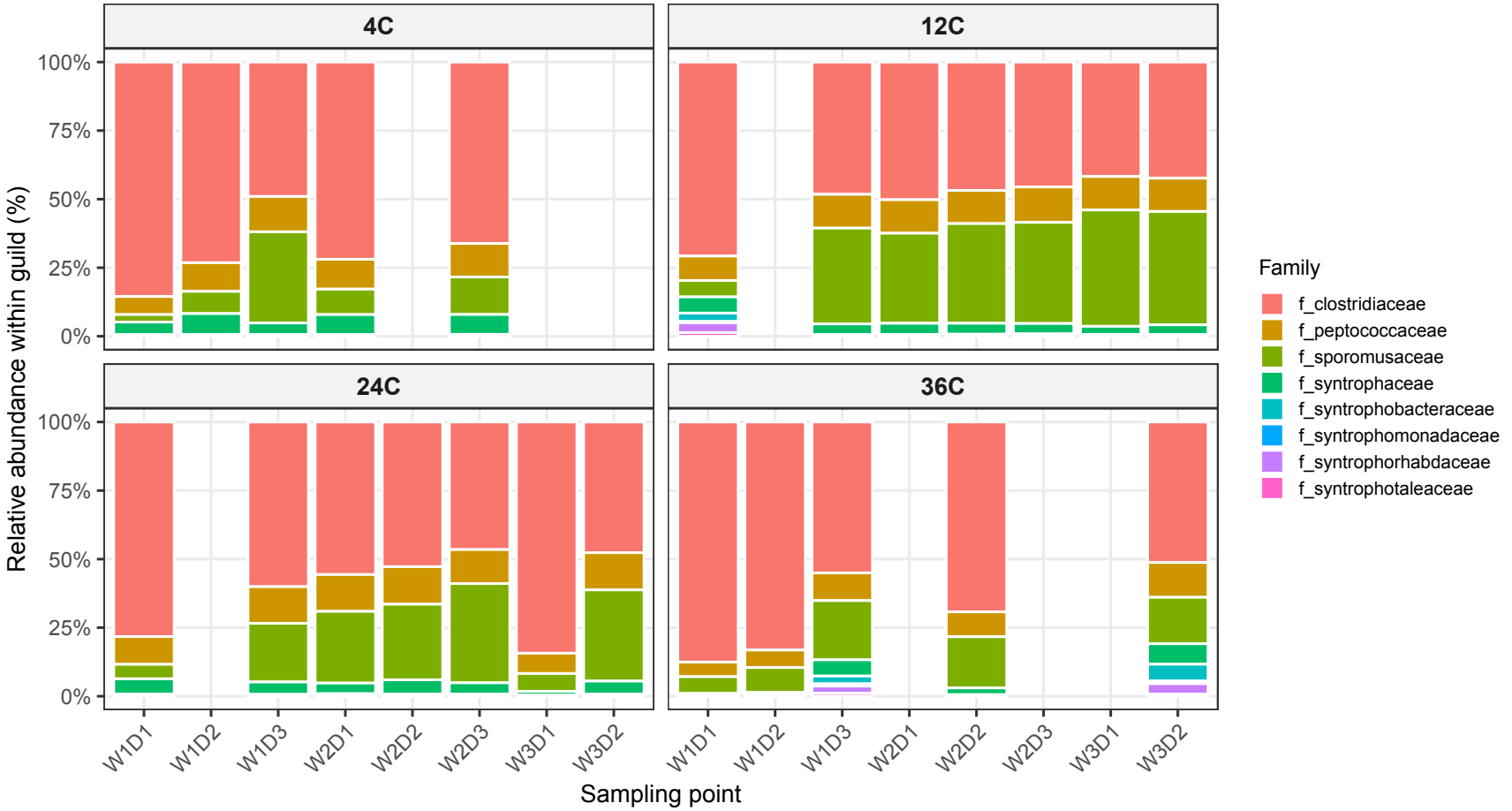

Figure 5.15: Relative abundance of syntrophic acetate-oxidising bacterial families across sampling points at four incubation temperatures (4 °C, 12 °C, 24 °C, and 36 °C). Stacked bar plots showing the relative abundance (%) of SAOB-associated families. Each bar represents a single sampling point.

5.5.5 Relative abundance of SRB families

The composition of SRB families varied markedly with temperature and incubation time, with a clear directional shift in dominant taxa across the temperature gradient (Figure 5.16).

At 4 °C, the SRB community was dominated by Desulfobulbaceae, with relatively high contributions from Desulfobacteraceae and low contributions from other families. Temporal variation was limited.

At 12 °C, SRB diversity increased, with multiple families contributing to the total. Desulfobulbaceae remained prominent, but other families, including Desulfomicrobiaceae, showed increasing relative contributions over time.

At 24 °C, a further shift in SRB composition was observed, with Desulfomicrobiaceae increasing substantially and Desulfobulbaceae remaining a prominent contributor. The relative abundance of Desulfobacteraceae declined relative to lower temperatures. Temporal trends indicated progressive increases in specific families across the incubation period.

At 36 °C, the SRB community underwent pronounced compositional change relative to lower temperatures. Desulfomicrobiaceae and Desulfobulbaceae remained key contributors, while Desulfobacteraceae – which had been prominent at 4 °C – was largely absent or at very low relative abundance. Temporal variability was high, with shifts in the relative proportions of these families across successive sampling points. The transition from Desulfobacteraceae dominance at 4 °C to Desulfomicrobiaceae and Desulfobulbaceae dominance at 36 °C represents a consistent and directional temperature-driven restructuring of the SRB guild across the experimental gradient.

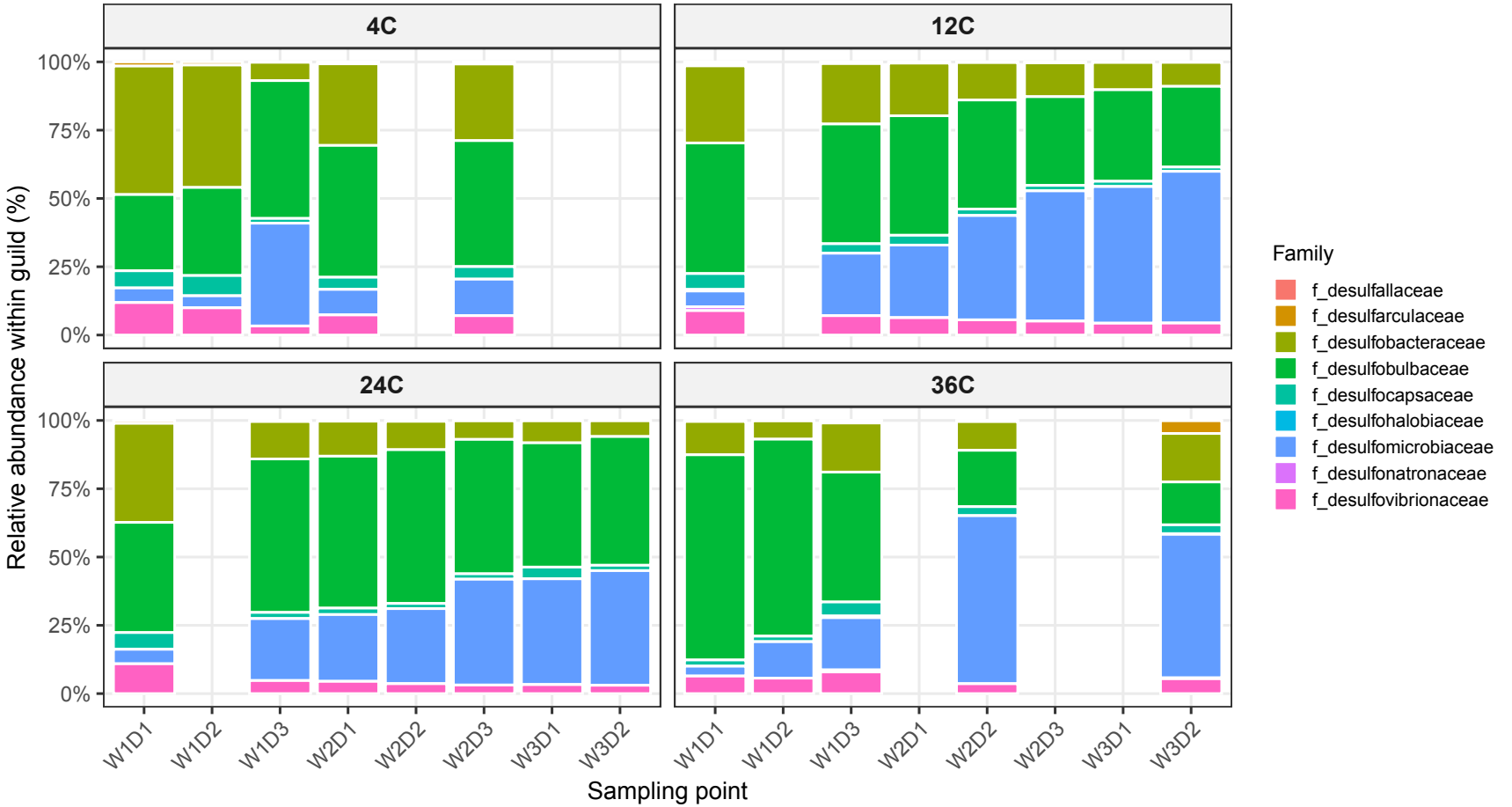

Figure 5.16: Relative abundance of sulphate-reducing bacterial families across sampling points at four incubation temperatures (4 °C, 12 °C, 24 °C, and 36 °C). Stacked bar plots showing the relative abundance (%) of SRB families. Each bar represents a single sampling point.

5.5.6 Relative abundance of viral families

Viral community composition was highly consistent across all temperatures and sampling points (Figure 5.17). At 4 °C, 12 °C, 24 °C, and 36 °C, the community was dominated by Siphoviridae, which comprised the majority of relative abundance in all samples. Podoviridae and Myoviridae were present as secondary components with relatively stable contributions across temperatures and time.

No clear temperature-dependent shift in dominant viral families was observed. Temporal variation within each temperature was limited, with only minor fluctuations in the relative proportions of Podoviridae and Myoviridae. The overall viral community structure at the family level therefore remained stable throughout the three-week incubation period across all temperature conditions, contrasting markedly with the pronounced temperature-driven restructuring observed in the prokaryotic community. Less abundant viral families – including Ackermannviridae, Autographiviridae, Demecriviridae, and Herelleviridae – contributed negligibly to the total viral community under most conditions, though as described in Section 5.6, these families showed notable correlation patterns at 36 °C that are not reflected in their low relative abundance values.

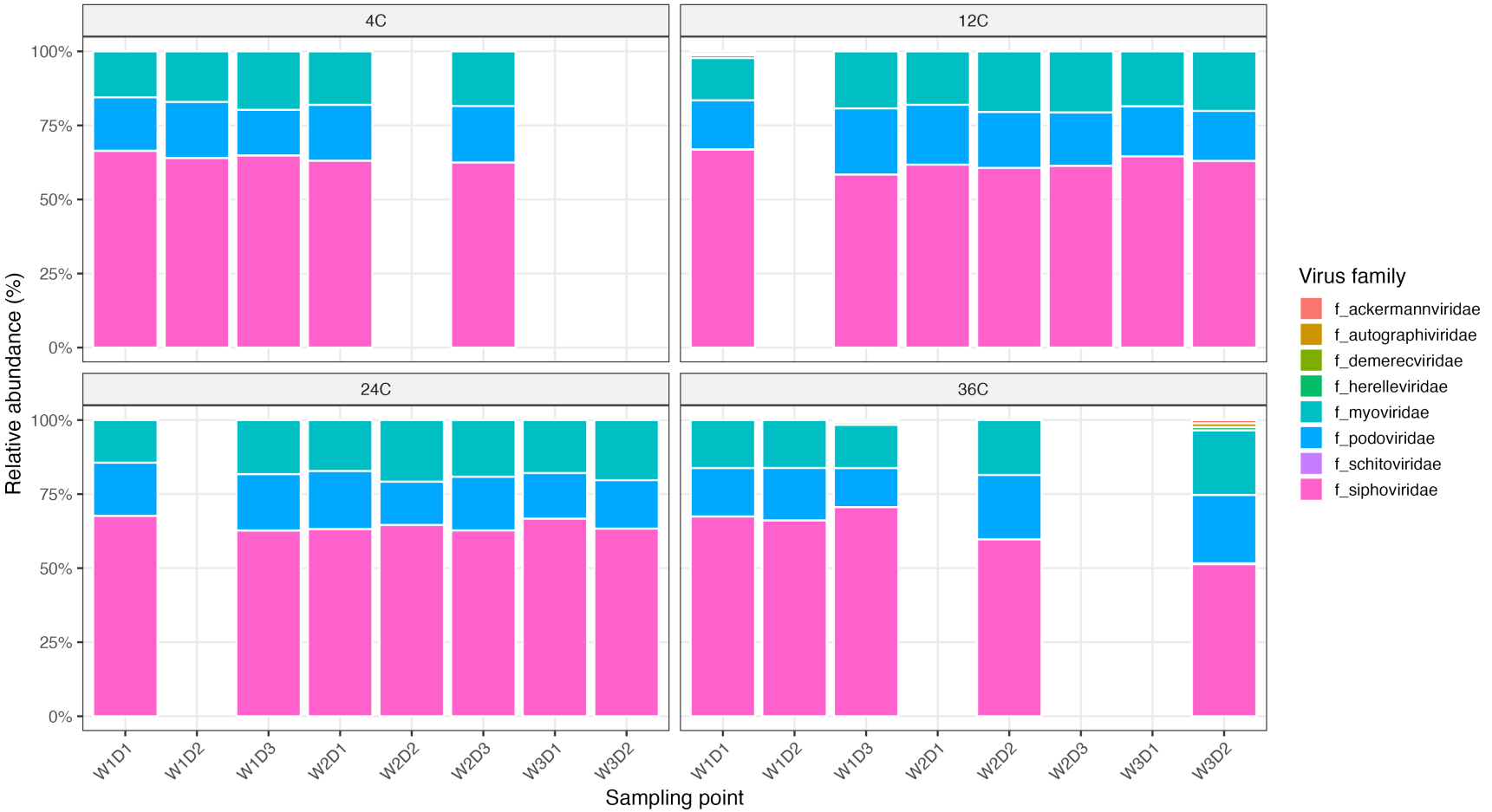

Figure 5.17: Relative abundance of phage families across sampling points at four incubation temperatures (4 °C, 12 °C, 24 °C, and 36 °C). Stacked bar plots showing the relative abundance (%) of phage families. Each bar represents a single sampling point.

5.5.7 Implications

The observed patterns indicate that temperature is associated with a systematic redistribution of AD functional guild composition over the incubation period, while viral community composition at the family level remains comparatively stable. The progressive increase in the relative abundance of SRB – and, to a lesser extent, SAOB – at 24 °C and 36 °C indicates that these guilds contribute more substantially to the overall community under higher temperature conditions, consistent with the enhanced metabolic rates and increased substrate competition expected under mesophilic conditions. The persistently high relative abundance of fermenters across all temperatures confirms that this guild is the dominant and most thermally resilient community component.

The consistently low relative abundance of methanogens across all temperatures indicates that this guild constitutes a minor fraction of the total microbial community under the experimental conditions. However, the directional increase in methanogenic family diversity and the increased relative contribution of Methanotrichaceae and Methanosarcinaceae at higher temperatures suggest that methanogenic community structure is sensitive to temperature even where overall methanogenic abundance remains low. The observed transition from Methanobacteriaceae dominance at low temperatures to a more diverse methanogenic assemblage at 36 °C has implications for methanogenic pathway diversity and functional redundancy under mesophilic conditions.

Within individual guilds, the direction and nature of family-level composition change differed across guilds. For SRB, the shift from Desulfobacteraceae dominance at 4 °C to Desulfomicrobiaceae dominance at later stages of incubation at 36 °C represents a clear and directional temperature response. For SAOB, the intermittent decline of Clostridiaceae dominance at 36 °C, replaced by more even contributions from Sporomusaceae and Peptococcaceae, suggests a less stable guild structure under upper mesophilic conditions. For fermenters and methanogens, compositional changes were more gradual and less directionally consistent.

In contrast to microbial communities, viral community composition remained consistent across all temperatures and sampling points at the family level. The persistent dominance of Siphoviridae, together with stable contributions from Podoviridae and Myoviridae, indicates that viral family-level composition is not strongly restructured by

temperature under the tested conditions, despite the pronounced changes in prokaryotic guild composition occurring simultaneously.

5.6 Temperature-dependent correlations between anaerobic microbial guilds, viral communities, and physicochemical parameters

5.6.1 Phage–host correlation patterns in laboratory-scale anaerobic reactors

4 °C

At 4 °C, the correlation structure was sparse and weakly organised, with only a limited number of statistically significant associations observed across the three-week incubation period (Figure 5.18). Significant correlations were concentrated among a small subset of viral families – primarily Myoviridae and Podoviridae – and selected fermentative and SRB taxa. Fermenters displayed mixed correlation patterns with alternating positive and negative associations depending on sampling timepoint, while methanogens exhibited minimal and weakly structured interactions with viral families. SAOB correlations were largely absent, and SRB associations were limited to a few isolated significant links. Overall, the 4 °C system was characterised by decoupled viral–microbial dynamics, reflecting reduced metabolic activity and weak ecological connectivity under psychrophilic conditions.

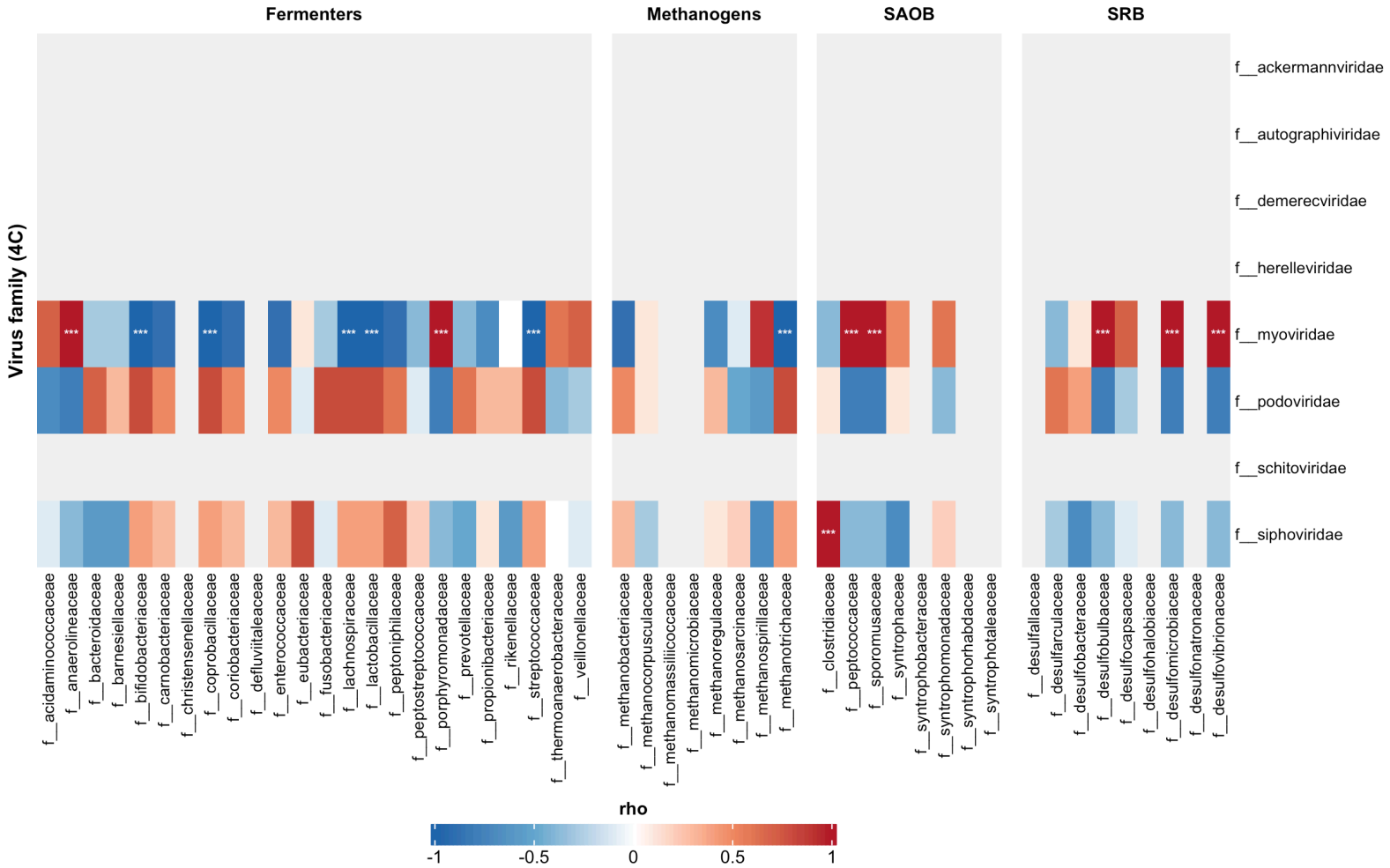

Figure 5.18: Spearman correlations between phage families and prokaryotic functional guilds at 4 °C. Heatmap showing Spearman correlation coefficients between phage families (rows) and prokaryotic families (columns). Colour scale indicates the direction and strength of correlations (–1 to +1). FDR-corrected significance levels: $q \leq 0.05$ (*), $q \leq 0.01$ (**), and $q \leq 0.001$ (***).

12 °C

At 12 °C, the correlation matrix exhibited a marked increase in both the number and strength of statistically significant associations, representing the highest degree of structured phage–host connectivity across the experimental gradient (Figure 5.19). Strong and consistent correlations emerged across all major viral families – including Podoviridae, Siphoviridae, Myoviridae, and Schitoviridae – and multiple microbial guilds. Fermentative bacteria showed widespread and repeated significant correlations, predominantly negative in direction, consistent across all sampling days, indicating stable and tightly coupled viral–host interactions. Methanogens, SAOB, and SRB displayed no statistically significant correlations with phages. A notable feature of the 12 °C heatmap was the presence of repeated, vertically aligned correlation patterns across viral families, indicating coherent temporal dynamics and synchronised shifts in community composition. Overall, 12 °C represents a transitional regime characterised by maximal interaction complexity, where microbial activity is sufficiently high to sustain robust phage–host coupling, yet has not stabilised into the more uniform structure observed at higher temperatures.

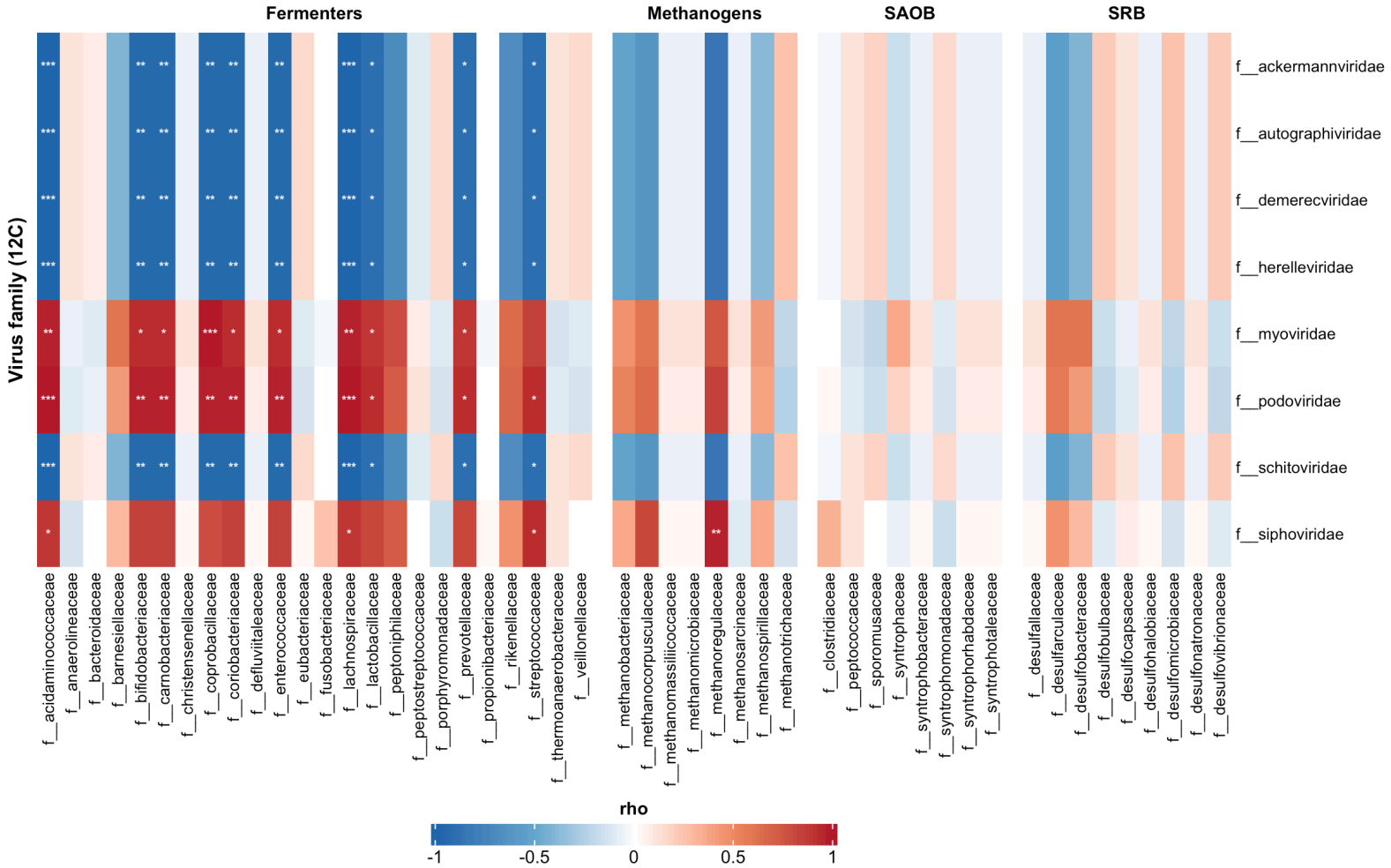

Figure 5.19: Spearman correlations between phage families and prokaryotic functional guilds at 12 °C. Heatmap showing Spearman correlation coefficients between phage families (rows) and prokaryotic families (columns). Colour scale indicates the direction and strength of correlations (–1 to +1). FDR-corrected significance levels: $q \leq 0.05$ (*), $q \leq 0.01$ (**), and $q \leq 0.001$ (***).

24 °C

At 24 °C, the correlation structure became less complex and more diffuse compared to 12 °C (Figure 5.20). Although structured patterns were visually present, no associations reached statistical significance following FDR correction. Fermentative bacteria exhibited heterogeneous correlation patterns, with alternating positive and negative associations that were not consistently significant. Methanogenic families showed moderate correlations, but these lacked statistical robustness and temporal consistency. SAOB correlations were limited and largely non-significant, indicating reduced coupling with viral dynamics. SRB displayed visually clear patterns, particularly involving Myoviridae and Podoviridae, but these associations did not persist following multiple testing correction. The reduction in statistically significant associations suggests that, under mesophilic conditions, the system transitions towards functional stability, where microbial processes are less tightly constrained by physicochemical gradients, and phage–host interactions become less synchronised at the community level.

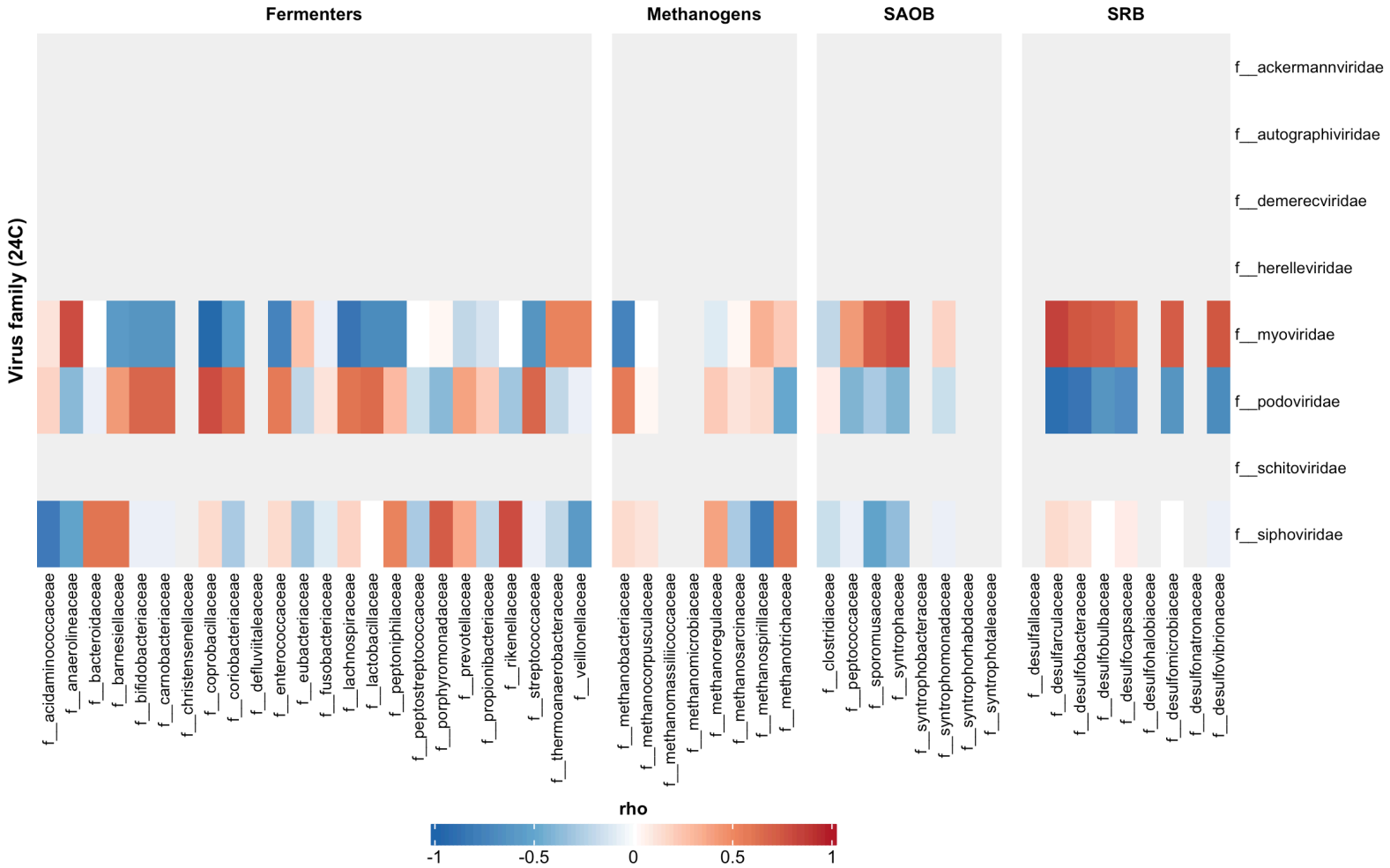

Figure 5.20: Spearman correlations between phage families and prokaryotic functional guilds at 24 °C. Heatmap showing Spearman correlation coefficients between phage families (rows) and prokaryotic families (columns). Colour scale indicates the direction and strength of correlations (–1 to +1). FDR-corrected significance levels: $q \leq 0.05$ (*), $q \leq 0.01$ (**), and $q \leq 0.001$ (***).

36 °C

At 36 °C, the correlation matrix exhibited a highly structured and strongly polarised pattern, characterised by extensive statistically significant associations across nearly all viral families and microbial guilds (Figure 5.21). Several key features distinguished this temperature condition. Strong and consistent correlations were observed between Myoviridae, Podoviridae, and multiple fermentative, methanogenic, and SRB families. Methanogens exhibited clear and repeated significant associations, suggesting tight coupling between viral dynamics and methanogenic pathways under upper mesophilic conditions. SAOB and SRB showed well-defined and highly significant correlation structures, indicating strong integration into the overall microbial–viral interaction network. Fermenter–phage associations were less prevalent at 36 °C than at 12 °C, where such correlations were most pronounced.

Importantly, correlation patterns were highly consistent across sampling days, reflecting temporal stability and reproducibility of interactions at elevated temperature. Overall, the 36 °C system represents a highly connected and strongly structured ecological network, indicative of enhanced metabolic rates, rapid turnover, and intensified phage–host interactions under upper mesophilic conditions.

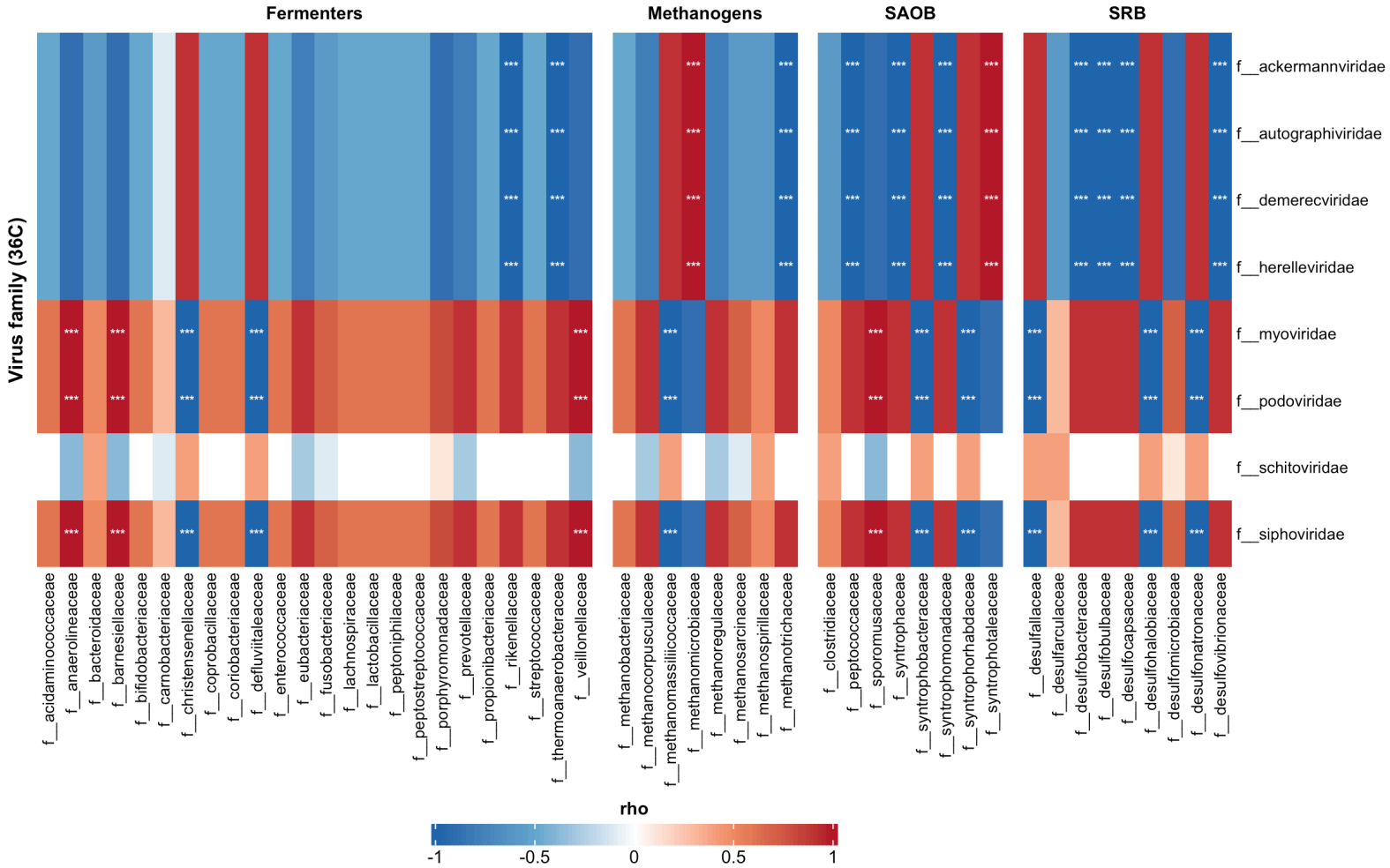

Figure 5.21: Spearman correlations between phage families and prokaryotic functional guilds at 36 °C. Heatmap showing Spearman correlation coefficients between phage families (rows) and prokaryotic families (columns). Colour scale indicates the direction and strength of correlations (–1 to +1). FDR-corrected significance levels: $q \leq 0.05$ (*), $q \leq 0.01$ (**), and $q \leq 0.001$ (***).

5.6.2 Temperature-dependent correlations between physicochemical parameters and microbial and viral communities

Spearman correlations between physicochemical parameters and the relative abundances of AD functional guilds and viral families revealed strong temperature-dependent structuring of abiotic-biotic relationships (Figures 5.22–5.23). Statistically significant correlations were unevenly distributed across the temperature gradient, with the highest number of robust associations observed at 24 °C, followed by 36 °C, whilst 4 °C and 12 °C exhibited limited or no statistically significant relationships following FDR correction.

4 °C

At 4 °C, correlations between physicochemical parameters and AD microbial families were weak and inconsistent, with no statistically significant associations detected following FDR correction. Although some visual patterns were apparent – particularly involving solids (TS, VS) and fermentative families – these did not persist after correction for multiple testing. Overall, these results indicate that under psychrophilic conditions, physicochemical gradients exert limited detectable influence on either microbial or viral community structure, consistent with reduced metabolic rates and constrained system responsiveness.

12 °C

At 12 °C, correlations remained largely non-significant following FDR correction, although more structured patterns began to emerge compared to 4 °C. Several fermentative and methanogenic families displayed negative correlations with TS and VS, suggesting reduced abundance under higher solids conditions. Viral communities showed a distinct pattern at this temperature: Ackermannviridae, Autographiviridae, Demecriviridae, and Herelleviridae exhibited a strong positive correlation with total and volatile solids, while Myoviridae, Podoviridae, and Siphoviridae showed a negative association with solids. However, as none of these associations remained significant following multiple testing correction, the 12 °C system is interpreted as a transitional regime in which physicochemical controls begin to emerge but remain statistically unresolved.

24 °C

At 24 °C, the correlation structure became highly organised and statistically robust, with multiple significant associations observed across microbial guilds and physicochemical parameters. A dominant feature of this temperature was the strong and consistent negative correlation between physicochemical parameters and a broad range of AD microbial families – particularly SRB, including Desulfobacteraceae, Desulfobulbaceae, and Desulfovibrionaceae, as well as syntrophic Clostridiaceae. These relationships were statistically significant across multiple parameters ($q \leq 0.05$), indicating that higher solids concentrations were associated with reduced relative abundance of key anaerobic functional groups, potentially through mass transfer constraints or reduced substrate accessibility. Conversely, several fermentative taxa and methanogens, including Methanobacteriaceae, showed positive correlations with COD and BOD_5, reflecting association with higher biodegradable substrate availability and suggesting that organic load promotes microbial activity at this temperature. In contrast, viral correlations at 24 °C were simplified and uniform. Negative correlations were observed between Myoviridae and COD/BOD_5 and positive correlations between Podoviridae and COD/BOD_5, but without statistically significant associations. This indicates that, despite strong microbial–environment coupling, viral dynamics remain comparatively decoupled from physicochemical gradients at this temperature.

36 °C

At 36 °C, the correlation structure remained highly organised but differed substantively from 24 °C. Only fermentative Carnobacteriaceae showed statistically robust strong positive correlations with COD, BOD_5, and solids, indicating tight coupling between microbial abundance and organic substrate availability under upped mesophilic conditions. Viral communities at 36 °C displayed no statistically robust correlations with physicochemical parameters.

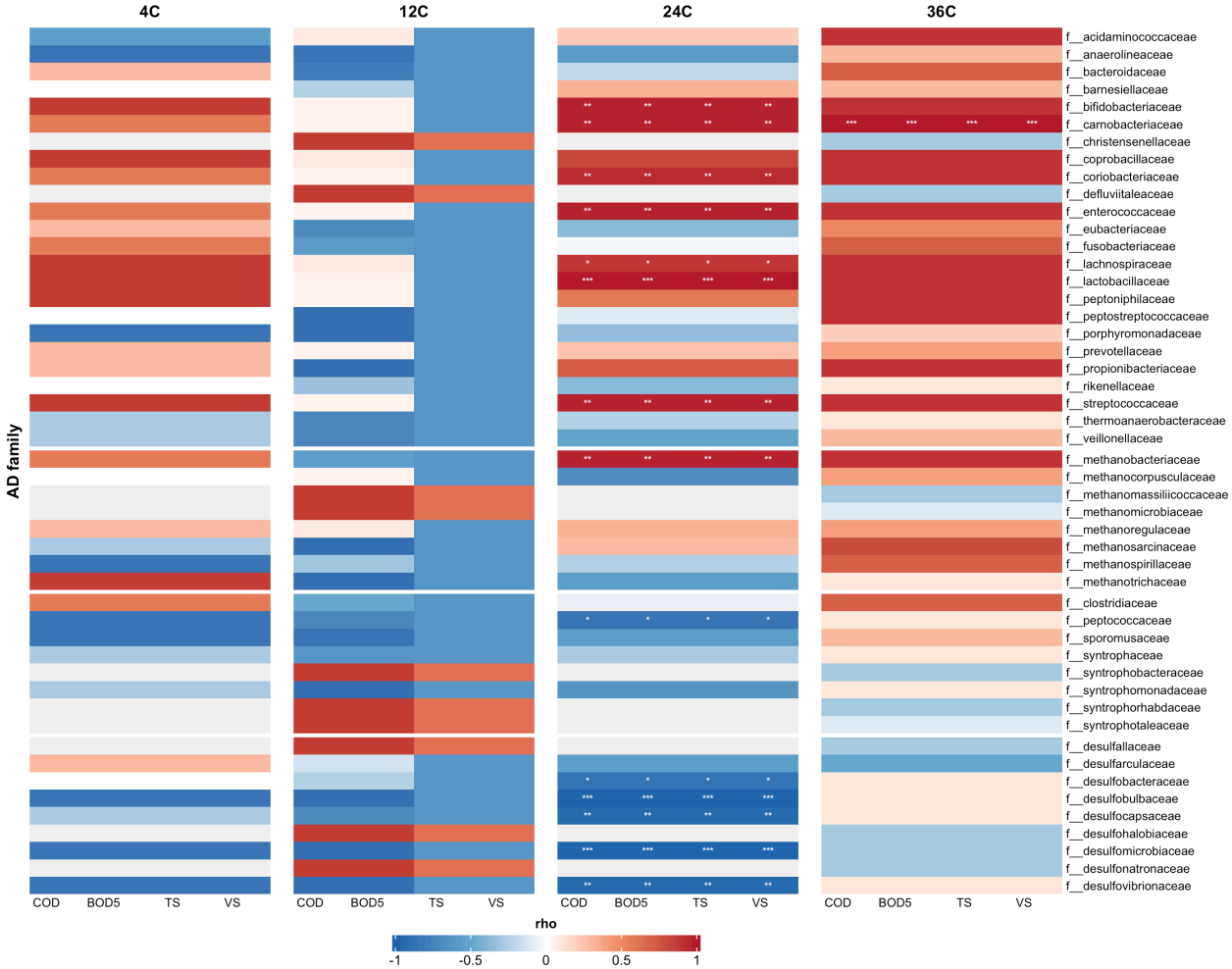

Figure 5.22: Spearman correlations between physicochemical parameters and prokaryotic functional guilds across temperatures. Heatmap showing Spearman correlation coefficients between physicochemical parameters (COD, BOD_5, TS, VS) and prokaryotic families across 4 °C, 12 °C, 24 °C, and 36 °C. Colour scale indicates the direction and strength of correlations (–1 to +1). FDR-corrected significance levels: $q \leq 0.05$ (*), $q \leq 0.01$ (**), and $q \leq 0.001$ (***).

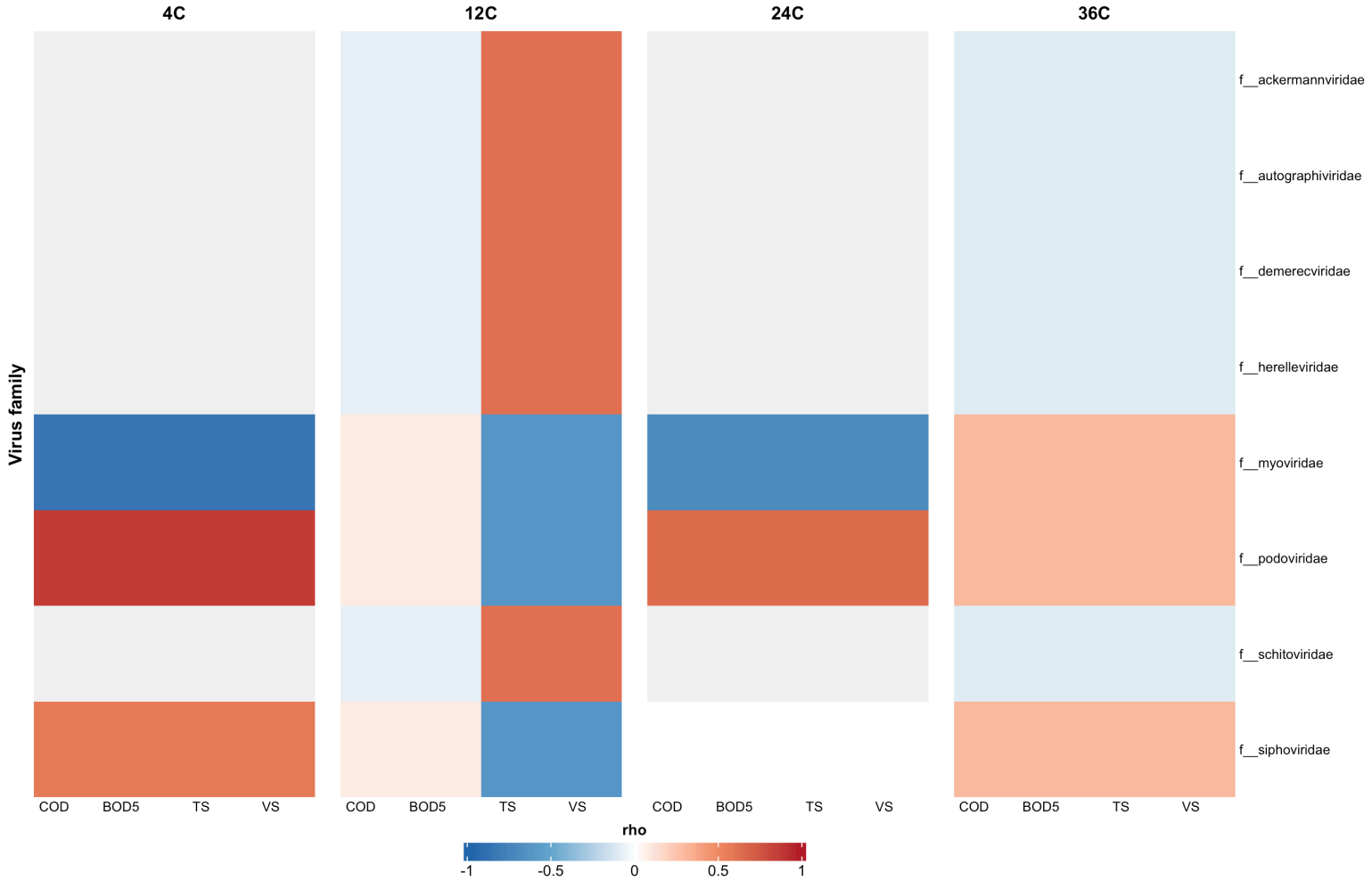

Figure 5.23: Spearman correlations between physicochemical parameters and phage families across temperatures. Heatmap showing Spearman correlation coefficients between physicochemical parameters (COD, BOD_5, TS, VS) and phage families across 4 °C, 12 °C, 24 °C, and 36 °C. Colour scale indicates the direction and strength of correlations (–1 to +1). FDR-corrected significance levels: $q \leq 0.05$ (*), $q \leq 0.01$ (**), and $q \leq 0.001$ (***).

5.6.3 Implications

In the laboratory batch reactor system, correlations between microbial and viral abundances and physicochemical parameters provide a functionally meaningful proxy for performance–community linkages, even where removal efficiency was not calculated as an explicit output variable. In a closed batch system, COD and BOD_5 concentrations decline progressively over the incubation period as organic matter is degraded, and TS and VS concentrations decline as particulate matter is hydrolysed and mineralised. Correlations between these parameters and guild abundance therefore reflect the relationship between the current state of treatment and the microbial community responsible for driving it, which is functionally equivalent to a performance-community correlation in this experimental context.

At 4 °C and 12 °C, the near-complete absence of statistically significant correlations following FDR correction indicates that under cold conditions the relationship between microbial community composition and physicochemical parameters was effectively decoupled, consistent with kinetically constrained metabolic rates insufficient to generate detectable co-variation between community structure and treatment state within the three-week incubation period. This decoupling does not indicate the absence of biological activity – measurable reductions in COD, BOD_5, TS, and VS were observed at both temperatures – but rather that the magnitude and rate of biological processing were insufficient to produce the degree of community–environment co-variation detectable by Spearman correlation analysis following FDR correction.

The absence of statistically significant correlations at 4 °C and 12 °C indicates that, under low-temperature conditions, measurable associations between these components are minimal or undetectable using correlation-based approaches.

At 24 °C, strong and statistically significant negative correlations were observed between solids-related parameters (TS and VS) and multiple SRB families including Desulfobacteraceae, Desulfobulbaceae, and Desulfovibrionaceae, as well as SAOB family Clostridiaceae ($q < 0.05$–0.001). This indicates that downstream processing guilds responsible for acetate oxidation and sulphate reduction were least abundant precisely when solids concentrations were highest – that is, when treatment efficiency was lowest – confirming that solids accumulation at mesophilic temperatures acts as a direct inhibitor

of the guilds most critical to organic matter mineralisation beyond the fermentation stage. At the same temperature, several fermentative families including Carnobacteriaceae, Coriobacteriaceae, and Lactobacillaceae showed significant positive correlations with COD and BOD_5 ($q < 0.01$–0.001), indicating that fermentative guild abundance was highest when biodegradable substrate concentrations were highest. This is consistent with fermenters responding to substrate availability rather than driving its removal and confirms the functional bottleneck at the hydrolysis–fermentation stage under high organic loading conditions at this temperature.

At 36 °C, the correlation structure shifted substantially: positive correlations between COD and BOD_5 and fermentative Carnobacteriaceae ($q < 0.01$) indicated substrate-driven guild expansion under conditions of adequate metabolic activity.

Viral communities showed a parallel but temporally offset pattern that is not statistically robust across the temperature gradient. These correlation patterns demonstrate that microbial guild abundance is systematically linked to treatment state in a temperature-dependent manner, a mechanistic pattern that provides direct microbiological explanation for the temperature-dependent performance variation documented across the experimental gradient.

5.7 Temperature-dependent variation in key anaerobic metabolic pathways

The relative abundance patterns of key anaerobic metabolic marker genes exhibited clear and systematic variation across the four temperature regimes (4 °C, 12 °C, 24 °C, and 36 °C), reflecting strong temperature control over functional pathway activity (Figure 5.24). Overall, the data indicate a progressive shift from selectively enriched but largely suppressed metabolic activity under psychrophilic conditions towards enhanced and more diversified pathway expression under mesophilic conditions, though the trajectories of individual genes differed substantially.

5.7.1 Acetate activation

Markers associated with acetate activation (*ackA* and *pta*) displayed contrasting and complementary responses to temperature. The *ackA* gene exhibited depletion at 4 °C and 12 °C, and high enrichment at 24 °C and 36 °C, indicating a progressive enhancement of acetate kinase-mediated acetate activation with increasing temperature. In contrast, *pta* showed strong enrichment at 4 °C, declining to near neutral at 12 °C, strong depletion at 24 °C, and strong enrichment at 36 °C, producing a pronounced U-shaped temperature response. This divergence suggests differential regulation of the two arms of the acetate activation pathway across the temperature gradient, potentially reflecting shifts between alternative acetate utilisation mechanisms or between the microbial taxa dominating under distinct thermal regimes. At the lowest temperature, *pta*-mediated phosphotransacetylase activity appears to represent the primary route of acetate activation, while at mesophilic temperatures *ackA*-mediated activity becomes increasingly dominant with both pathways highly active at 36 °C.

5.7.2 Syntrophic acetate oxidation / Wood–Ljungdahl pathway

Markers associated with syntrophic acetate oxidation and the Wood–Ljungdahl pathway demonstrated heterogeneous but structured responses reflecting distinct thermal optima among pathway components. K00194 was enriched at 4 °C, depleted at 12 °C, near neutral at 24 °C, and enriched again at 36 °C, suggesting a bimodal temperature response with activity at both low and upper mesophilic conditions. K00198 showed moderate enrichment at 4 °C, 12 °C and 24 °C but strong depletion at 36 °C, indicating optimal activity at intermediate temperatures and suppression under upper mesophilic conditions. K01491 was strongly enriched at 4 °C and progressively depleted with increasing temperature, reaching strong depletion at 36 °C, indicating a psychrophilic or cold-favoured profile for this marker. K03388 showed strong depletion at 4 °C, enrichment at 12 °C, a moderate decline at 24 °C, and strong enrichment at 36 °C – a pattern that does not follow a simple gradient but instead shows a local minimum at 24 °C before recovering at upper mesophilic temperature.

These contrasting trends confirm that the SAO/WL functional guild is not homogeneous but comprises multiple pathways or taxa with distinct thermal optima,

and that SAO-related processes may operate under both cold and warm conditions via different metabolic routes or microbial consortia.

5.7.3 Sulphate reduction

Sulphate-reducing pathway markers (*dsrA* and *dsrB*) showed a clear and directional temperature-dependent enhancement, though the two marker genes differed in their patterns at 36 °C. The *dsrA* gene exhibited strong depletion at 4 °C, near neutral abundance at 12 °C, strong enrichment at 24 °C, and sustained strong enrichment at 36 °C, indicating a consistent increase in sulphate reduction activity from psychrophilic to mesophilic conditions. The *dsrB* gene showed a similar pattern at 4 °C through 24 °C – strong depletion at 4 °C, near-neutral at 12 °C, and strong enrichment at 24 °C – but differed at 36 °C, where *dsrB* returned to near-neutral abundance rather that remaining strongly enriched. The divergence between *dsrA* and *dsrB* at 36 °C indicates that these two subunits of the dissimilatory sulphite reductase complex do not follow identical temperature trajectories under upper mesophilic conditions, which may reflect differential regulation, or gene copy number variation among SRB taxa.

5.7.4 Core methanogenesis

The methanogenesis marker gene *mcrA* displayed a distinct non-monotonic response across the temperature gradient. Relative abundance was strongly depleted at 4 °C, increased to moderate enrichment at 12 °C, declined to clear depletion at 24 °C, and reached strong enrichment at 36 °C. The pronounced decline at 24 °C may reflect transitional dynamics between competing terminal electron-accepting pathways, particularly sulphate reduction, which also peaked at 24 °C and may outcompete methanogenesis for shared substrates such as hydrogen and acetate at this temperature. The subsequent strong recovery of *mcrA* at 36 °C, despite continuing strong sulphate reduction, suggests that substrate availability is sufficient at upper mesophilic temperatures to support both pathways simultaneously, and that distinct methanogenic niches – particularly hydrogenotrophic methanogenesis – remain thermodynamically favourable despite competing pressure.

5.7.5 Implications

Taken together, these results demonstrate a clear temperature-driven reorganisation of anaerobic metabolic pathways. At 4 °C, metabolic activity is generally reduced, with selective enrichment of *pta*-mediated acetate activation, specific SAO/WL pathway components, particularly K00194 and K01491, and near complete suppression of sulphate reduction and methanogenesis. This suggests that psychrophilic conditions favour a fermentation-dominated metabolism with limited downstream processing at 12 °C, pathway activity diversifies, with moderate enrichment of SAO/WL markers and the beginnings of SRB activity, while methanogenesis recovers partially from its 4 °C minimum. At 24 °C, sulphate reduction reaches high activity while methanogenesis is simultaneously suppressed, suggesting competitive exclusion or substrate partitioning between these terminal pathways. At 36 °C, a shift toward comprehensive metabolic activation occurs, with high activity across acetate activation, sulphate reduction, and methanogenesis, though with notable differences between *dsrA* and *dsrB* and between individual SAO/WL markers suggesting ongoing pathway-level heterogeneity even under optimal thermal conditions.

5.8 Pathogen-associated community composition across temperatures

The relative abundance of pathogen-associated bacterial families varied across the temperature gradient in ways that reflect both the thermal preferences of individual taxa and the overall shift in community structure with increasing temperature (Figure 5.25).

At 4 °C, the pathogen-associated community was dominated by Brucellaceae, which comprised approximately 70% of total relative abundance within the pathogen community. Pseudomonadaceae was present as a secondary component, with minor contributions from Enterobacteriaceae, Moraxellaceae, and Campylobacteraceae. The overall community structure was highly skewed toward a single dominant family under psychrophilic conditions.

At 12 °C, Brucellaceae remained the dominant family, though its relative abundance decreased compared to 4 °C. Pseudomonadaceae increased substantially as a proportion

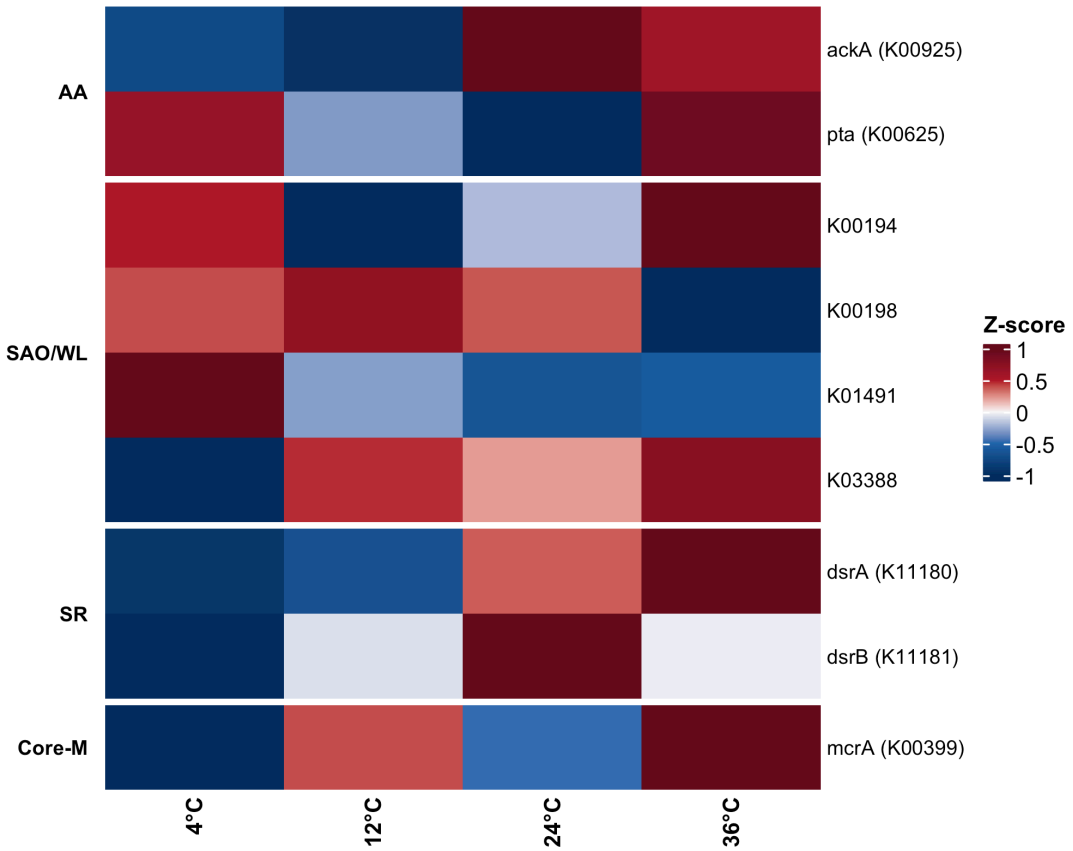

Figure 5.24: Temperature-dependent variation in key functional genes associated with AD pathways. Heatmap showing *Z-score* normalised abundances of selected functional genes across temperature conditions (4 °C, 12 °C, 24 °C, and 36 °C). Genes are grouped by functional category. Colour scale represents relative enrichment (red) or depletion (blue) relative to the cross-temperature mean.

of the community, and additional families contributed small but noticeable fractions, resulting in a more distributed composition while still characterised by single-family dominance.

At 24 °C, the composition shifted substantially away from Brucellaceae dominance. Pseudomonadaceae and Moraxellaceae increased markedly, becoming the dominant families at this temperature. The overall community became more heterogeneous, with multiple families contributing moderate fractions and no single taxon accounting for the majority of relative abundance.

At 36 °C, Pseudomonadaceae was again a substantial component, though at lower relative abundance than at 4 °C, and the community was more evenly distributed across multiple taxa. Pseudomonadaceae and Peptostreptococcaceae increased relative to 4 °C, and the overall composition showed greater evenness across families than at lower temperatures. Across all temperatures, the majority of other pathogen-associated families – including Vibrionaceae, Neisseriaceae, Rickettsiaceae, Helicobacteraceae, and Pasteurellaceae – remained at very low relative abundances.

5.8.1 Implications

The observed patterns indicate that temperature is associated with systematic shifts in the composition of pathogen-associated bacterial families, with the dominant taxon differing across the temperature gradient. Brucellaceae dominance at low temperatures (4 °C and 12 °C) indicates that this family is the most abundant pathogen-associated group under cold conditions, while the shift to Pseudomonadaceae and Moraxellaceae dominance at 24 °C and a more distributed community at 36 °C reflects differential thermal responses among pathogen-associated families. The transition from concentrated single-family dominance at low temperatures to greater compositional evenness at higher temperatures is consistent with the broader pattern of increased community heterogeneity observed across the prokaryotic dataset with increasing temperature. The persistence of Pseudomonadaceae as a consistent secondary contributor across all temperature conditions indicates that this family represents a thermally resilient component of the pathogen-associated community, active across the full range of temperatures tested.

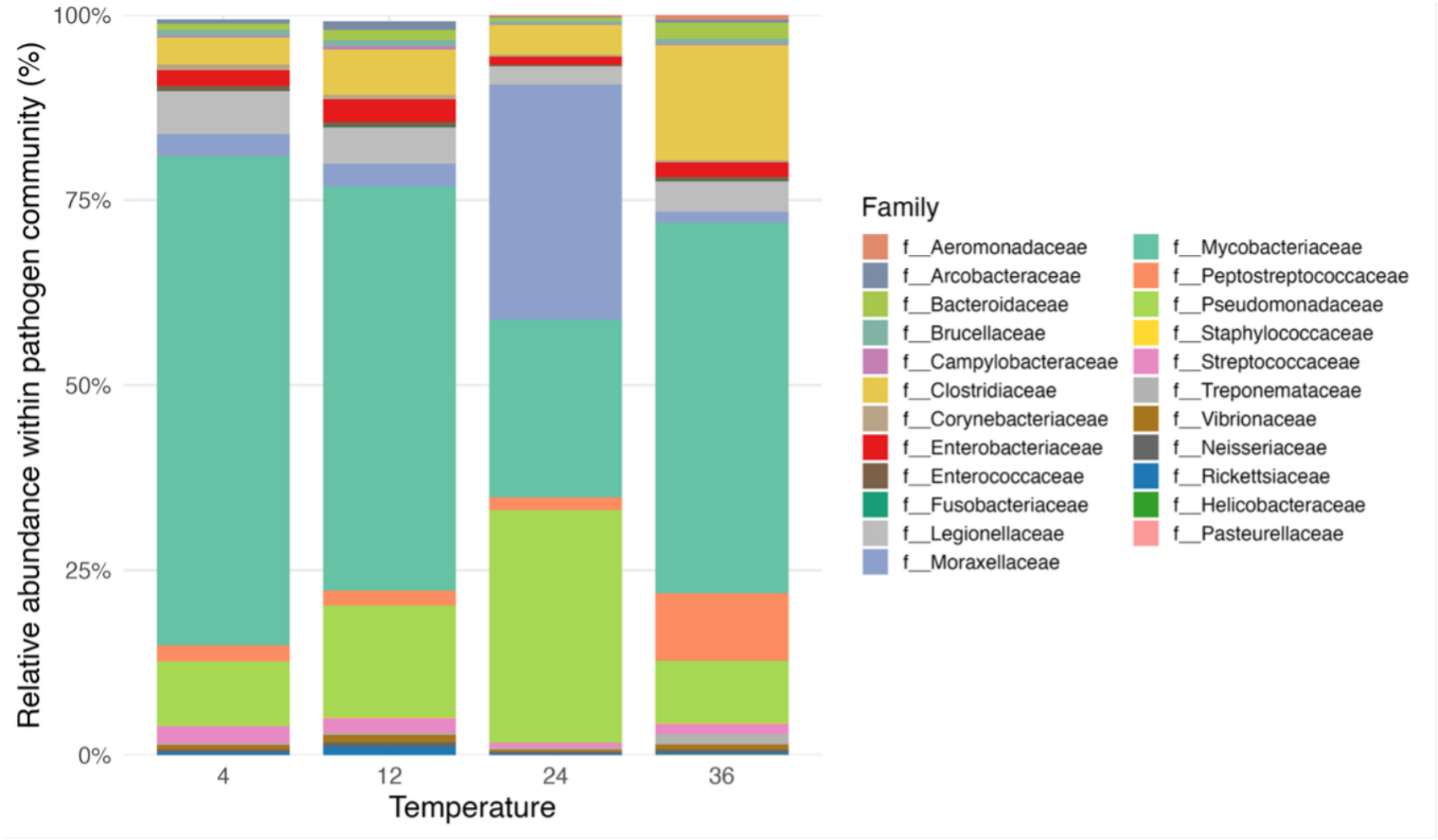

Figure 5.25: Relative abundance of pathogen-associated bacterial families across temperature conditions. Stacked bar plot showing the relative abundance (%) of pathogen-associated bacterial families at 4 °C, 12 °C, 24 °C, and 36 °C. Colours indicate individual families. Relative abundances are normalised to 100% within each temperature condition.

5.9 Summary of key findings

The laboratory-scale anaerobic reactor experiments demonstrated a clear and consistent relationship between temperature and physicochemical treatment performance under controlled conditions. Across all temperature regimes, concentrations of COD, BOD_5, TS, and VS declined over the 21-day incubation period, confirming sustained anaerobic activity throughout the experiment. The observed enhancement of organic matter and solids removal at higher temperatures is consistent with well-established findings from anaerobic digestion studies reporting increased microbial activity and reaction rates under warmer mesophilic conditions (Carballa et al., 2011; Guo et al., 2014; Pussayanavin et al., 2015).

Temperature-dependent differences in microbial community composition structure were observed at multiple levels of organisation. Fermentative bacteria remained the dominant functional group across all temperature conditions. SRB and SAOB increased progressively in relative abundance from 12 °C upward, becoming substantially more prominent at 24 °C and 36 °C. Within guilds, directional shifts in family-level composition were observed: Desulfobacteraceae dominance in SRB at 4 °C transitioned to Desulfobulbaceae dominance at 36 °C, Methanobacteriaceae was progressively supplemented by Methanotrichaceae and Methanosarcinaceae at higher temperatures, and Clostridiaceae dominance in SAOB became intermittent at 36 °C. The pathogen-associated community shifted from Brucellaceae dominance at low temperatures to Moraxellaceae dominance at 24 °C, with greater evenness across families at 36 °C.

Community dissimilarity analyses revealed three distinct structural regimes rather than a linear temperature gradient: stable and homogeneous communities at 4 °C and 24 °C, a transitional and highly variable regime at 12 °C characterised by low median but extreme variance in within-group distances, and consistently high compositional turnover at 36 °C. Prokaryotic and viral beta diversity were strongly concordant (Mantel $r = 0.86$, $p = 0.001$), though this concordance was primarily driven by between-temperature contrasts at the compositional extremes. Viral community composition at the family level remained stable across all temperatures despite pronounced prokaryotic restructuring.

Analysis of functional marker genes revealed temperature-driven reorganisation of anaerobic metabolic pathways. Key patterns included the shifts from *pta*-dominated to

ackA-dominated acetate activation with increasing temperature, the suppression of *mcrA* at 24 °C coinciding with peak sulphate reduction activity, and the heterogeneous thermal optima of individual SAO/WL pathway markers. The divergence between *dsrA* and *dsrB* at 36 °C indicates pathway-level complexity within sulphate reduction not captured by guild-level abundance metrics alone.

Phage–host correlations were more extensive and statistically robust at 12 °C and 36 °C, with near-complete absence of significant associations at 4 °C and reduced connectivity at 24 °C. Physicochemical correlations with microbial community composition were strongest at 24 °C. Viral communities showed no significant physicochemical correlations at any temperature gradient.

The persistence of measurable treatment performance at low temperatures aligns with previous reports of psychrophilic and low-temperature anaerobic systems, where biological processes continue despite substantially reduced kinetics (Luostarinen, 2005; Luostarinen et al., 2007; Zhou et al., 2018), and is consistent with interpretations in the literature suggesting that reduced low-temperature performance is governed by kinetic limitation rather than complete biological inhibition.

Beyond confirming established temperature–performance relationships, the present laboratory experiment extends existing knowledge by systematically evaluating reactor behaviour across a broad thermal range within a single, harmonised experimental framework. Many previous laboratory studies have focused on narrow temperature intervals or single operating regimes, limiting their ability to resolve temperature-driven transitions in anaerobic performance (Luostarinen, 2005; Luostarinen et al., 2007; Garcia et al., 2013; Pussayanavin et al., 2015; Petropoulos et al., 2017). By contrast, the inclusion of cold, transitional, and upper mesophilic conditions in this study provides a continuous reference against which seasonal variability can be assessed.

As such, the laboratory-scale results serve both as validation of prior findings and as a novel, integrative dataset for examining temperature sensitivity across the full range of conditions relevant to septic tank operation in cold-climate regions. This controlled temperature framework provides a valuable foundation for interpreting the field-scale observations presented in subsequent section of this thesis.

6 Results: Physicochemical and microbial dynamics in household septic tanks

Extracts of this chapter have appeared in Paper 2 and Paper 3 (see Section 1.7).

6.1 Overview of dataset and experimental design

This chapter presents the physicochemical, microbial, viral, and phage–host results obtained from ten household septic tanks monitored over a twelve-month period. In total, 74 metagenomic assemblies were generated, including technical and biological replicates; following quality control and composition similarity assessment, these were consolidated into 50 high-quality datasets used for downstream analyses (Table C.1).

The chapter begins with an overview of seasonal patterns in wastewater physicochemical parameters, followed by analyses of microbial, viral, and pathogen-associated community composition and their temporal dynamics. Correlation analyses are then used to explore relationships between abiotic factors, microbial functional guilds, and viral families. Functional potential is examined using unigene-level annotation, where unigenes represent non-redundant gene sequences generated by clustering highly similar predicted genes, alongside KEGG Orthology annotation (Quince et al., 2017; HaploX, 2024). The chapter concludes with iPHoP-derived phage–host predictions that characterise key ecological interactions within the anaerobic digestion system.

6.2 Physicochemical characteristics

Field-scale physicochemical analyses showed clear seasonal variation in wastewater quality across the ten domestic septic tanks monitored between January and December 2024. A complete record of COD, BOD_5, TS, and VS measurements for all ten septic tanks over the full monitoring period is provided in Appendix C (Figures C.2–C.5).

The results presented in this section focus on the subset of 50 samples that met the concentration and quality criteria required for shotgun metagenomic sequencing (Table C.1). Accordingly, the physicochemical data discussed here correspond directly to the samples used for microbial and viral analyses and provide the environmental context for the metagenomic results presented later in this chapter.

6.2.1 pH and temperature buffering

pH remained consistently near-neutral (6.6–7.3) throughout the study period (Figure 6.1). Ambient air temperatures varied widely between −27 °C in January and 31 °C in June; however, septic tank wastewater temperatures exhibited a narrower and buffered range, increasing gradually from 4 °C in winter to 15 °C in summer (Figure 6.2). These observations confirm substantial thermal buffering within septic tanks relative to external climatic conditions.

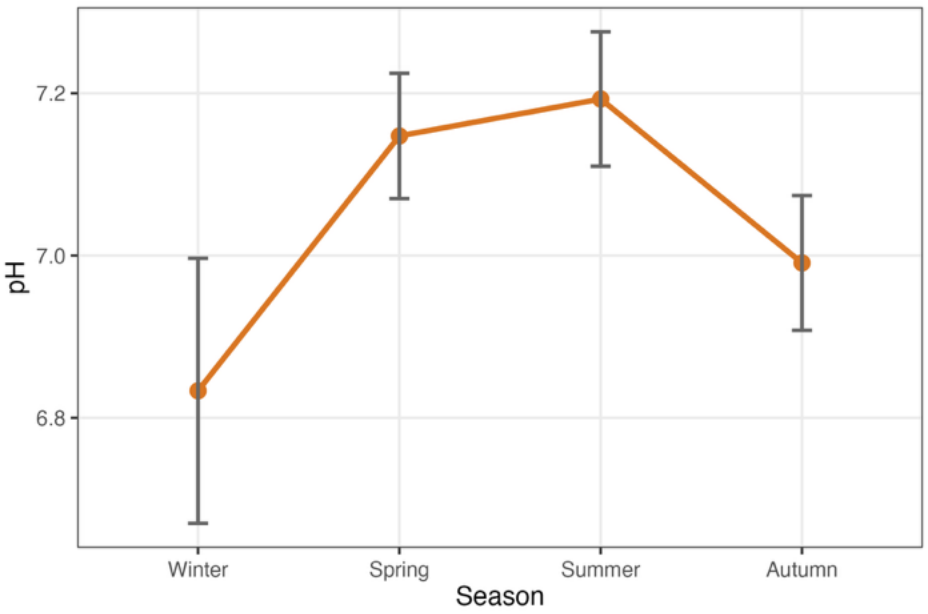

Figure 6.1: Seasonal variation in septic tank wastewater pH (mean ± standard deviation (SD)).

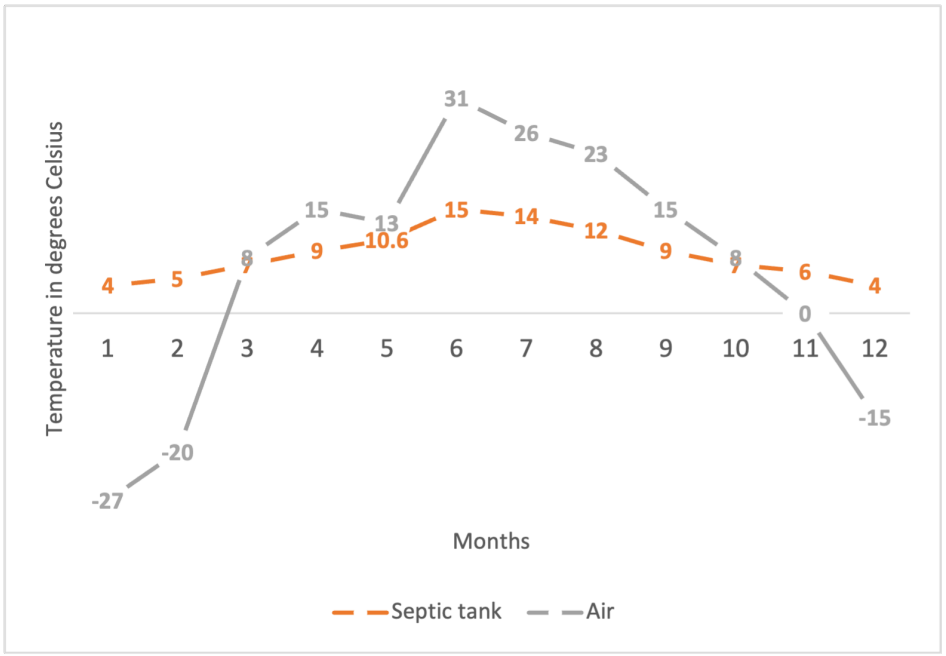

Figure 6.2: Monthly variation in ambient air and septic tank wastewater temperatures across the study period of January (month 1) to December (month 12) 2024.

6.2.2 COD and BOD_5

Organic loading varied across seasons, with the highest COD values recorded in winter (1375 ± 1160 mg/L) and autumn (1572 ± 1462 mg/L), and the lowest in summer (411 ± 160 mg/L) (Table 6.1; Figure 6.3). BOD_5 followed the same pattern, ranging from 485 ± 442 mg/L in winter to 166 ± 101 mg/L in summer (Table 6.1; Figure 6.4). Despite differences in absolute concentrations, BOD_5/COD ratios remained within ranges typical of domestic wastewater (approximately 0.3–0.6), indicating broadly consistent proportions of biodegradable versus non-biodegradable organic matter across seasons (Tchobanoglous et al., 2003; Metcalf and Eddy, 2013).

Table 6.1: Seasonal physicochemical parameters (mean ± SD). Concentrations are reported in mg/L; VS% is reported as percent.

Season	COD		BOD_5		TS		VS		VS%	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Winter	1375	1160	485	442	1313	1742	1001	1440	76	83
Spring	721	379	231	108	2540	5625	2498	5675	98	101
Summer	411	160	166	101	8	7	7	7	88	100
Autumn	1572	1462	370	279	6	9	6	8	100	89

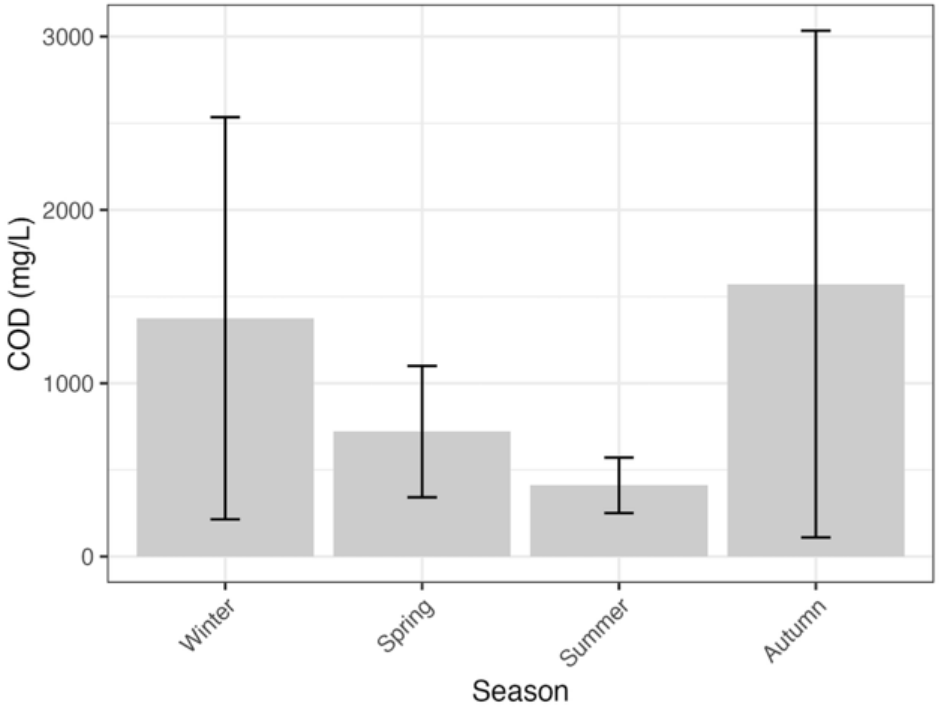

Figure 6.3: Seasonal variation in COD concentrations in septic tank wastewater.

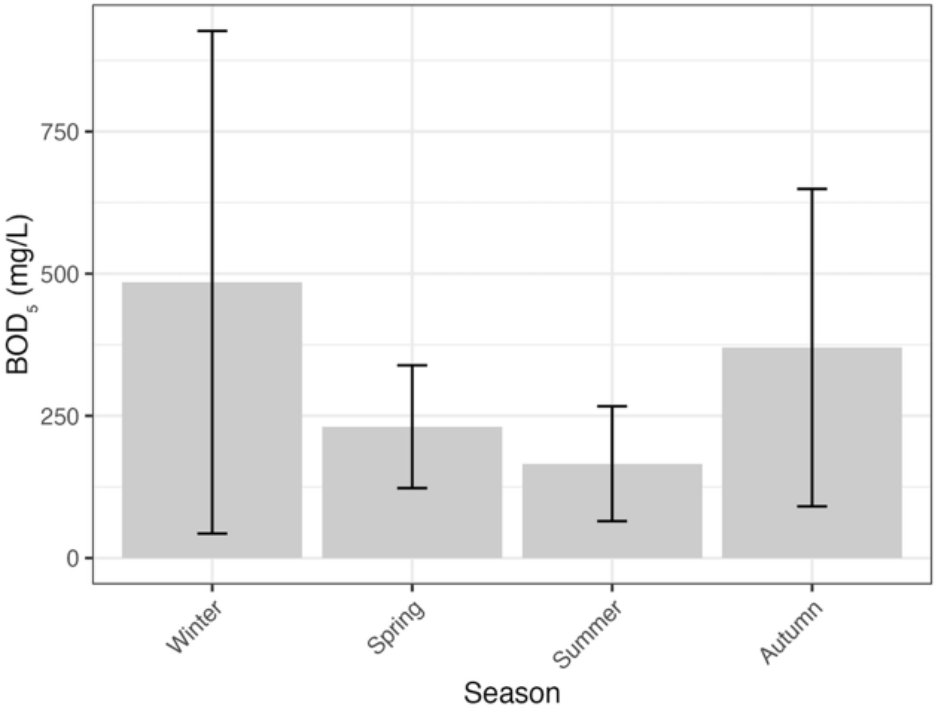

Figure 6.4: Seasonal variation in BOD_5 concentrations in septic tank wastewater.

6.2.3 TS and VS

Both TS and VS exhibited pronounced variation (Table 6.1; Figures 6.5 and 6.6). The highest seasonal means were observed in spring (TS: 2540 ± 5625 mg/L; VS: 2498 ± 5675 mg/L), with substantially lower values in summer and autumn. The wide standard deviations, particularly in spring, reflect substantial heterogeneity among tanks and between sampling events (Table 6.1).

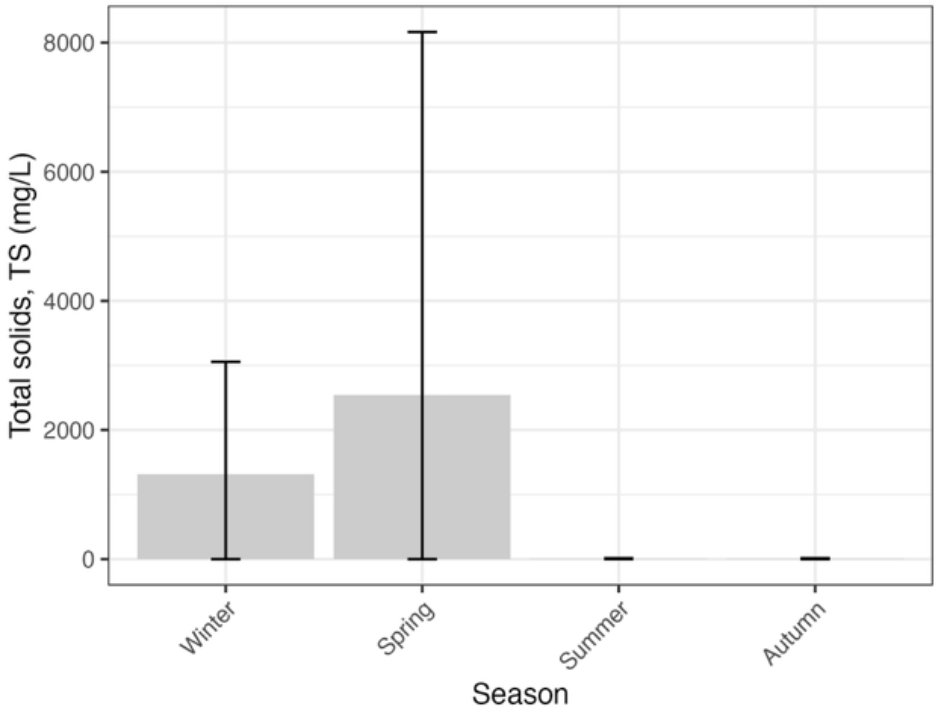

Figure 6.5: Seasonal variation in TS concentrations in septic tank wastewater.

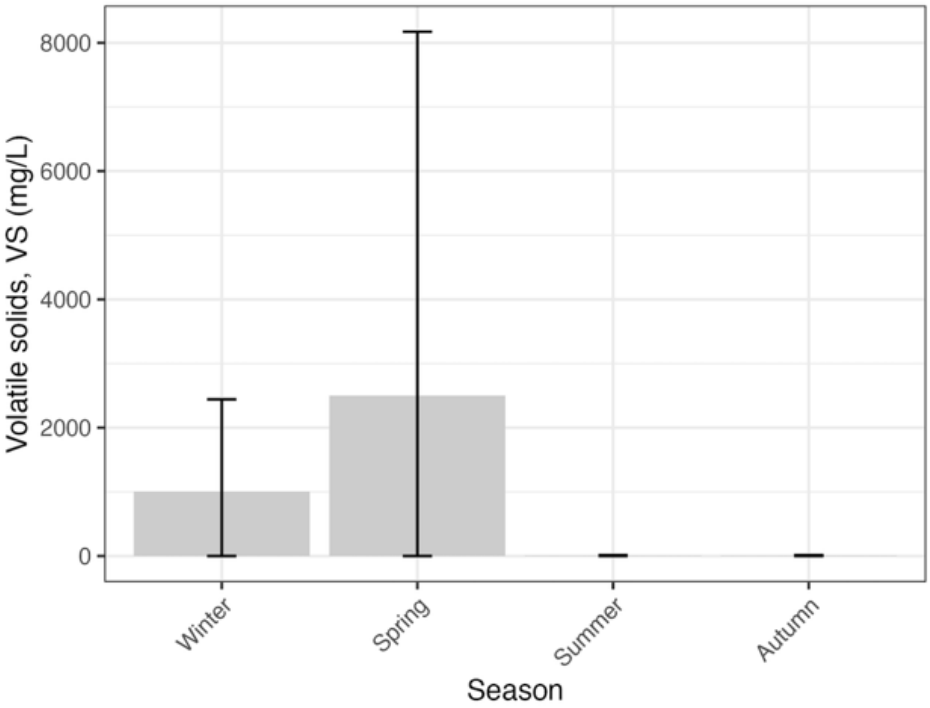

Figure 6.6: Seasonal variation in VS concentrations in septic tank wastewater.

6.2.4 Organic fraction

The volatile solids fraction (VS%) remained high across the year (Table 6.1; Figure 6.7), with seasonal means ranging from approximately 76% (winter) to near 100% (spring and autumn). However, very low solids content in summer and autumn resulted in unstable percentage estimates and occasional VS% values exceeding 100%. VS% values from winter and spring are therefore the most reliable indicators of the organic fraction of stored solids.

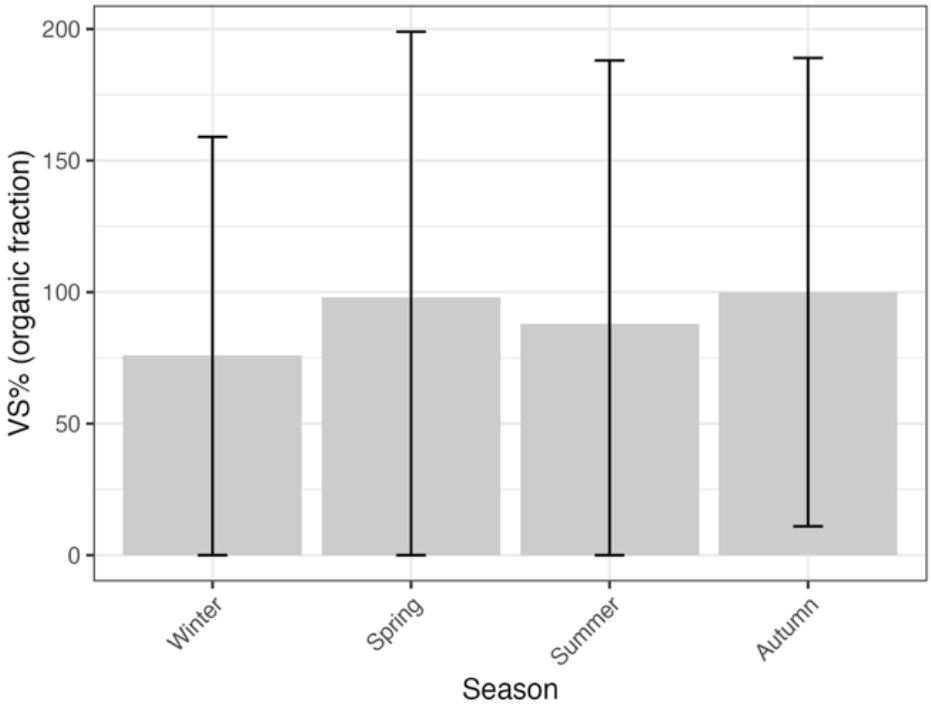

Figure 6.7: Seasonal variation in the organic fraction of solids (VS%) across ten domestic septic tanks.

6.2.5 Temperature and solids

Temperature exerted a pronounced seasonal influence on solids accumulation and organic loading. Thawing conditions in spring corresponded to a marked rise in suspended solids and associated parameters, whereas samples collected in summer and autumn frequently consisted of clarified supernatant with very low TS and VS concentrations. These seasonal shifts define the environmental conditions in which microbial and viral communities operated.

6.2.6 Heterogeneity between tanks

Large standard deviations across COD, BOD_5, TS, VS, and VS% reflect substantial between-tank variability characteristic of decentralised systems. Differences in household water use, influent strength, desludging intervals, retention times, and tank configuration likely contribute to this divergence. For this reason, seasonal averaging is used only for physicochemical parameters. Microbial and viral community analyses are based on individual samples, as averaging would mask meaningful ecological variation between tanks.

6.2.7 Implications

- The observed thermal buffering of septic tanks relative to ambient air temperature indicates that microbial communities experience a more stable thermal environment than external climatic conditions suggest. Similar thermal damping effects in buried septic tanks have been reported in cold-climate settings, where soil insulation moderates short-term temperature fluctuations (Bian et al., 2024).

- The spring increase in solids concentrations is consistent with the physical mobilisation and resuspension of solids accumulated during winter, driven by thaw-related changes in hydraulic conditions. Similar springtime mobilisation of stored solids has been reported in cold-climate on-site sanitation systems, where freeze–thaw processes and increased hydraulic disturbance result in transient releases of particulate matter (Kinsley et al., 2017). This physical disturbance mechanism is consistent with the pronounced variability observed in microbial community composition during the spring period.

- The very low TS and VS concentrations observed during summer and autumn indicate that microbial and viral community data from these seasons predominantly reflect organisms present in comparatively clarified wastewater rather than sludge-rich material. This pattern may reflect relatively frequent desludging practices at the studied households, which would reduce solids accumulation within the tanks. Such practices appear more intensive than those typically reported in the on-site sanitation literature and may reflect local operational behaviours rather than standardised management guidelines.

- The relative stability of BOD_5/COD ratios across seasons suggests that, despite pronounced seasonal variation in absolute organic loads, the proportion of biodegradable to non-biodegradable organic matter remained characteristic of domestic wastewater. Similar observations have been reported for domestic wastewater and on-site sanitation systems (Gray, 2004; Rudaru et al., 2022).

6.3 Microbial and viral community structure

This section examines α-diversity patterns, β-diversity relationships, and multivariate ordination analyses describing the structure and seasonal dynamics of prokaryotic and viral communities in household septic tanks. All multivariate analyses were performed using Aitchison distances following CLR transformation to appropriately account for the compositional nature of metagenomic relative abundance data.

6.3.1 Alpha-diversity patterns

Clear differences were observed between prokaryotic and viral diversity across the household septic tank samples (Figures 6.8– 6.10). Prokaryotic richness was relatively consistent, ranging from approximately 24 to 270 detected families, whereas viral richness was substantially lower, typically ranging from 4 to 10 families per sample.

A weak but statistically significant negative correlation was observed between prokaryotic and viral richness (Spearman $\rho = -0.34$, p = 0.016; Figure 6.8), indicating that samples with higher prokaryotic richness tended to contain slightly fewer viral families. In contrast, Shannon diversity showed a positive but non-significant association between domains (Spearman $\rho = 0.17$, p = 0.24; Figure 6.9), while inverse Simpson diversity showed minimal correspondence between prokaryotic and viral communities (Spearman $\rho = 0.05$, p = 0.71; Figure 6.10).

Overall, prokaryotic and viral α-diversity metrics showed limited correspondence, with only a modest inverse relationship observed for richness. This suggests that viral diversity did not scale proportionally with host taxonomic richness. Instead, viral community structure appears to be shaped primarily by host identity, infection dynamics, and broader ecological interactions within the septic tank environment. Such decoupling is consistent with engineered systems exposed to strong seasonal forcing and disturbance, where viral diversity reflects context-dependent interaction processes rather than simple numerical coupling with host communities (Zhang et al., 2017; Fan et al., 2023).

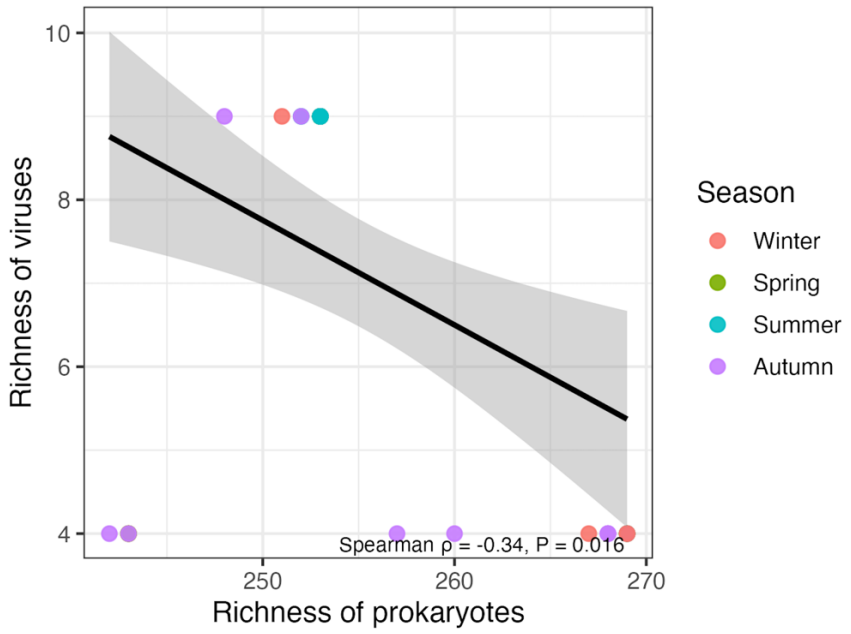

Figure 6.8: Richness relationship between viral and prokaryotic communities. Scatterplot showing viral richness versus prokaryotic richness across all ten monitored septic tanks, coloured by season. Spearman's rank correlation coefficient (ρ) and p-value are shown in the figure.

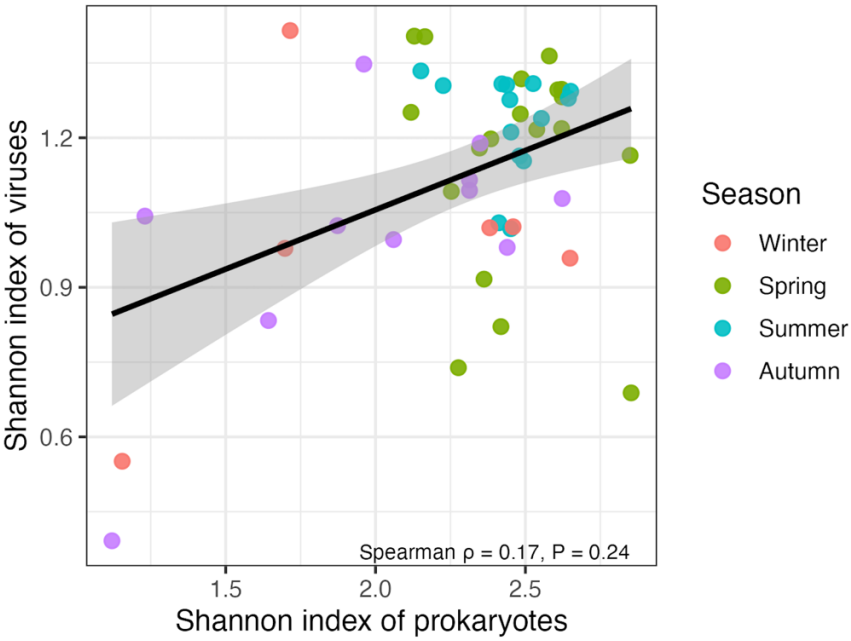

Figure 6.9: Shannon diversity of viral and prokaryotic communities. Scatterplot comparing Shannon diversity indices of viruses and prokaryotes across seasons and all ten monitored septic tanks. Spearman coefficient (ρ) and p-value are shown in the figure.

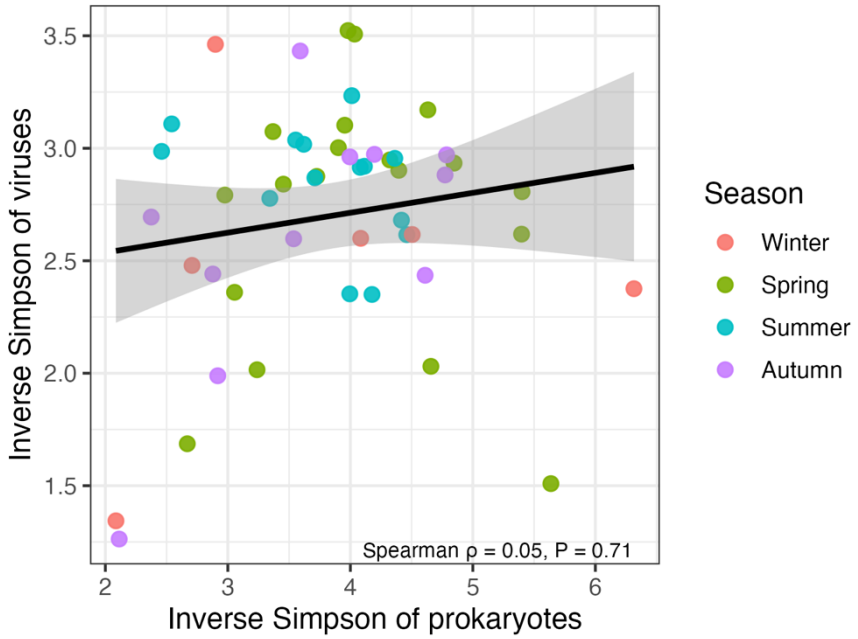

Figure 6.10: Inverse Simpson diversity of viral and prokaryotic communities in 10 septic tanks. Scatterplot comparing inverse Simpson diversity indices for viruses and prokaryotes across all samples from ten monitored septic tanks, coloured by season. Each point represents one sampling event; Spearman's rank correlation coefficient (ρ) and p-value are shown in the figure.

6.3.2 Beta-diversity patterns

Correlation between viral and prokaryotic β-diversity

Community-level dissimilarity patterns of prokaryotic and viral assemblages were assessed using Aitchison distances derived from CLR–transformed relative abundances. Mantel analysis revealed a strong and statistically significant association between prokaryotic and viral community dissimilarity (r = 0.68, p = 0.001; Figure 6.11), indicating that samples that were compositionally distinct in their prokaryotic communities were also distinct in their viral communities. This demonstrates coordinated community turnover across the annual cycle, consistent with strong host–virus coupling within the anaerobic environment (Zhang et al., 2017; Fan et al., 2023).

Seasonal patterns in β-diversity

PERMANOVA, with permutations constrained by septic tank identity to account for repeated measures, demonstrated a significant effect of season on both prokaryotic and viral community composition. Season explained approximately 21% of the variation in

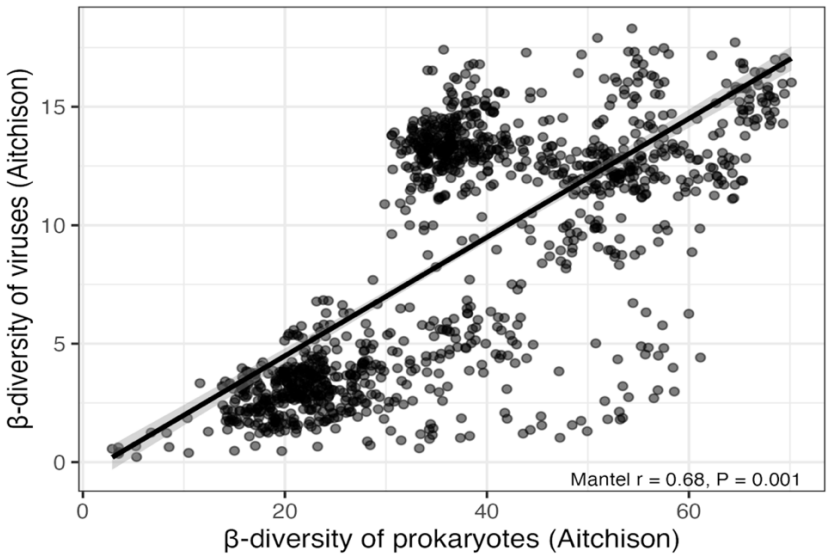

Figure 6.11: Mantel test comparing Aitchison distance matrices of viral and prokaryotic communities across all ten monitored septic tanks. Mantel r and p-values are shown in the figure.

prokaryotic community structure ($R^2 = 0.21$, $p = 0.001$) and 34% of the variation in viral community structure ($R^2 = 0.34$, $p = 0.001$). The stronger seasonal signal observed for viral communities indicates a higher sensitivity of viral assemblages to seasonal forcing relative to their prokaryotic hosts.

Seasonal comparisons revealed substantial differences in within-season and between-season β-diversity (Figure 6.12). Summer samples exhibited the lowest within-season dispersion, indicating relatively homogenous microbial and viral communities across tanks during this period. In contrast, spring samples displayed the highest within-season dispersion, spanning nearly the full observed range of Aitchison distances, reflecting elevated heterogeneity and transitional community states. Winter and autumn showed intermediate variability.

Across all seasons, between-season distances were consistently higher than within-season distances, demonstrating systematic seasonal restructuring of both prokaryotic and viral communities rather than random temporal fluctuation.

Tests for homogeneity of multivariate dispersion revealed significant differences in dispersion among seasons for prokaryotic communities ($p = 0.009$), suggesting that seasonal differences in prokaryotic β-diversity reflected both shifts in centroid location and changes in within-season variability. In contrast, viral community dispersion showed only marginal seasonal differences ($p = 0.056$), indicating that the observed seasonal effect

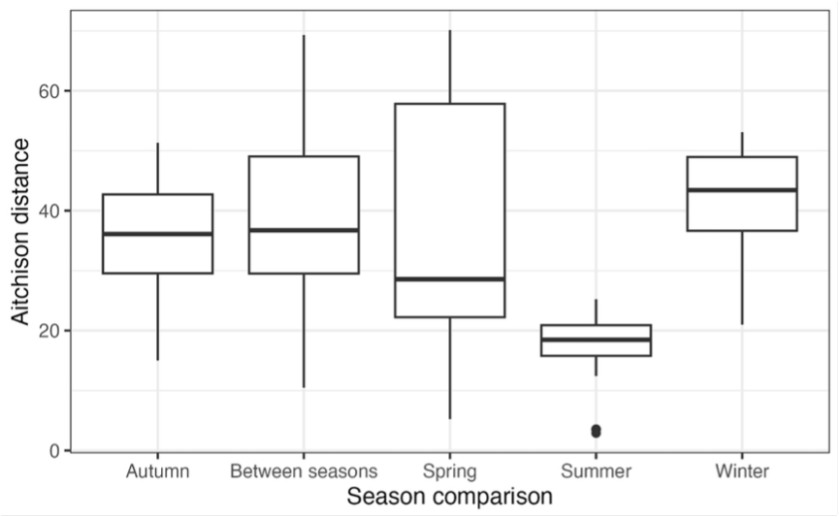

Figure 6.12: Boxplots of pairwise Aitchison distances showing within-season and between-season β-diversity patterns.

on viral community structure was driven primarily by shifts in community composition rather than dispersion alone.

Tank-level β-diversity and system comparability

Pronounced variability was also observed at the individual tank level (Figure 6.13). Tanks 6 and 9 showed low within-tank distances, indicating relatively stable temporal community structure, whereas tanks 7, 8, and 10 showed higher temporal dispersion, reflecting more dynamic within-tank variability. Tanks 1-5 displayed intermediate behaviour.

Importantly, across the dataset, within-tank distances remained consistently lower than between-tank distances, indicating that each tank maintained a distinct yet coherent microbial and viral signature over time.

6.3.3 PCoA and NMDS ordination analyses

Principal coordinates analysis based on Aitchison distances revealed clear seasonal structuring of both prokaryotic and viral communities.

Prokaryotic communities

Prokaryotic samples clustered strongly by season (Figure 6.14). Summer samples formed a compact and well-defined cluster, indicating relatively homogenous community composition during this period. Spring samples were widely dispersed, consistent with

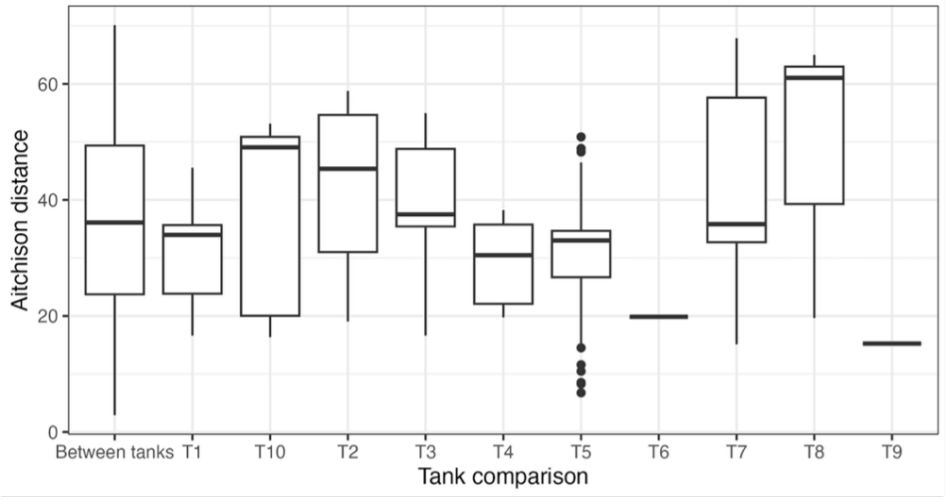

Figure 6.13: Boxplots of pairwise Aitchison distances showing within-tank and between-tank β-diversity patterns.

elevated β-diversity and greater between-sample variability. Winter and autumn occupied intermediate and partially overlapping positions, suggesting transitional community states. PCoA1 and PCoA2 explained 42.6% and 10.4% of the total variance, respectively, indicating that seasonal effects accounted for a substantial proportion of prokaryotic community structuring.

Viral communities

Viral communities showed a similar but more strongly expressed seasonal structure (Figure 6.15). Summer samples again formed a dense cluster, while spring samples displayed the greatest dispersion. Winter and autumn samples were more variable and partially overlapping. Notably, PCoA1 alone explained 71.9% of the total variance in viral community composition, indicating that seasonal dynamics constituted a dominant structuring gradient for viral assemblages.

Non-metric multidimensional scaling analyses supported the PCoA results (Appendix D; Figures D.1- D.2). Prokaryotic NMDS achieved a stress value of 0.0917, indicating a reliable two-dimensional representation, while viral NMDS achieved a stress value of 0.0350, indicating excellent ordination fit. Across both ordination methods, summer samples clustered tightly and spring samples were highly dispersed, reinforcing the robustness of the observed seasonal patterns.

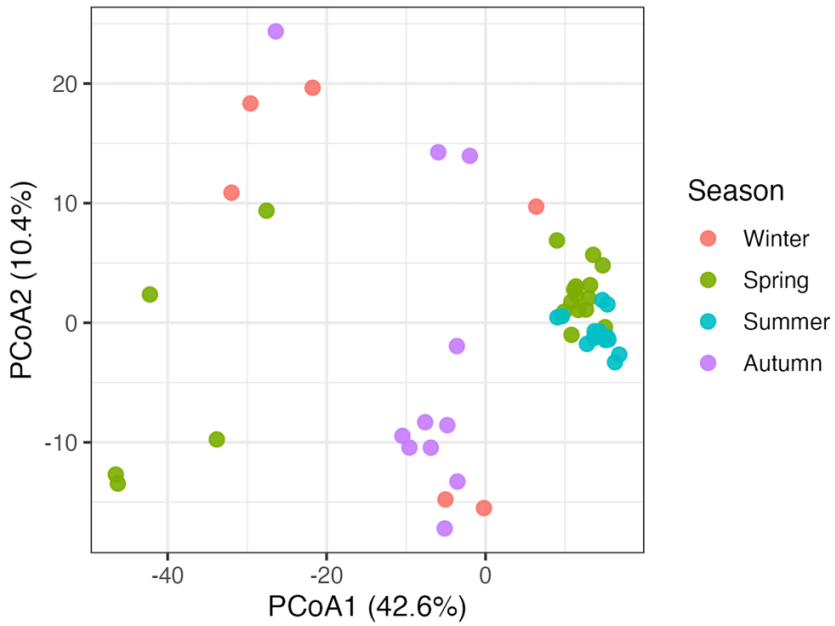

Figure 6.14: PCoA ordination of prokaryotic communities (Aitchison distance). Principal coordinates analysis showing seasonal clustering of prokaryotic community composition across all ten monitored septic tanks.

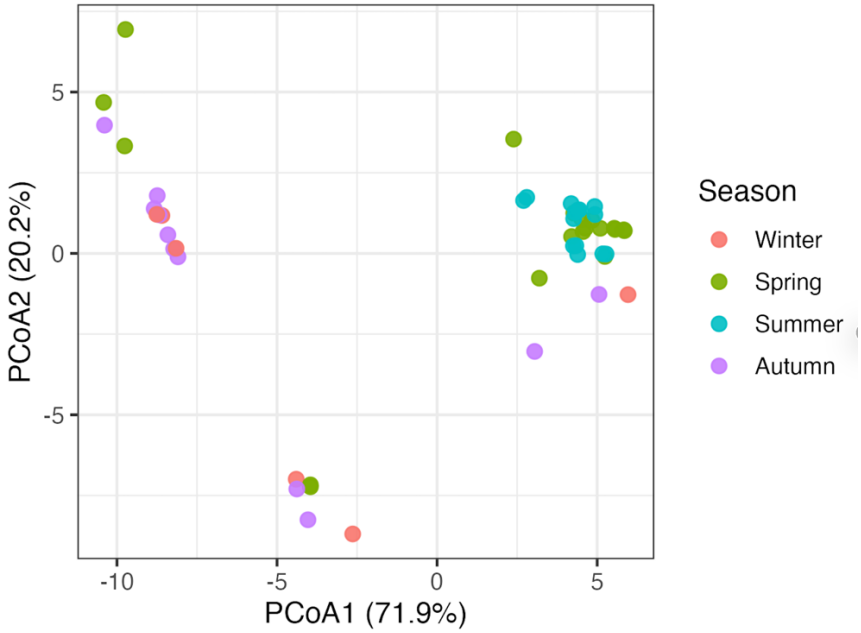

Figure 6.15: PCoA ordination of viral communities (Aitchison distance). Principal coordinates analysis showing seasonal clustering of viral community composition across all ten monitored septic tanks.

6.3.4 Implications

The observed decoupling between prokaryotic and viral α-diversity metrics indicates that within-sample richness and evenness responded differently across domains, despite coordinated changes in community composition at the β-diversity level. This pattern is consistent with observations from anaerobic treatment systems, where viral diversity is not directly proportional to host richness but is instead shaped by host identity, metabolic states, and interaction dynamics (Gradilla-Hernández et al., 2018; Trubl et al., 2021). The present study indicates that similar diversity relationships occur in decentralised septic tank systems operating under field conditions, even under pronounced seasonal temperature variability.

In contrast, the strong correspondence between prokaryotic and viral β-diversity indicates that shifts in prokaryotic community composition were closely mirrored by viral communities across the annual cycle. This finding suggests that viral assemblages responded primarily to changes in host community structure rather than to absolute host diversity. While coordinated host–virus community patterns have been reported in laboratory-scale and full-scale anaerobic digesters, this study provides field-based, year-round evidence that such coupling is also expressed in operational septic tanks in cold-climate settings.

Although each septic tank retained a recognisable microbial and viral signature over time, multivariate analyses consistently identified that seasonal forcing was the dominant source of variation across the dataset. Samples clustered primarily by season rather than by tank identity in both prokaryotic and viral ordinations, and Mantel analysis confirmed strong coupling between domains. Together, these results indicate that the ten septic tanks functioned as comparable system replicates responding coherently to seasonal temperature dynamics, rather than representing fundamentally dissimilar sanitation systems.

These findings are consistent with previous studies reporting substantial similarity of microbial communities across on-site sanitation systems at higher taxonomic levels and among dominant taxa, including the presence of stable core communities across containment types and operational histories (Sam et al., 2025). The present results further demonstrate that such apparent stability at higher taxonomic resolution can coexist with

pronounced seasonal differences in community composition at finer taxonomic resolution and within viral communities. In this context, persistence of core anaerobic assemblages does not preclude strong seasonal variability in relative abundance patterns and host–virus associations.

Elevated variability observed during spring across multiple analytical approaches highlights a period of pronounced community heterogeneity, likely reflecting transitional physicochemical conditions associated with seasonal thaw. In contrast, reduced variability observed during summer indicates convergence towards more similar microbial and viral assemblages under warmer and more stable conditions, consistent with sustained anaerobic metabolic activity and reduced physical disturbance (Gray, 2004; Hernández and Vives, 2020; Trubl et al., 2021).

Persistent tank-specific signatures, evidenced by consistently lower within-tank than between-tank dissimilarities, confirm that each septic tank followed a distinct microbial trajectory over time. However, the concordance among β-diversity metrics, ordination analyses, and Mantel results demonstrates that the seasonal organisation of community structure was robust across multiple analytical frameworks.

Collectively, these findings indicate that prokaryotic and viral community structure in septic tanks reflects the combined influence of seasonal temperature variation and system-specific conditions. Importantly, coordinated changes in prokaryotic and viral community composition occurred within an otherwise conserved anaerobic framework, providing new insight into microbial–viral organisation in decentralised sanitation systems operating under cold-climate conditions.

6.4 Functional composition of AD guilds and viral communities

This section examines the functional composition and seasonal behaviour of AD-associated microbial guilds and viral communities across ten household septic tanks monitored throughout an annual cycle. Prokaryotic taxa were assigned to four functional guilds central to AD processes: fermentative bacteria, syntrophic acetate-oxidising bacteria, methanogenic archaea, and sulphate-reducing bacteria. Seasonal changes in the relative abundance and taxonomic composition of these guilds, together with associated viral communities, are presented to characterise how anaerobic microbial organisation varies

under field conditions in a cold-climate decentralised sanitation context.

6.4.1 Seasonal abundance of AD functional guilds

Across all seasons, fermentative bacteria constituted the dominant functional guild, accounting for approximately 70–90% of the prokaryotic community assigned to AD-related functions (Figure 6.16). This persistent dominance reflects the central role of fermentation in the initial stages of AD, encompassing hydrolysis and acidogenesis, and the continuous input of readily degradable organic matter into household septic tanks. In contrast, SAOB, SRB, and methanogens together comprised smaller but functionally critical fractions of the community associated with downstream anaerobic processes.

Despite the overall dominance of fermenters, seasonal variation was evident in the relative contributions of the non-fermentative guilds. During winter, methanogens and SRB were present at comparatively low relative abundances. In spring, the proportional representation of SAOB and SRB increased. Summer samples exhibited the highest relative contribution of SRB, accompanied by a reduced proportion of methanogens. By autumn, SAOB, SRB, and methanogens each contributed moderate and more balanced proportions to the functional community.

These results indicate that while fermentative bacteria consistently dominate the AD-associated microbial assemblage, the relative representation of downstream anaerobic guilds varies systematically across seasons under *in situ* operating conditions.

6.4.2 Seasonal balance between SRB and methanogens

Clear seasonal shifts were observed in the relative balance between SRB and methanogenic archaea across the ten monitored household septic tanks (Figure 6.17). During winter, methanogens accounted for approximately 60% of the combined SRB–methanogen assemblage. This balance shifted in spring, when SRB increased to represent around 65% of the assemblage. In summer, SRB dominance intensified further, accounting for nearly 90% of the combined relative abundance, while methanogens were strongly reduced. By autumn, methanogens partially recovered, contributing approximately 25% of the assemblage.

Together, these results indicate a consistent seasonal restructuring of the SRB and

Figure 6.16: Relative abundance of AD functional guilds across seasons. Bars represent the mean proportional contribution of fermenters, SAOB, SRB, and methanogens to the prokaryotic community across all ten monitored septic tanks in each season. Values are normalised to 100% of all classified taxa.

methanogens assemblage, with each season characterised by a distinct balance between these two competing anaerobic guilds across all monitored systems.

6.4.3 Seasonal dynamics of dominant AD guilds and viral community composition

Seasonal variation was evident in the taxonomic composition of all major anaerobic functional guilds and viral families, although the magnitude and structure of change differed between groups. Seasonal relative abundances of dominant families within each guild are presented in Figures 6.18– 6.22.

Fermentative bacteria

Fermentative bacteria exhibited the highest taxonomic diversity among all functional groups and showed pronounced seasonal restructuring (Figure 6.18). Winter communities comprised a diverse assemblage of families, including Bacteroidaceae, Bifidobacteriaceae, Carnobacteriaceae, Lachnospiraceae, Porphyromonadaceae, Prevotellaceae, and

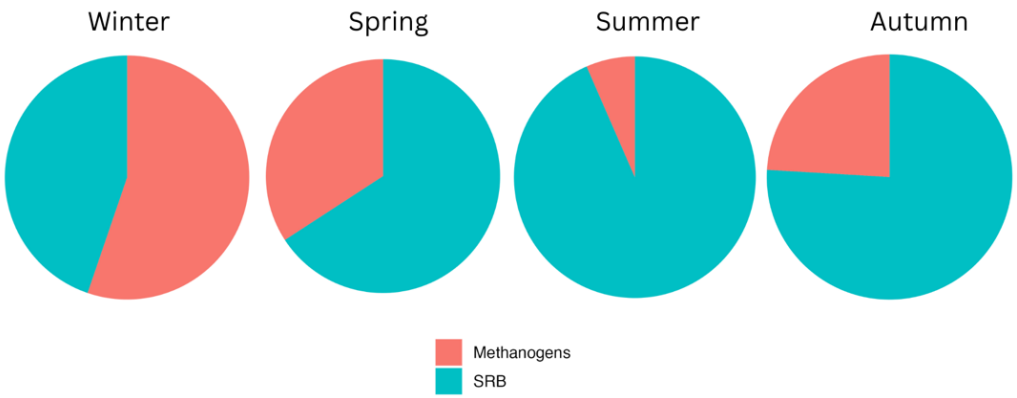

Figure 6.17: Seasonal variations in the SRB-to-methanogen ratio across all ten monitored household septic tanks.

Propionibacteriaceae, alongside numerous low-abundance taxa. In spring and summer, the fermentative community shifted towards dominance by Bacteroidaceae, Porphyromonadaceae, and Prevotellaceae, accompanied by a reduced contribution from Bifidobacteriaceae, Carnobacteriaceae, and Propionibacteriaceae. By autumn, fermentative composition shifted again, with a marked increase in Carnobacteriaceae and a decline in several families that dominated during summer.

Methanogenic archaea

Methanogenic community composition varied markedly across seasons (Figure 6.19). Winter and spring communities were dominated by acetoclastic methanogens, primarily Methanotrichaceae, with smaller contributions from Methanosarcinaceae. During summer and autumn, the relative abundance of Methanotrichaceae declined sharply, coinciding with increased dominance of hydrogenotrophic methanogens, particularly Methanocorpusculaceae. Other hydrogenotrophic families, including Methanobacteriaceae, Methanomicrobiaceae, Methanoregulaceae, and Methanospirillaceae, exhibited moderate seasonal fluctuations, while the methylotrophic Methanomassiliicoccaceae remained consistently present at low relative abundance throughout the year. Overall, methanogenic communities exhibited two contrasting seasonal compositional states: acetoclastic dominance in winter and spring, and hydrogenotrophic dominance in summer and autumn.

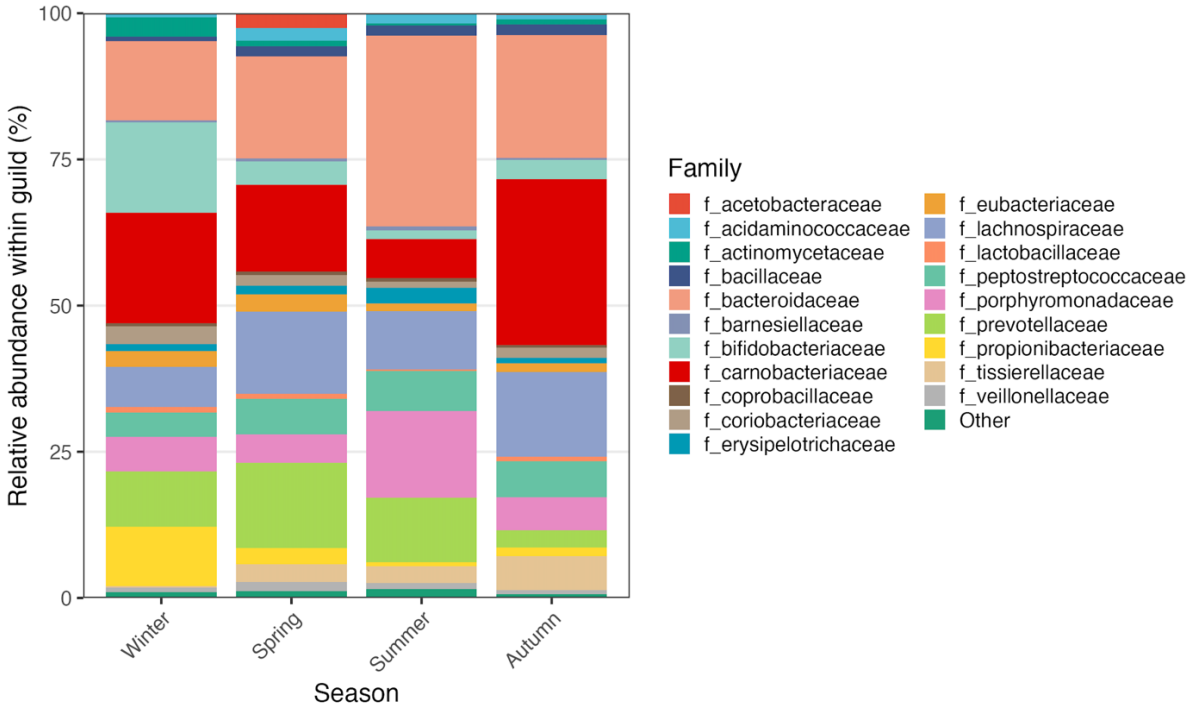

Figure 6.18: Seasonal relative abundance of the top 20 fermentative bacterial families (100%-stacked) across all ten household septic tanks.

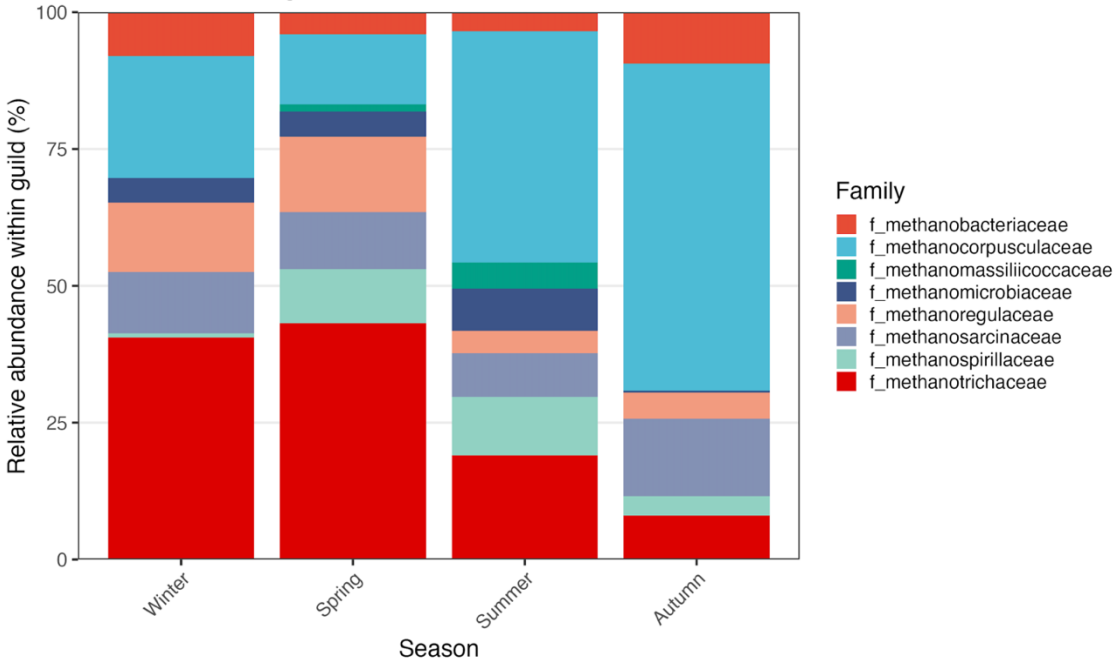

Figure 6.19: Seasonal relative abundance of methanogenic families (100%-stacked) across all ten household septic tanks.

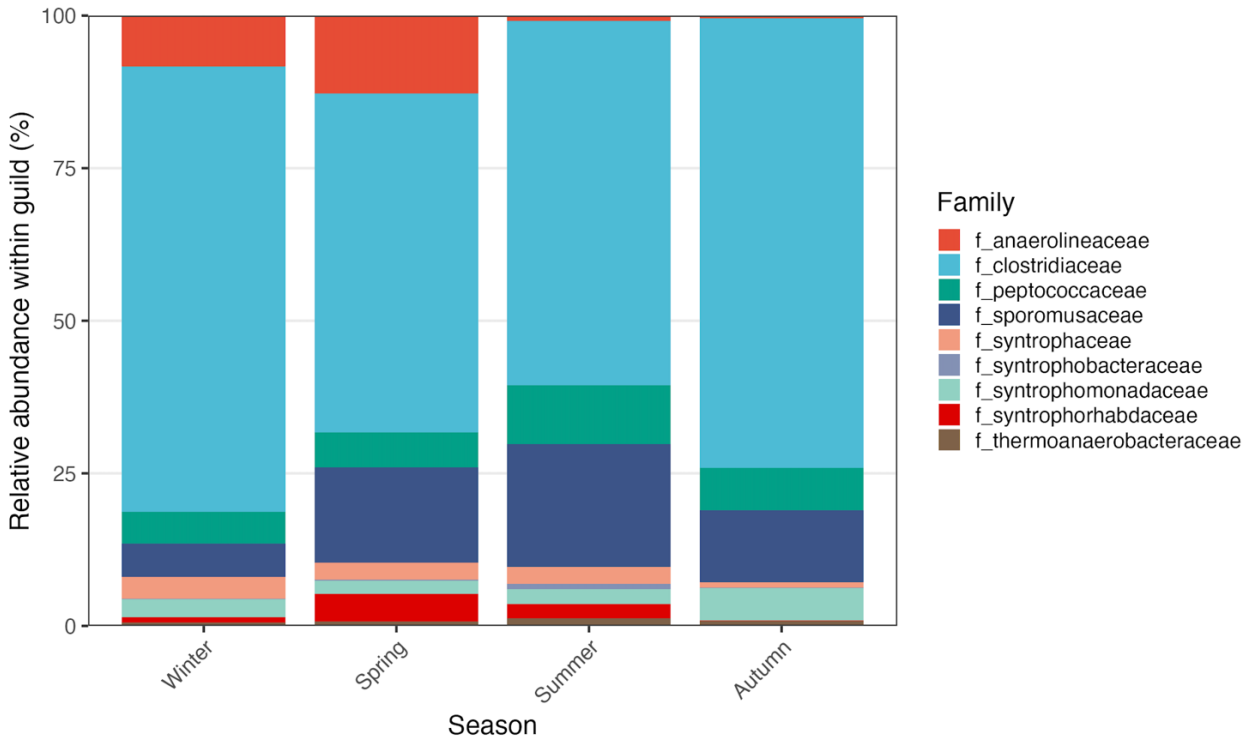

Figure 6.20: Seasonal relative abundance of syntrophic acetate-oxidising bacterial families (100%-stacked) across all ten household septic tanks.

Syntrophic acetate-oxidising bacteria

Compared with other functional guilds, SAOB showed relatively limited seasonal variation in taxonomic composition (Figure 6.20). Across all seasons, Clostridiaceae overwhelmingly dominated the SAOB assemblage, with minor contributions from Sporomusaceae, Peptococcaceae, Syntrophaceae, and Syntrophomonadaceae. Slight increases in Sporomusaceae and Peptococcaceae were observed during spring and summer, while autumn communities closely resembled winter, with near-exclusive dominance by Clostridiaceae.

Sulphate-reducing bacteria

The SRB guild exhibited pronounced seasonal restructuring at the family level (Figure 6.21). Desulfobulbaceae, Desulfomicrobiaceae, Desulfovibrionaceae, and Desulfobacteriaceae dominated across all seasons, but their relative contributions varied substantially. Desulfobulbaceae declined during spring and summer before returning to winter-like levels in autumn. Desulfomicrobiaceae and Desulfobacteriaceae increased progressively from winter to summer and declined thereafter. Desulfovibrionaceae showed

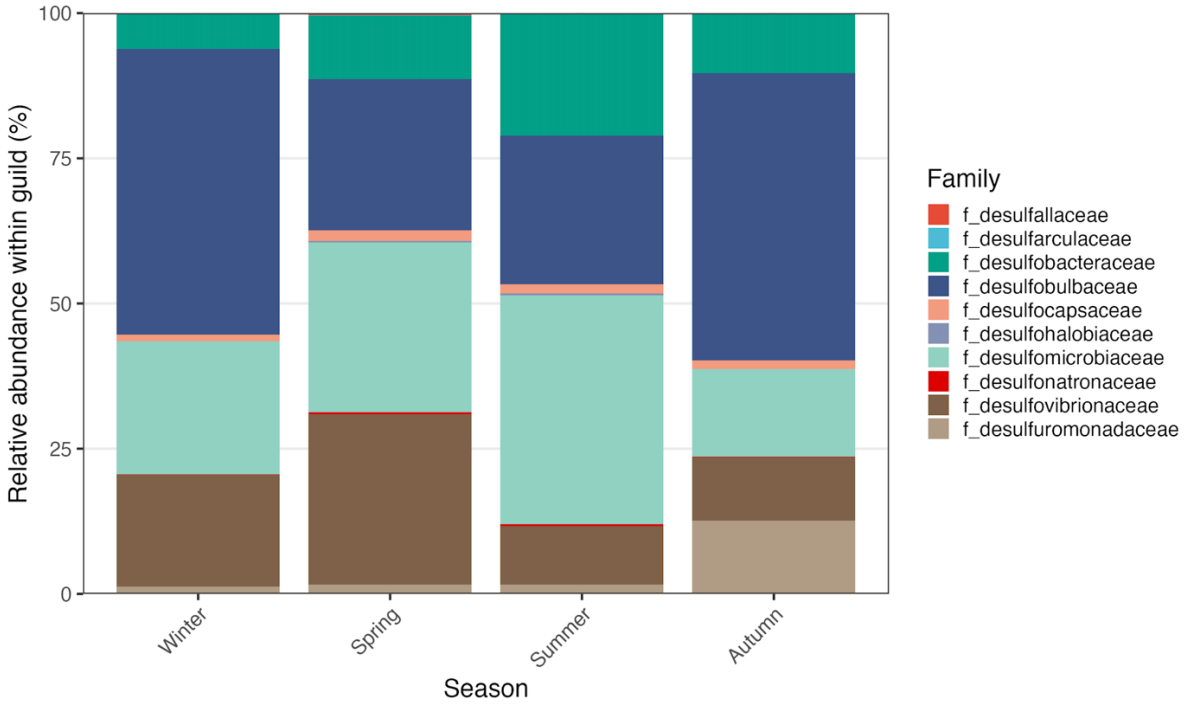

Figure 6.21: Seasonal relative abundance of sulphate-reducing bacterial families across seasons (100%-stacked) across all ten household septic tanks.

a pronounced spring peak, while Desulfuromonadaceae increased substantially during autumn.

Viral families

Seasonal variation in viral family composition was less pronounced than for prokaryotic guilds (Figure 6.22). Across all seasons, viral assemblages were consistently dominated by Siphoviridae, Myoviridae, and Podoviridae, which together accounted for the majority of viral reads. Schitoviridae contributed a smaller but stable fraction. Seasonal changes were primarily confined to lower-abundance families, including Autographiviridae, Herelleviridae, Demecriviridae, and Ackermannviridae, which increased modestly during spring and summer before declining in autumn and winter.

6.4.4 Implications

The results presented in this section document clear seasonal variation in the relative abundance and taxonomic composition of AD-associated microbial guilds and viral communities in household septic tanks under field conditions. The persistent dominance of fermentative bacteria across all seasons is consistent with previous studies of AD systems,

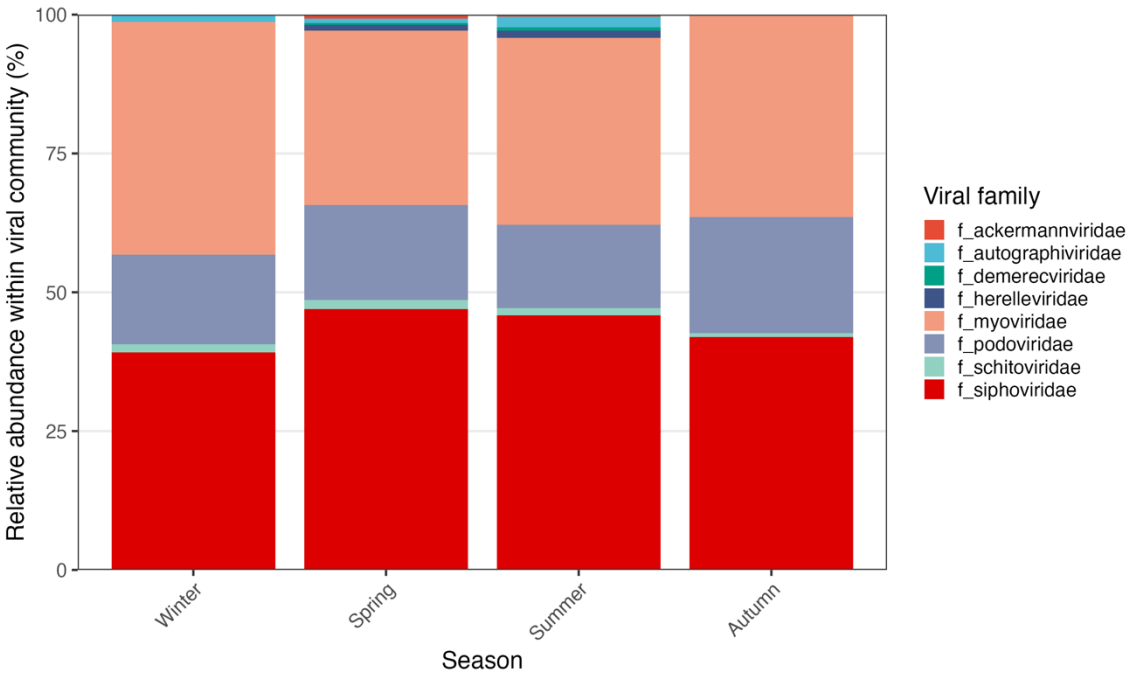

Figure 6.22: Seasonal relative abundance of viral families (100%-stacked) across all ten household septic tanks.

including laboratory-scale digesters and wastewater treatment environments, where fermentative taxa commonly comprise the largest fraction of the microbial community. This observation aligns with existing literature describing the central role of fermentation in anaerobic treatment systems (Mara and Horan, 2008; Parajuli et al., 2022; Rama et al., 2024).

In contrast, the observed seasonal alternation in the relative abundance between SRB and methanogenic archaea extends previous observation by providing year-round field data from decentralised septic tanks in a cold-climate setting. While prior studies have reported temperature-associated differences in the relative abundance of SRB and methanogens in controlled systems (Mara and Horan, 2008; Appels et al., 2008; Stams and Plugge, 2009), comparable seasonal patterns have rarely been documented across multiple operational household septic tanks. The present results therefore contribute new descriptive evidence on seasonal structuring of these guilds *in situ,* without implying changes in reaction rates or pathway dominance, which is discussed in later sections.

Seasonal shifts in methanogenic community composition were observed across the monitored septic tanks, with acetoclastic methanogens contributing a larger fraction of the methanogenic assemblage during winter and spring and hydrogenotrophic methanogens

dominating during summer and autumn. This observation contrasts with previous studies of septic tank systems. For example, Shaw and Dorea (2021) reported no evidence of acetoclastic methanogens in any septic tanks described in the published literature. The consistent detection of acetoclastic methanogenic taxa in the present study therefore extends existing observations by demonstrating their presence in operational household septic tanks under cold-climate field conditions.

The SAOB guild exhibited comparatively limited seasonal variation in taxonomic composition, with persistent dominance of Clostridiaceae across all seasons. Similar persistence of Clostridiaceae has been reported in other anaerobic environments, particularly in mesophilic laboratory scale AD reactors (Regueiro et al., 2015; Gao et al., 2019). The present study extends these observations by documenting the consistent presence of Clostridiaceae-dominated SAOB assemblages under *in situ* field conditions spanning psychrophilic and mesophilic temperature ranges in operational household septic tanks.

Viral community composition was comparatively stable at the family level across seasons, with dominance by Siphoviridae, Myoviridae, and Podoviridae, consistent with previous reports from AD environments (Tamaki et al., 2012; Balleste et al., 2022). Seasonal variation was largely restricted to lower-abundance viral families.

Overall, the findings presented in this section are primarily descriptive and compositional in nature. They confirm several patterns previously reported in anaerobic reactors while extending the available evidence through systematic, year-round field observations in household septic tanks. The results establish a robust structural framework for microbial and viral community organisation across seasons, which provides contextual support for subsequent analyses of functional gene presence.

6.5 Seasonal correlations between anaerobic microbial guilds, viral communities, and physicochemical parameters

Seasonal Spearman correlation analyses (ρ) were applied to assess statistically robust associations among anaerobic microbial guilds, viral families, and physicochemical (abiotic) parameters. Correlations were calculated separately for winter, spring, summer, and autumn. P-values were adjusted for multiple testing using the Benjamini–Hochberg

FDR procedure, and only associations with q $<$ 0.05 were considered statistically significant. Seasonal correlation matrices are presented in Figures 6.23– 6.26.

6.5.1 Seasonal phage–host correlation patterns across anaerobic functional guilds

Winter

The winter correlation matrix (Figure 6.23) contained a limited number of statistically significant associations concentrated among a small set of dominant viral families. Strong positive correlations were observed between several viral families (Schitoviridae, Herelleviridae, Demerecviridae, and Ackermannviridae) and SRB families (Desulfohalobiaceae and Desulfonatronaceae) as well as the methylotrophic methanogenic family Methanomassiliicoccaceae (Spearman $\rho = 1.00$, q = 0.00). The occurrence of perfect rank correlations likely reflects the limited number of winter samples combined with coherent seasonal shifts in relative abundance across functionally linked taxa.

Spring

Spring exhibited the highest number and diversity of significant phage–prokaryote associations (Figure 6.24). Both strong positive and negative correlations were broadly distributed across viral families, with significant associations observed with SRB, SAOB, methanogens, and fermenters. Many associations absent in winter emerged during spring, resulting in the most heterogenous and interconnected correlation structure observed across the annual cycle. This expansion of statistically significant associations coincided with transitional physicochemical conditions and pronounced shifts in microbial community composition.

Summer

In summer, the overall correlation structure became less complex (Figure 6.25). Although visually apparent patterns were present, relatively few phage–guild associations remained statistically significant following FDR correction. This suggests a reduction in tightly coupled viral–microbial co-variation during the warmest period.

Autumn

Autumn samples displayed an intermediate pattern between spring and summer (Figure 6.26). Strong correlations re-emerged between viral families and both SRB (Desulfohalobiaceae, Desulfonatronaceae) and methanogenic (Methanomassiliicoccaceae) groups, while several associations were unique to autumn. As in summer, however, only a limited number of associations remained statistically significant after FDR correction, indicating partial reorganisation of the phage–microbial association structure as the system transitioned toward winter.

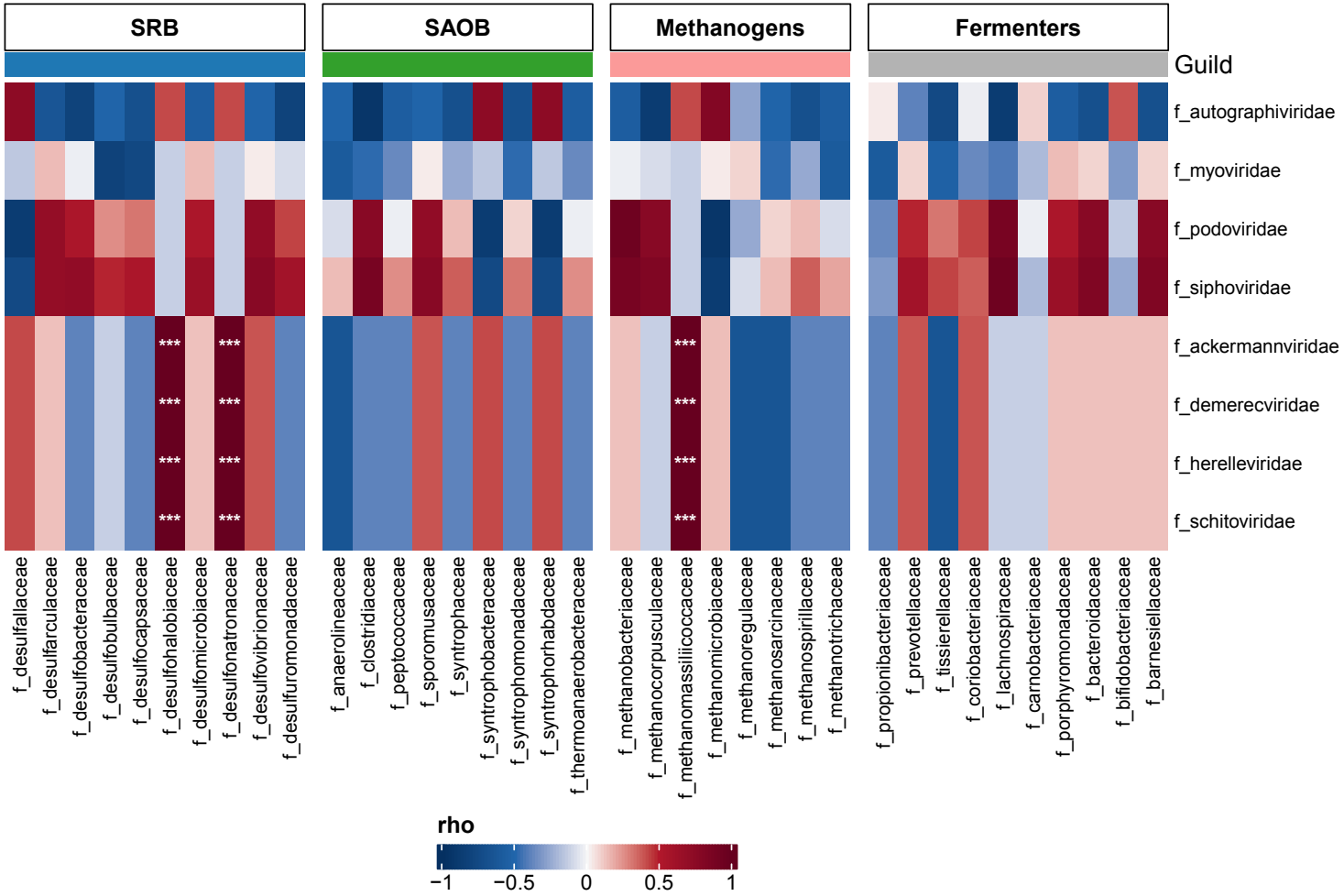

Figure 6.23: Spearman correlations between phage families and prokaryotic functional guilds in winter septic tank samples across all ten monitored septic tanks. FDR-corrected significance levels are indicated as: q ≤ 0.05 (*), q ≤ 0.01 (**), q ≤ 0.001 (***).

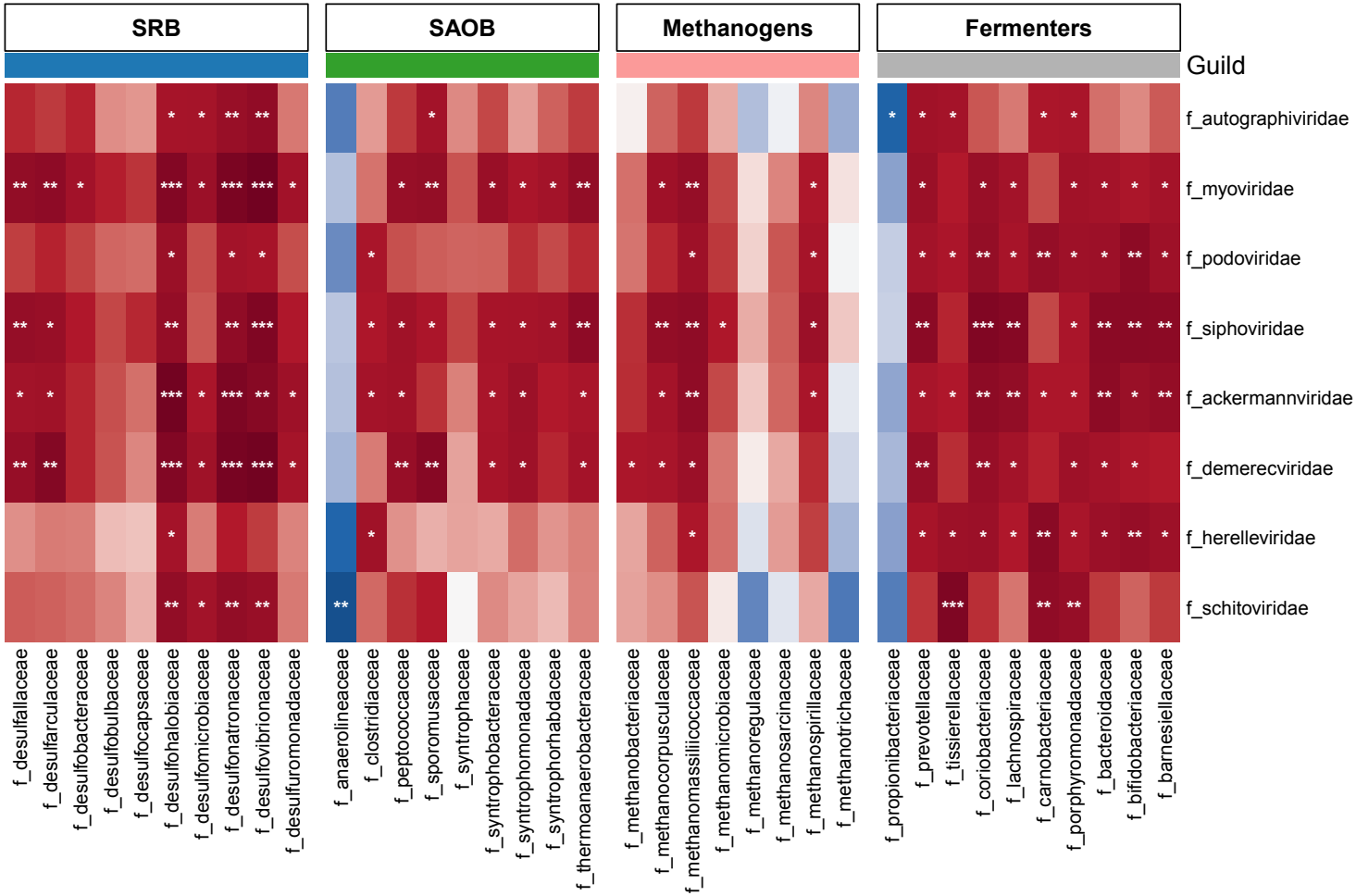

Figure 6.24: Spearman correlations between phage families and prokaryotic functional guilds in spring septic tank samples across all ten monitored septic tanks. FDR-corrected significance levels are indicated as: q $\leq$ 0.05 (*), q $\leq$ 0.01 (**), q $\leq$ 0.001 (***).

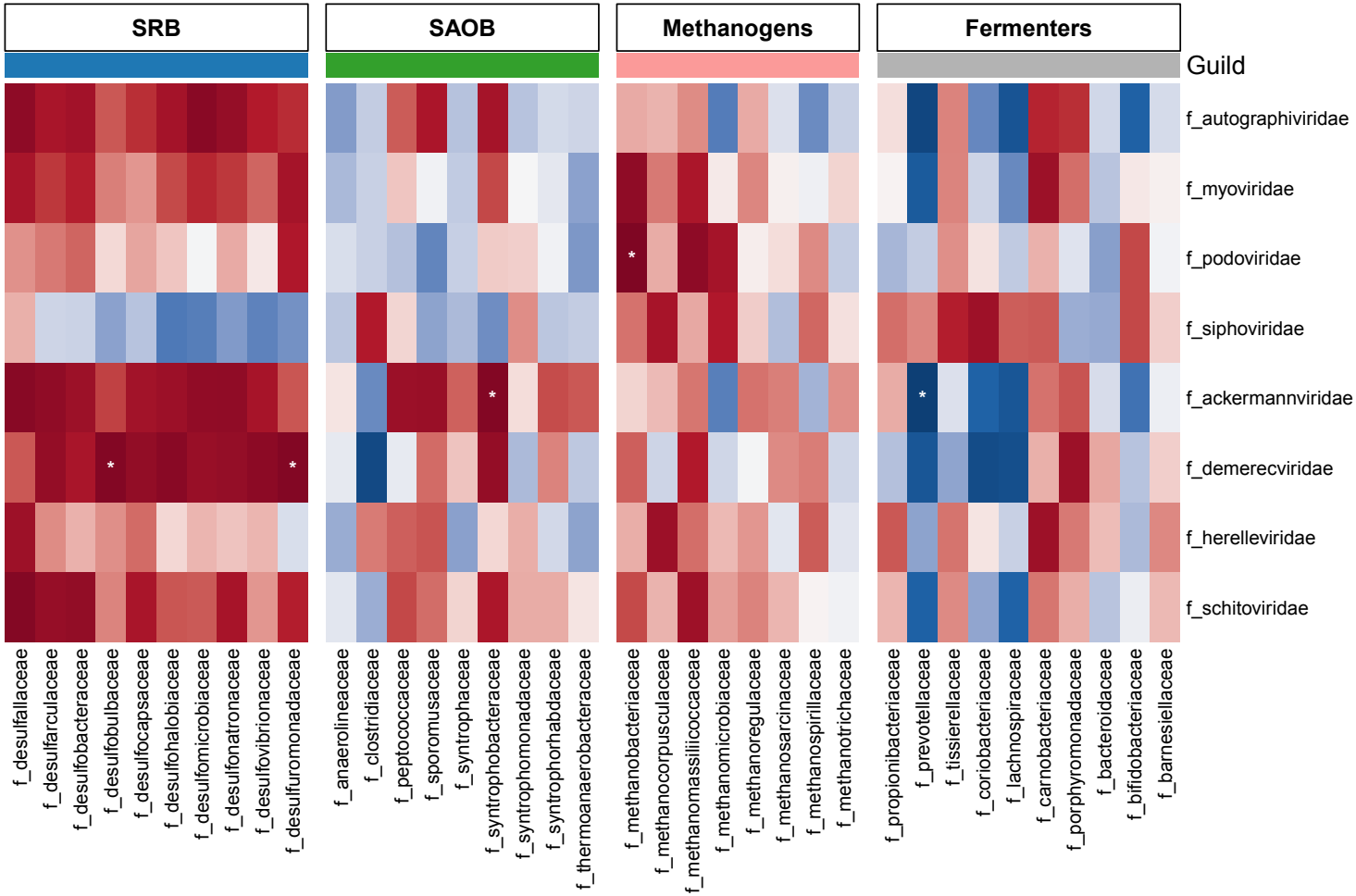

Figure 6.25: Spearman correlations between phage families and prokaryotic functional guilds in summer septic tank samples across all ten monitored septic tanks. FDR-corrected significance levels are indicated as: q $\leq$ 0.05 (*), q $\leq$ 0.01 (**), q $\leq$ 0.001 (***).

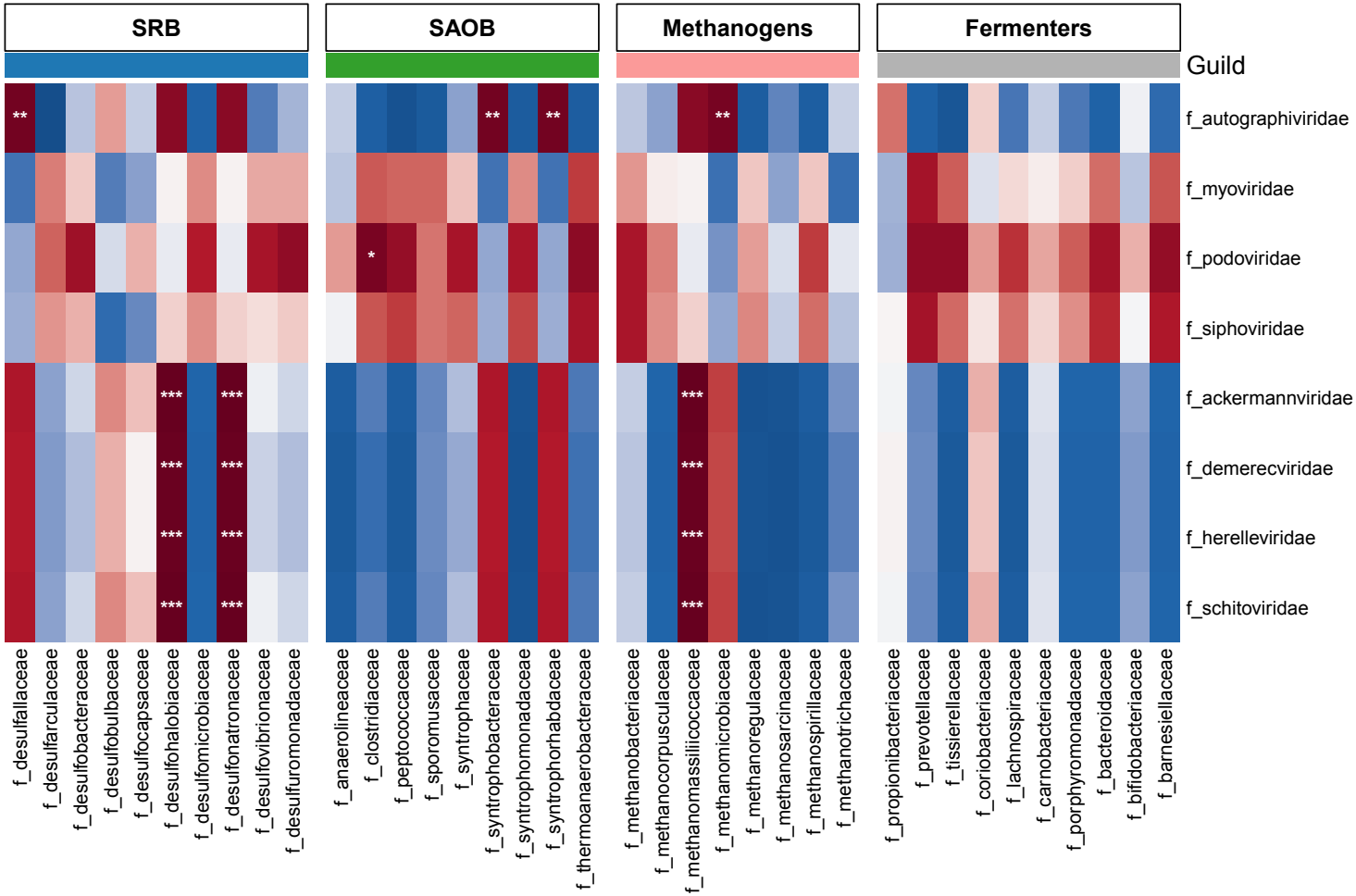

Figure 6.26: Spearman correlations between phage families and prokaryotic functional guilds in autumn septic tank samples across all ten monitored septic tanks. FDR-corrected significance levels are indicated as: q $\leq$ 0.05 (*), q $\leq$ 0.01 (**), q $\leq$ 0.001 (***).

6.5.2 Seasonal correlations between microbial guilds, viral families, and physicochemical parameters

Seasonal Spearman correlations were calculated between physicochemical parameters (wastewater temperature, pH, BOD_5, COD, VS, TS) and the relative abundances of dominant microbial guilds and viral families. P-values were adjusted for multiple testing using the Benjamini-Hochberg FDR procedure, with $q \leq 0.05$ considered statistically significant. Seasonal heatmaps summarising these relationships are presented in Figures 6.27 and 6.28.

Across the annual cycle, statistically robust correlations were detected only during spring. Although structured patterns were apparent in the winter, summer, and autumn heatmaps, none of these associations remained significant following FDR correction and are therefore interpreted as descriptive rather than inferential.

Winter

In winter, correlation heatmaps exhibited structured but statistically unsupported trends, particularly along pH and gradients of solids (TS, VS). SRB families tended to co-occur with higher solids concentrations and lower pH, while several methanogenic families appeared more abundant in samples with higher pH or lower solids. Certain viral families also appeared enriched in samples with higher COD and solids. However, none of these associations reached FDR-adjusted significance, and winter correlations are therefore not interpreted further.

Spring

Spring was the only season in which robust, FDR-corrected relationships were detected between physicochemical parameters and both microbial and viral groups. Significant correlations were clustered into three broad categories: solids content, wastewater temperature, and biodegradable organic load.

Methylotrophic methanogens of the Methanomassiliicoccaceae family showed a strong positive correlation with temperature ($\rho = 0.77$, $q = 0.0045$). Several SRB families, including Desulfohalobiaceae, Desulfonatronaceae ($\rho = 0.77$, $q = 0.0045$), and

Desulfovibrionaceae ($\rho = 0.72$, q = 0.0125), also correlated positively with the temperature of wastewater. Among fermentative bacteria, Prevotellaceae, Porphyromonadaceae, Bacterioidaceae, Barnesiellaceae exhibited strong positive temperature correlations (ρ = 0.76, q = 0.0045), alongside Carnobacteriaceae (ρ = 0.74, q = 0.0077) and Coriobacteriaceae ($\rho = 0.68$, q = 0.0327). Collectively, these findings indicate coordinated temperature-linked co-variation across multiple trophic levels during spring.

Several viral families exhibited significant negative correlations with solids. Autographiviridae showed strong negative correlations with both VS (ρ = –0.75, q = 0.0011) and TS (ρ = –0.79, q = 8.5 x 10^{-4}). Schitoviridae were negatively correlated with VS (ρ = –0.66, q = 0.0089) and TS (ρ = –0.69, q = 0.0052), while Herelleviridae were negatively correlated with VS (ρ = –0.56, q = 0.0465). These patterns show that several phage families were more abundant in samples with lower solids content.

A broad group of viral families showed strong positive associations with wastewater temperature, with Spearman coefficients ranging from $\rho = 0.74$ to 0.77 (q = 8.5x10^{-4} to 0.0013). These included Autographiviridae, Myoviridae, Podoviridae, Ackermannviridae, Demecriviridae, Herelleviridae, Schitoviridae, and Siphoviridae, indicating consistent co-variation between viral abundance and the seasonal temperature gradient during spring.

In addition, Ackermannviridae exhibited a significant positive correlation with BOD_5 (ρ = 0.56, q = 0.0465), indicating association with samples characterised by higher biodegradable organic load.

Summer

In summer, correlation heatmaps showed several visually structured patterns involving SRB, methanogens, and viral families. However, none of these associations remained statistically significant following FDR correction. Statistically, summer represents a period in which correlations between abiotic parameters and microbial or viral abundances were weak or non-detectable.

Autumn

Autumn samples also exhibited visually structured heatmap patterns, particularly involving solids, COD, pH, SRB, and several viral families. As in summer, however, no correlations reached FDR-adjusted significance. Consequently, autumn correlations are presented descriptively only and are not interpreted further.

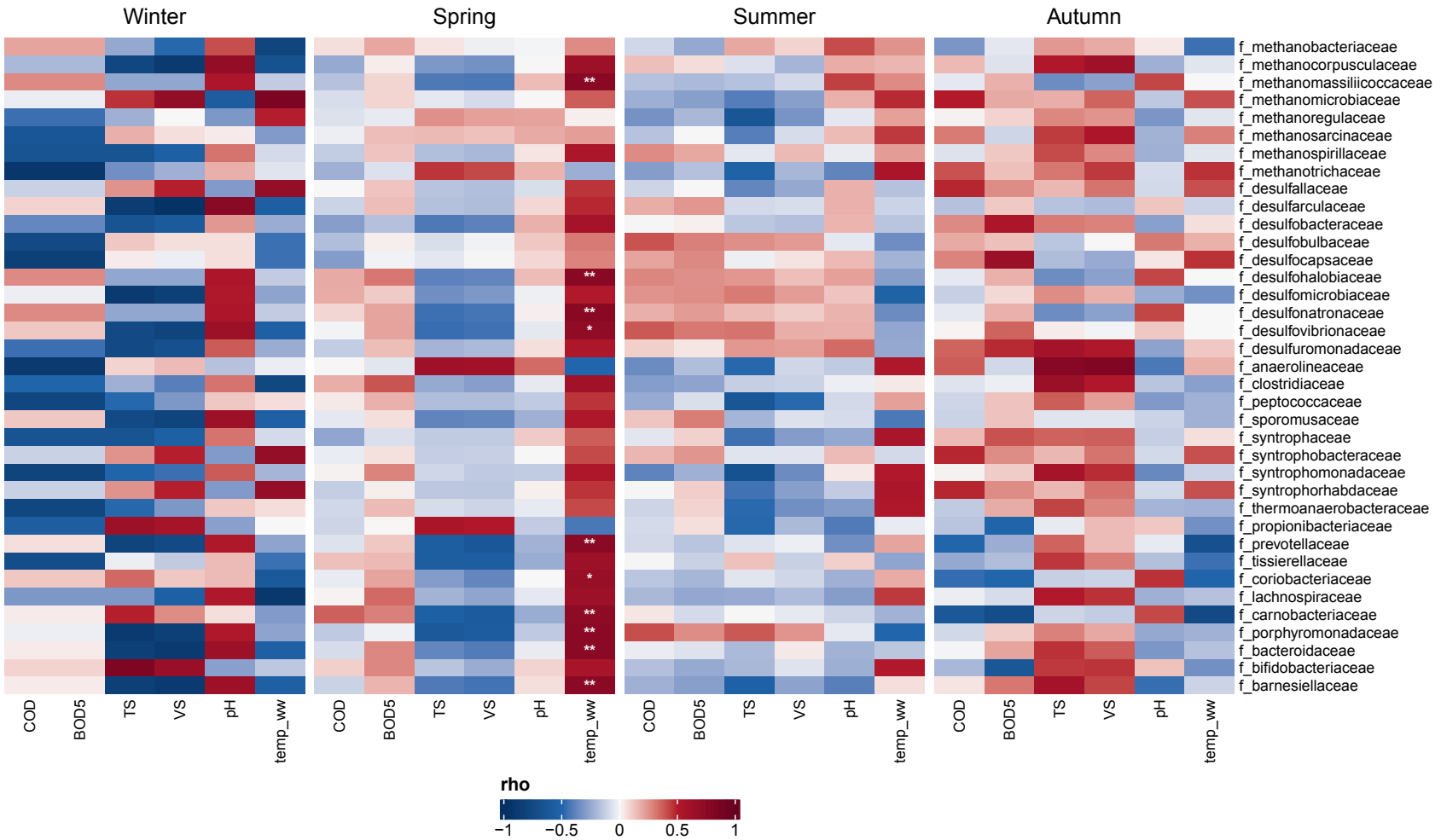

Figure 6.27: Spearman correlations between prokaryotic functional guilds and physicochemical wastewater parameters across all ten monitored septic tanks. FDR-corrected significance levels are indicated as: q $\leq$ 0.05 (*), q $\leq$ 0.01 (**), q $\leq$ 0.001 (***).

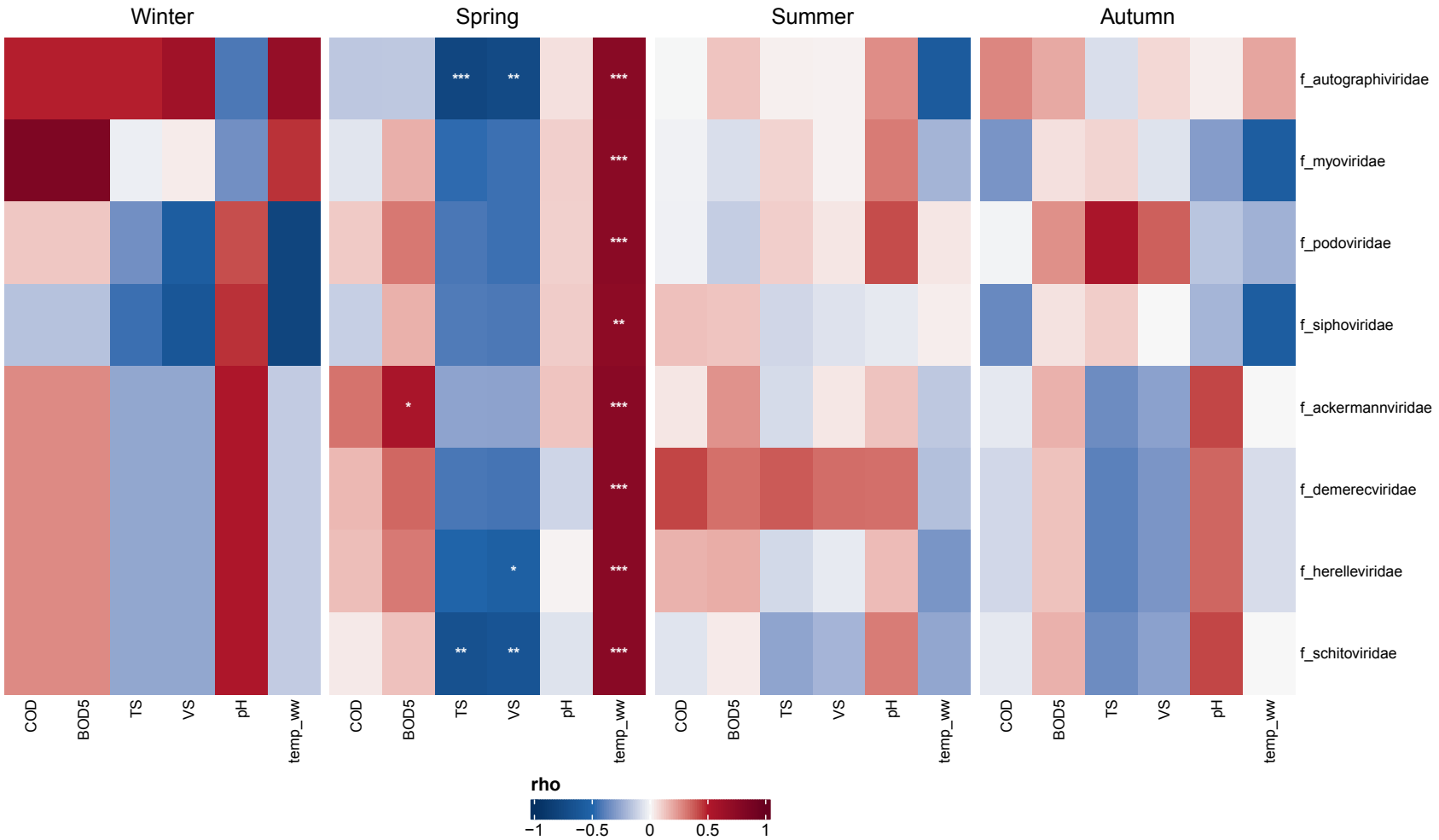

Figure 6.28: Spearman correlations between phage families and physicochemical wastewater parameters across all ten monitored septic tanks. FDR-corrected significance levels are indicated as: q $\leq$ 0.05 (*), q $\leq$ 0.01 (**), q $\leq$ 0.001 (***).

6.5.3 Implications

The seasonal correlation analyses presented in this section demonstrate that statistically robust associations among abiotic factors, viral families, and anaerobic functional guilds in household septic tanks are strongly season dependent. Across the annual cycle, statistically significant correlations between microbial or viral groups and physicochemical parameters were detected exclusively during spring, whereas associations observed in winter, summer, and autumn did not withstand correction for multiple testing. These results indicate that periods of statistically detectable microbial–viral–abiotic co-variation are temporally constrained under *in situ* field conditions.

Recent studies have highlighted the potential importance of viral communities in anaerobic digestion systems. For example, Fan et al. (2023) reported that viral variables explained a substantial proportion of prokaryotic community variation relative to measured abiotic parameters, suggesting that viral contributions to prokaryotic community structuring may be underrepresented when environmental drivers are considered alone.

Similar reductions in the number and diversity of statistically reported phage–host and microbial– or virus–environment associations under low-temperature or low-activity conditions have been reported in engineered anaerobic systems and wastewater environments (Zhang et al., 2017; Wang et al., 2024). The present study extends these observations by documenting seasonally constrained correlation structures in operational household septic tanks monitored across four seasons, with wastewater temperatures ranging from 4 °C to 15 °C, a decentralised sanitation context in which such seasonal contraction of statistically supported correlation patterns has not previously been described.

6.6 Seasonal variation in functional marker genes associated with key anaerobic pathways in household septic tank samples

Distinct seasonal patterns were observed across functional marker genes representing the principal anaerobic digestion pathways involved in AD, including acetate activation,

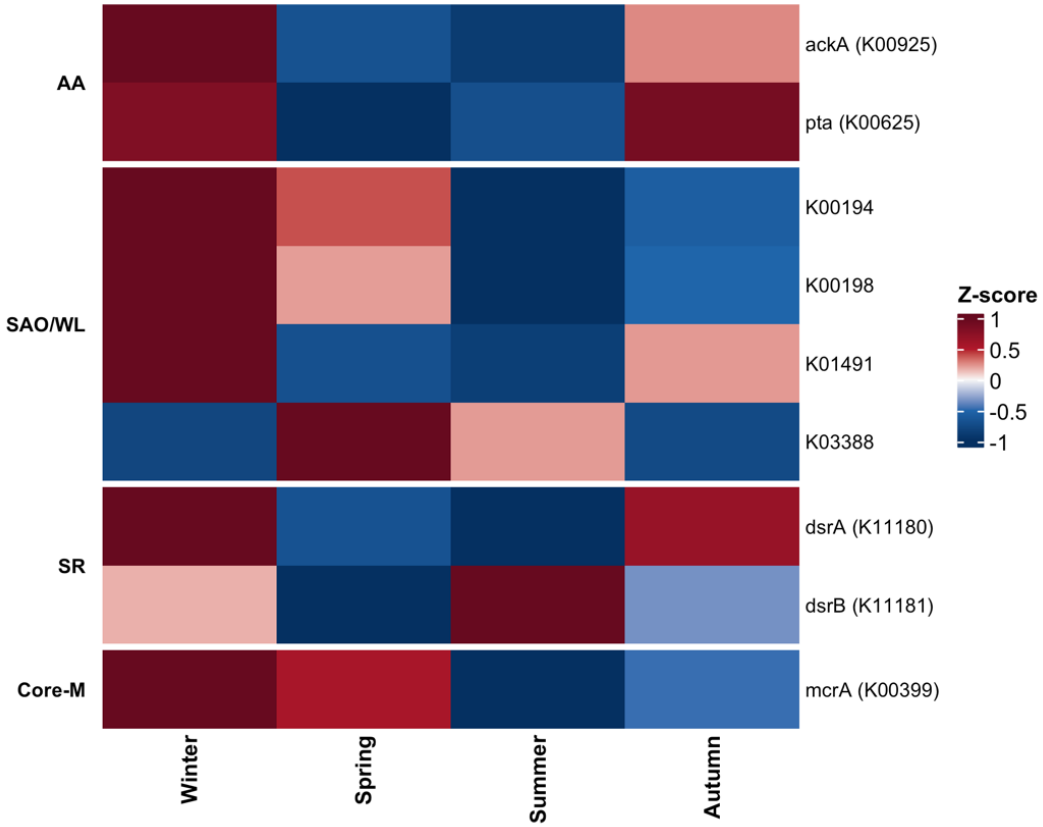

Figure 6.29: Seasonal variation in functional marker genes associated with key AD pathways across all ten monitored household septic tanks.

WL pathway–associated syntrophic acetate oxidation, core methanogenesis, and sulphate reduction (Figure 6.29; Figure D.3 in Appendix D). Gene abundances were $\log_{10}$-transformed and subsequently z-scored within each gene following Fan et al. (2023). Accordingly, the heatmap illustrates relative seasonal enrichment and depletion rather than absolute gene abundance, facilitating comparison of seasonal trends across functional groups independent of differences in baseline gene abundance.

6.6.1 Acetate activation

Markers associated with acetate activation, *ackA* (K00925) and *pta* (K00625), showed clear seasonal variation in relative abundance. Both genes were relatively enriched during winter, followed by lower relative abundance in spring and summer and partial recovery in autumn. This seasonal pattern suggests a shift in the relative representation of bacterial acetate activation genes across the annual cycle. The winter enrichment of *ackA* and *pta* suggests increased representation of acetate-activating bacterial pathways during colder period, based on gene-level abundance patterns.

6.6.2 Wood–Ljungdahl pathway–associated C_1 metabolism

Genes associated with acetate-centred C_1 metabolism via the WL pathway (K00194, K00198, K01491, K03388) exhibited pronounced seasonal variation in relative abundance. Several WL-associated genes were relatively enriched during winter and spring, whereas summer samples were characterised by consistently lower relative abundance across most genes in this pathway. Among these markers, K03388 showed particularly strong enrichment during spring.

Overall, these results demonstrate clear seasonal structuring in the representation of WL pathway genes, with higher relative abundance during cold and transitional seasons and reduced representation during summer. This pattern is consistent with enhanced acetate-centred metabolism under low-temperature conditions, where homoacetogenesis and related C_1 transformations are favoured. While these genes may also participate in syntrophic acetate oxidation under specific system constraints, their presence alone does not distinguish between homoacetogenic and syntrophic pathways, and therefore indicates acetate-centred metabolism rather than definitive SAO activity.

6.6.3 Core methanogenesis

The core methanogenesis marker *mcrA* (K00399) also displayed clear seasonal variation in relative abundance. *mcrA* was most abundant during winter and spring, followed by a decline during summer and autumn. This pattern indicates seasonal differences in the representation of methanogenesis-associated functional potential within the microbial community.

6.6.4 Sulphate reduction

Sulphate reduction markers, *dsrA* (K11180) and *dsrB* (K11181), exhibited marked but distinct seasonal patterns. *dsrA* showed higher relative abundance during winter and autumn, whereas *dsrB* reached its highest relative abundance during summer. This divergence indicates gene-specific seasonal structuring within sulphate reduction functional repertoire. These patterns describe differences in the relative representation of sulphate reduction–related genes across seasons, without implying differences in sulphate reduction rates or process intensity.

6.6.5 Implications

Collectively, the seasonal profiles of functional gene markers demonstrate that the functional potential of anaerobic metabolism in household septic tanks is seasonally structured rather than static, varying systematically across the annual cycle. Differences in the relative abundance of markers associated with AA, SAO/WL, sulphate reduction, and methanogenesis reflect seasonal reorganisation of the genetic repertoire underpinning key anaerobic processes.

Winter samples were characterised by higher relative representation of genes associated with acetate activation, sulphate reduction, and methanogenesis, whereas summer samples showed lower relative abundance across multiple anaerobic functional markers. Transitional seasons, particularly spring, exhibited more heterogeneous functional gene profiles, with concurrent enrichment of selected markers across multiple pathways. These patterns mirror the seasonal restructuring observed at the taxonomic level and indicate that functional gene composition varies in parallel with microbial community composition. While Wang et al. (2022) proposed mechanistic roles for viruses in methanogenic and sulphate-reducing systems based on viral genome content, the present study provides complementary, field-based evidence of seasonally resolved co-variation between viral communities and anaerobic functional guilds in household septic tanks.

Taken together, these findings show that seasonal environmental forcing is associated not only with shifts in microbial and viral community structure but also with changes in the distribution of functional genes linked to key anaerobic pathways in septic tank systems. This seasonal organisation of functional potential provides essential context for interpreting microbial and viral dynamics under real-world operating conditions, while highlighting the importance of seasonality when assessing anaerobic processes in cold-climate on-site sanitation systems.

6.7 Seasonal structuring of pathogen-associated communities in septic tanks

6.7.1 Seasonal variation in relative abundance

Pathogen-associated bacterial families exhibited pronounced seasonal restructuring, with clear shifts in dominance across the year (Figure 6.30). Winter communities were strongly dominated by Moraxellaceae, alongside contributions from Pseudomonadaceae and Arcobacteraceae, while other families were present at comparatively low relative abundances. In spring, the pathogen-associated communities became more evenly distributed, with increased contributions from Pseudomonadaceae, Arcobacteraceae, and Aeromonadaceae, and a marked decline in Moraxellaceae. During summer, the community shifted further towards dominance by Arcobacteraceae and Pseudomonadaceae, accompanied by notable contributions from Bacteroidaceae and Aeromonadaceae. By autumn, the composition shifted again, with a substantial increase in Pseudomonadaceae, which became the dominant family, while Arcobacteraceae and Moraxellaceae declined relative to their earlier seasonal peaks. Families such as Mycobacteriaceae, Treponemataceae, Neisseriaceae, and Vibrionaceae persisted at low relative abundances across all seasons. Additional information on the reported characteristics of these families is provided in Table G.1 in Appendix G.

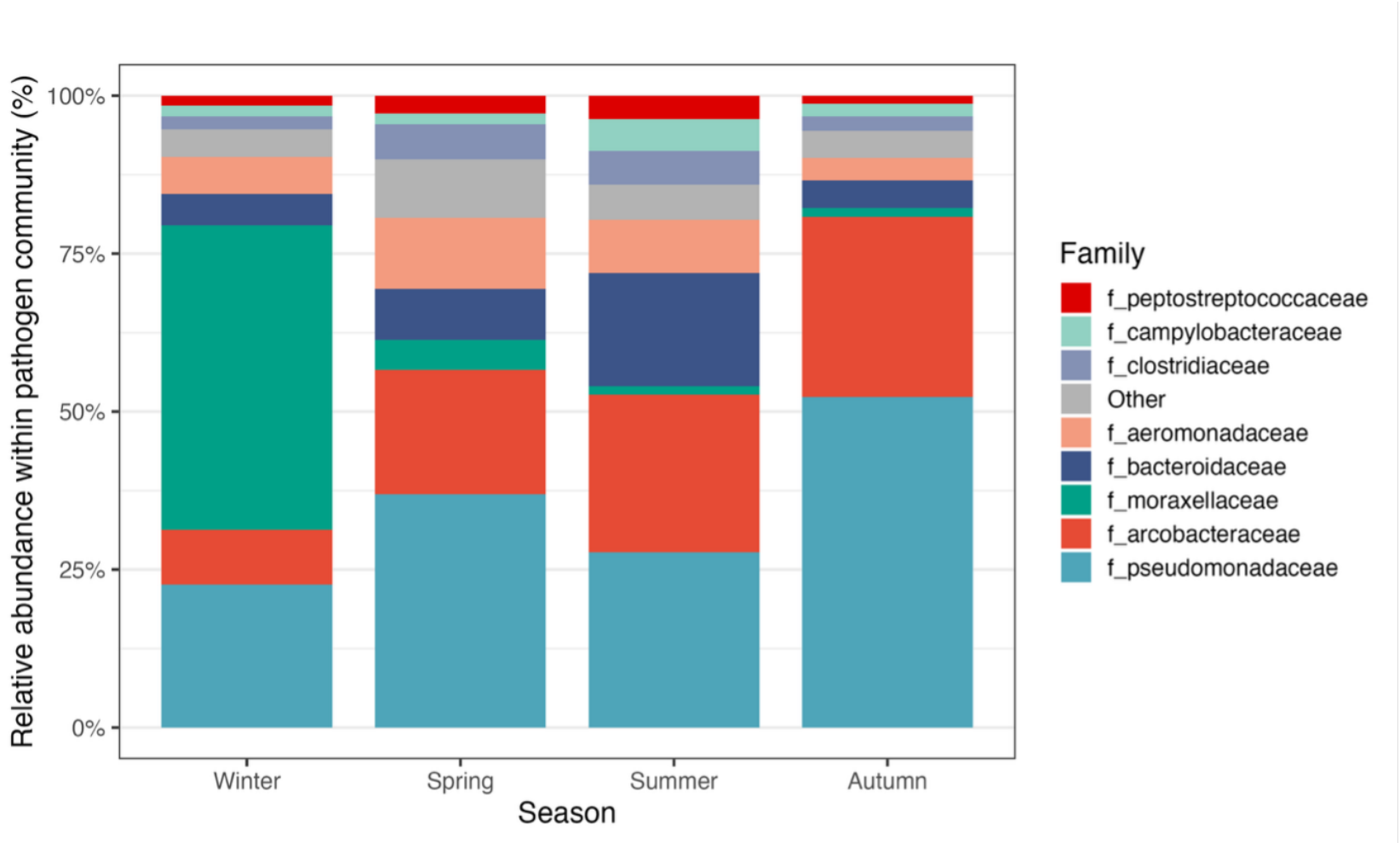

Figure 6.30: Seasonal relative abundance of dominant pathogenic families (100%-stacked) across all ten household septic tanks

6.7.2 Implications for environmental and public health risk

The consistent detection of pathogen-associated bacterial families across all seasons reflects sustained influent inputs of faecal-derived microorganisms. This pattern indicates that septic tank systems act as selective ecological environments in which microbial communities are differentially retained and restructured, rather than fully attenuated. Accordingly, the presence of these taxa is interpreted as the result of continuous input combined with environmentally mediated selection processes within the system.

The comparatively low relative abundances of multiple pathogen-associated families during winter do not indicate reduced input but are consistent with temperature-constrained microbial dynamics. Under low-temperature conditions, reduced metabolic activity and slower growth rates limit the proliferation of many taxa, resulting in a community structure characterised by reduced diversity and the dominance of psychrotolerant groups such as Moraxellaceae. Other families remain detectable at low relative abundances, reflecting constrained growth and competitive capacity rather than absence.

In contrast, the increased diversity and redistribution of pathogen-associated families during warmer seasons are consistent with enhanced microbial activity and more complex community structuring. The greater relative contributions of families such as Arcobacteraceae, Pseudomonadaceae, and Aeromonadaceae correspond to conditions that support a broader range of taxa commonly associated with wastewater and faecal environments. These seasonal patterns therefore reflect temperature-associated changes in community composition rather than variation in influent characteristics.

The detection of families such as Vibrionaceae and Mycobacteriaceae indicates that groups known to include pathogenic species are present in the system. However, because the analysis is based on family-level classification, it is not possible to determine whether specific pathogenic species are present. Therefore, these results indicate potential rather than confirmed health relevance.

Taken together, these results indicate that septic tank systems do not fully remove pathogen-associated microorganisms but are associated with changes in their relative abundance and community structure across seasons. As such, septic tank effluent may represent a continuous source of microbially derived inputs to the surrounding

environment, with patterns that vary in relation to temperature-dependent microbial dynamics.

6.8 iPHoP-based phage–host predictions in the second household septic tank

The iPHoP-based host predictions provide a seasonally resolved, genome-informed overview of which microbial taxa are putatively associated with viral populations in septic tank 2. These results enable interpretation of host identity and guild-level structure of predicted virus–host associations, but do not provide information on infection rates, lytic versus lysogenic replication, or AD reaction kinetics (Roux, 2022). Accordingly, the results discussed here are restricted to patterns of host association and seasonal structure, rather than direct functional or mechanistic effects (Roux, 2022). The seasonal relative abundances of all anaerobic functional guilds for septic tank 2 are presented in Figures D.4 - D.7 in Appendix D.

6.8.1 Phage–host predictions

Fermentative bacteria

Fermentative bacteria displayed the highest host-family diversity in winter and autumn, with predicted phage–host links distributed across multiple families (Figure 6.31). During spring and summer, fermenter-associated predictions were strongly dominated by Bacteroidaceae, indicating a pronounced seasonal shift towards host–family dominance under warmer conditions.

Methanogens

Methanogenic archaea exhibited highly constrained phage–host associations (Figure 6.32). Predictions were limited to winter and autumn, with dominance by obligate acetoclastic methanogen Methanotrichaceae in winter and facultative acetoclastic methanogen Methanosarcinaceae in autumn. No methanogen-associated phage–host links were detected in spring or summer, and no additional methanogenic families were identified across seasons.

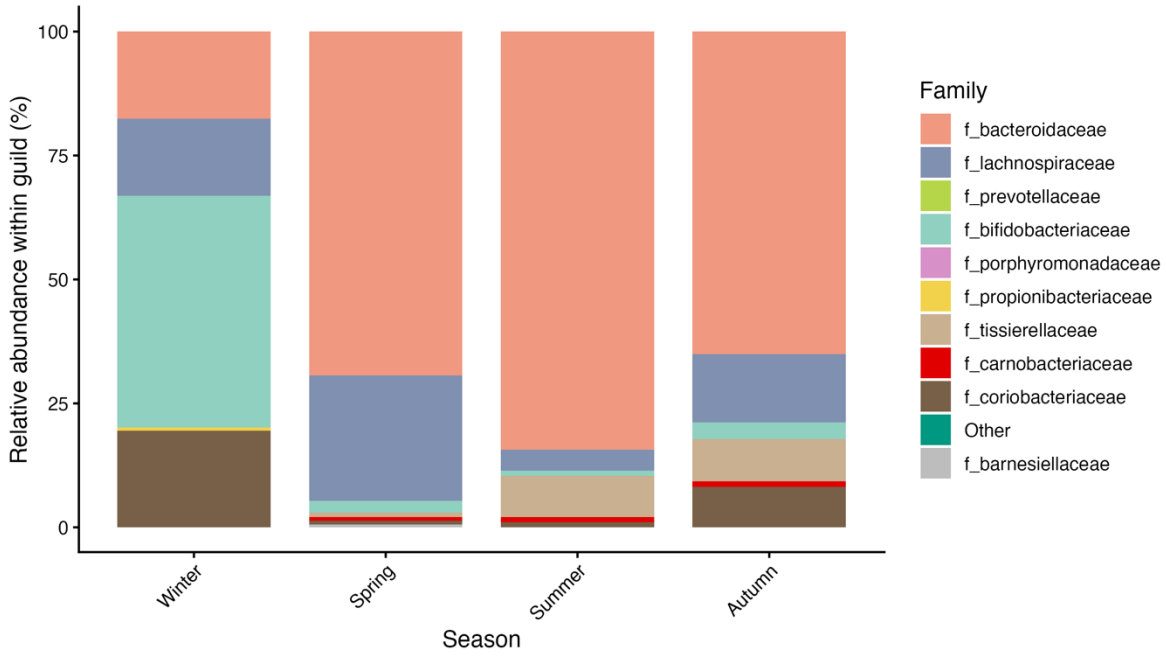

Figure 6.31: Seasonal distribution of phage–host associations for fermentative guilds in Septic Tank 2. 100% stacked bar plots showing the relative seasonal composition of predicted phage–host associations inferred using iPHoP.

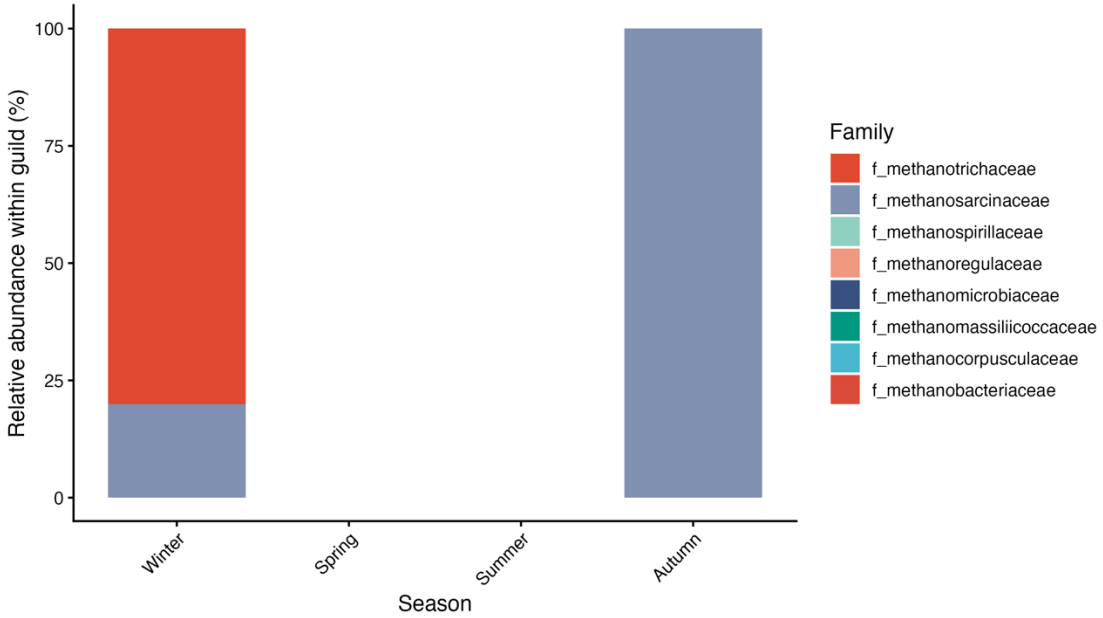

Figure 6.32: Seasonal distribution of phage–host associations for methanogenic guilds in Septic Tank 2. 100% stacked bar plots showing the relative seasonal composition of predicted phage–host associations inferred using iPHoP.

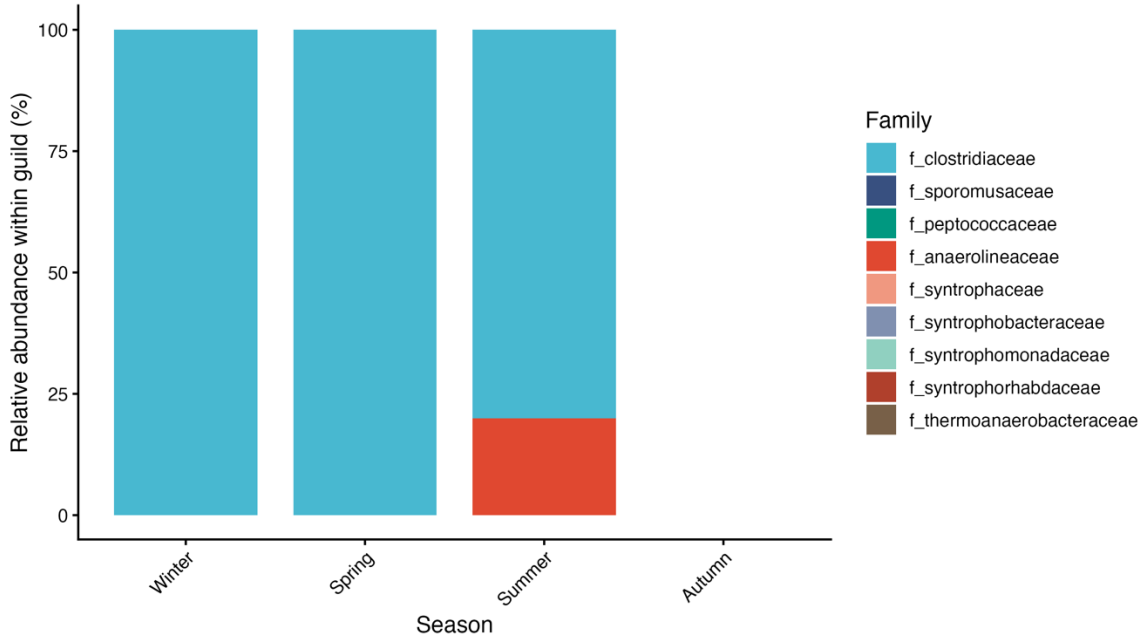

Figure 6.33: Seasonal distribution of phage–host associations for syntrophic guilds in Septic Tank 2. 100% stacked bar plots showing the relative seasonal composition of predicted phage–host associations inferred using iPHoP.

Syntrophic acetate-oxidising bacteria

SAOB showed strongly restricted phage–host associations dominated by Clostridiaceae in winter, spring, and summer, with a minor contribution from Anaerolineaceae in summer (Figure 6.33). No SAOB-associated phage–host predictions were detected in autumn.

Sulphate-reducing bacteria

SRB exhibited pronounced seasonality. In spring, all predicted SRB-associated phage–host links were assigned to Desulfovibrionaceae (Figure 6.34). In summer, SRB-associated predictions diversified substantially, with Desulfomicrobiaceae becoming dominant alongside contributions from Desulfobulbaceae, Desulfovibrionaceae, and other SRB families. No SRB-associated phage–host predictions were detected in winter or autumn.

Across all functional guilds, sustained dominance by a single host family across all seasons was not observed. Instead, predicted phage–host associations showed strong temporal filtering, with shifts in host-family composition occurring between seasons and differing markedly among microbial guilds.

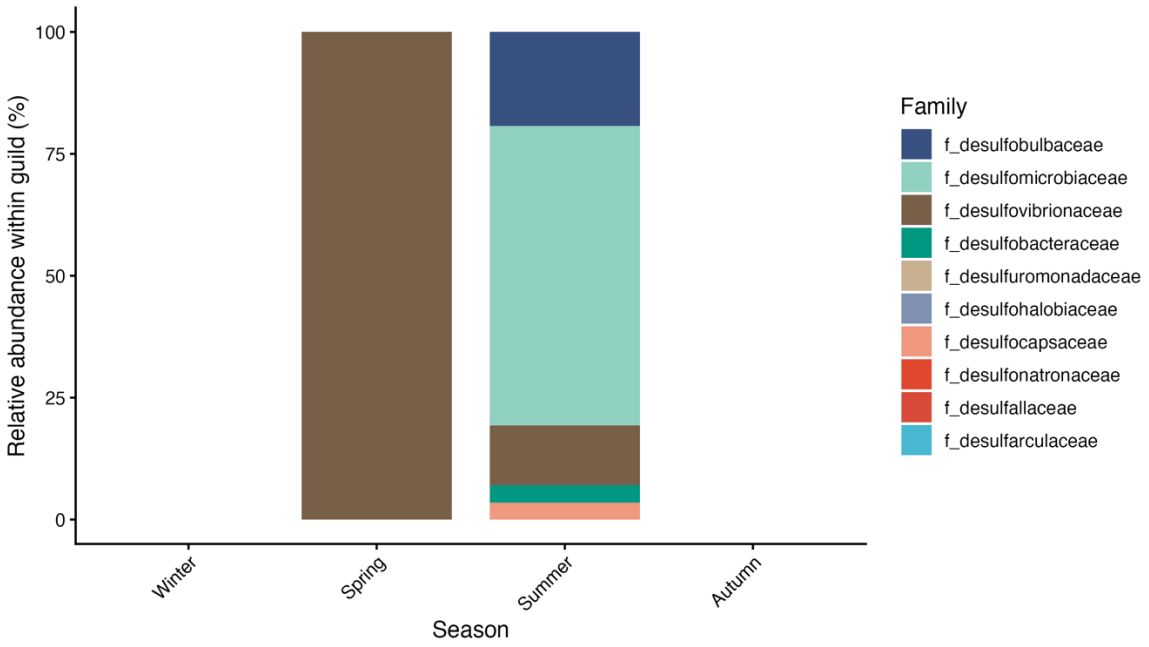

Figure 6.34: Seasonal distribution of phage–host associations for sulphate-reducing guilds in Septic Tank 2. 100% stacked bar plots showing the relative seasonal composition of predicted phage–host associations inferred using iPHoP.

6.8.2 Implications

The integration of seasonal guild composition with iPHoP-based phage–host predictions demonstrates that virus–prokaryote interactions in household septic tank 2 are structured by host functional dominance and seasonal stability, rather than by continuous lytic control. Across AD guilds, phage associations were most consistently detected for fermentative bacteria, whereas methanogenic archaea, SAOB, and SRB exhibited more restricted, intermittent, or seasonally absent predicted infections. This pattern has important implications for how microbial stability and functional resilience are maintained in cold-climate septic systems.

The persistence of phage associations with fermentative guilds across all seasons suggests that viruses are preferentially linked to abundant, metabolically central, and compositionally stable host populations. Fermenters dominate organic matter hydrolysis and acidogenesis and maintain high relative abundance throughout the year. Their consistent representation among predicted infected hosts is therefore most consistent with a Piggyback-the-Winner–type interaction regime, in which viruses are maintained predominantly through lysogenic or low-impact associations with dominant hosts, rather than through strong lytic regulation (Zhang et al., 2017; Fan et al., 2023; Castledine

and Buckling, 2024). This interpretation aligns with recent observations from anaerobic digestion systems showing that viruses frequently track host abundance and metabolic activity rather than driving host collapse (Zhang et al., 2017; Fan et al., 2023).

In contrast, methanogenic archaea exhibited highly seasonal and taxonomically restricted phage associations, despite clear seasonal shifts in methanogenic community composition and functional gene markers. Notably, acetoclastic methanogens were detected among predicted infected hosts in autumn and winter, despite their generally low abundance. This finding extends previous literature syntheses reporting little or no evidence for acetoclastic methanogens in septic tanks (Shaw and Dorea, 2021), by demonstrating that these taxa can persist and engage in virus–host interactions under *in situ* cold-climate field conditions, even if they do not dominate methanogenic flux. Nevertheless, the absence of widespread phage associations across methanogenic taxa and seasons suggests that methanogens in this system are not primarily regulated by lytic Kill-the-Winner dynamics.

Similarly, SAOB and SRB guilds, although seasonally dominant at the taxonomic and functional levels (Figures D.6 - D.7), showed limited and discontinuous predicted phage infections (Figures 6.33 - 6.34), particularly during autumn. This decoupling between guild dominance and detectable phage association suggests that these groups may operate under physiological or ecological conditions that reduce viral detectability, such as slow growth rates, tight syntrophic coupling, or prolonged lysogenic states. Importantly, the absence of predicted infected hosts in autumn does not imply the absence of viruses, but rather indicates a transition toward a more stable, low-turnover viral ecological regime, in which viral–host interactions are weakly expressed and therefore less readily captured by correlation- and prediction-based approaches.

Across guilds, the seasonal restriction of detectable phage–host associations to specific periods suggests that viral impacts on anaerobic communities are temporally gated, becoming most apparent during phases of environmental or community reorganisation rather than during steady-state operation. This pattern is consistent with recent work indicating that viral influences on prokaryotic communities may be underestimated when analyses rely solely on abiotic drivers or assume constant interaction strength (Fan et al., 2023). In cold-climate septic tanks, viruses appear to contribute to community persistence

and modulation, rather than exerting strong, continuous top-down control.

The principal novelty of this study lies in providing the first seasonally resolved, field-based application of genome-informed phage–host prediction to an operational household septic tank. Existing virus–host interaction frameworks and empirical evidence are derived largely from laboratory studies or engineered anaerobic digesters, whereas this work extends these concepts to a decentralised on-site sanitation system operating under real-world seasonal forcing. By combining year-round monitoring, functional gene profiling, and iPHoP-based host prediction, it demonstrates that virus–prokaryote interactions in on-site sanitation systems are guild-specific, seasonally constrained, and predominantly lysogeny-biased, rather than governed by classical lytic control frameworks derived from laboratory or high-rate engineered systems.

These findings advance current understanding by showing that viral ecology must be interpreted within the context of seasonal stability, host functional dominance, and low-temperature operation, which together shape microbial resilience and functional continuity in decentralised sanitation infrastructure.

6.9 Summary of key findings

This chapter provides a year-round, field-based assessment of seasonal dynamics in wastewater physicochemical characteristics, microbial and viral community structure, anaerobic functional potential, pathogen-associated communities, and predicted phage–host associations in cold-climate household septic tanks. Field monitoring, community profiling, and functional gene analyses were conducted across all ten operational household septic tanks, while genome-informed phage–host predictions using iPHoP were performed for a representative system (septic tank 2) to enable high-resolution host–virus interpretation.

Field measurements revealed pronounced seasonal fluctuations in COD, BOD_5, TS, and VS across all ten septic tanks, with solids accumulation peaking in spring and declining through summer and autumn. Wastewater temperatures were partially buffered relative to ambient conditions but nonetheless exerted a strong influence on system performance. Microbial and viral community analyses across all ten septic tanks exhibited clear seasonal restructuring. Summer samples were compositionally more uniform across

sites, whereas spring displayed the greatest within-season variability, consistent with transitional environmental conditions. Beta-diversity analyses revealed strong coupling between prokaryotic and viral community dynamics, and ordination analyses confirmed distinct seasonal clustering for both domains. Statistically significant correlations between physicochemical parameters and viral and prokaryotic communities were largely confined to spring, indicating that detectable microbial-viral-environment coupling is temporally constrained and most pronounced during periods of seasonal transition.

Seasonal restructuring was also evident at the level of anaerobic functional groups across the ten septic tanks. Marker genes associated with acetate activation, Wood–Ljungdahl pathway–associated acetate-centred metabolism, core methanogenesis, and sulphate reduction exhibited distinct but coordinated seasonal patterns, with generally higher relative representation during winter and transitional spring conditions and reduced representation during summer. These trends demonstrate that functional gene composition reorganises seasonally in parallel with microbial community structure under field conditions.

Pathogen-associated bacterial communities showed consistent presence across all seasons with temperature-dependent restructuring. Despite this sustained input, clear seasonal shifts in relative abundance and dominance patterns were observed. Winter communities were characterised by reduced diversity and selective dominance of psychrotolerant taxa, whilst warmer seasons supported a broader and more evenly distributed assemblage of pathogen-associated families. These patterns indicate that temperature-dependent microbial processes regulate the proliferation and relative abundance of these taxa rather than their presence or absence. The detection of families that include microorganisms associated with waterborne and opportunistic infections indicates that septic tank effluent may act as a continuous, seasonally variable source of microbial inputs to the surrounding environment. However, as the analysis is based on family-level classification, the presence of specific pathogenic species cannot be confirmed. These findings therefore indicate potential rather than demonstrated pathogenic risk.

Targeted iPHoP-based phage–host predictions for septic tank 2 revealed guild-specific and seasonally constrained patterns of predicted phage infection, with fermentative bacteria exhibiting the most persistent associations across seasons, while SRB, SAOB,

and methanogenic archaea showed more restricted, intermittent, or seasonally absent predicted patterns. These findings indicate that phage–host interactions in septic tanks are structured by host functional dominance and seasonal stability, rather than by continuous lytic regulation.

Collectively, the septic tank results demonstrate that temperature-driven seasonal dynamics govern physicochemical conditions, microbial and viral community composition, anaerobic functional potential, pathogen-associated community structuring, and phage–host interactions in cold-climate septic tank systems. By combining multi-tank field monitoring with a genome-informed, tank-specific investigation of phage–host interactions, this study provides the first seasonally resolved, field-based framework linking viral ecology to anaerobic functional guilds in decentralised sanitation infrastructure, offering new insight into microbial stability and functional resilience under cold-climate operating conditions.

7 Discussion

7.1 Overview of key findings

This study provides an integrated, field-based examination of how seasonal temperature variation governs physicochemical conditions, microbial and viral community dynamics, anaerobic functional pathways, and phage–host interactions in household septic tanks operating under cold-climate conditions. To the author's knowledge, this is the first study to combine stakeholder perceptions, year-long *in situ* monitoring, laboratory-scale reactor experiments, shotgun metagenomics, functional gene analysis, and genome-informed phage–host prediction to explain cold-season septic tank underperformance under real operating conditions. By linking user-reported failures to microbial and viral processes, the research moves beyond descriptive surveys and laboratory inference to establish mechanistic explanations relevant to decentralised sanitation systems.

Stakeholder surveys consistently identified winter as the most problematic operational period, with recurrent reports of odour emissions, wastewater backflow, increased solids accumulation, and heightened desludging frequency. These perceptions are strongly supported by field observations and biological data generated in this study. Although wastewater temperatures within septic tanks were partially buffered from ambient extremes, seasonal variation within the tanks (approximately 4 °C to 15 °C) was sufficient to induce pronounced shifts in microbial community structure and functional potential, resulting in measurable changes in treatment performance, solids accumulation, and microbial risk profiles. Functional gene profiles showed that relatively modest temperature changes were associated with substantial reorganisation of acetate activation, syntrophic acetate oxidation, methanogenesis, and sulphate reduction pathways. This provides new field-based evidence that anaerobic digestion in septic tanks responds to temperature through pathway-level restructuring rather than uniform suppression or enhancement, challenging assumptions commonly derived from laboratory-scale or high-rate engineered systems (Angelidaki and Ahring, 1993; Luostarinen et al., 2007; Almansa et al., 2023). Taken together, these findings demonstrate that septic tanks in cold climates do not

operate as steady-state systems, but rather as seasonally dynamic anaerobic environments governed by temperature-dependent biological processes.

7.2 Stakeholder perceptions and system-level constraints

Stakeholder responses differed meaningfully across stakeholder groups in ways that reflect their distinct roles, responsibilities, and operational realities within the sanitation service chain. Institutional actors primarily identified governance-related constraints, including limited enforcement of construction standards and weak coordination across responsible authorities, indicating systemic gaps in regulation and oversight. In contrast, operational actors emphasised practical limitations associated with system access and maintenance under cold conditions, while households experienced the direct consequences of these combined constraints in the form of reduced usability and reliability. The role of NGOs as intermediaries highlights the importance of knowledge transfer and support mechanisms, although their effectiveness is constrained by limited technical capacity and integration within formal governance structures. Taken together, this differentiation demonstrates that system underperformance is not driven by user behaviour alone, but arises from the combined effects of design deficiencies, operational constraints, and institutional limitations. This challenges the simplified narratives that attribute failure primarily to poor maintenance practices and instead highlights the importance of addressing upstream structural factors.

The stakeholder findings from Kazakhstan are consistent with evidence from other cold-climate and decentralised sanitation contexts, indicating that the observed challenges are not unique to the study area but reflect broader environmental and operational constraints. In Mongolia, residents similarly identify sanitation as a primary concern, shaped by everyday experiences of inadequate and seasonally constrained facilities, while demonstrating high awareness of associated risks and a willingness to improve conditions (Sigel, 2010). This suggests that unsafe practices are not driven by lack of knowledge, but by structural limitations and absence of viable alternatives (Sigel, 2010). Across Central Asia, direct evidence on stakeholder perceptions is limited, with available data focusing primarily on system conditions. Nevertheless, the scale and visibility of infrastructure degradation, service gaps, and environmental risks imply that sanitation is widely experienced as a persistent and systemic problem. The continued prioritisation of water

supply over sanitation further suggests that, although deficiencies are recognised, they remain insufficiently addressed, reinforcing a pattern in which awareness does not translate into effective action (Vinokurov et al., 2024).

In contrast, evidence from high-income Nordic countries, including Finland, Sweden, and Norway, indicates that even where regulatory frameworks are well established, stakeholder perceptions are characterised by uncertainty, financial concern, and low prioritisation (Laukka et al., 2022). Property owners often perceive on-site sanitation as a costly and burdensome requirements, exacerbated by limited understanding of environmental benefits, leading to postponement of system upgrades and reduced compliance (Laukka et al., 2022). At the institutional level, authorities recognise the importance of managing diffuse pollution but frequently perceive sanitation as a lower strategic priority, constrained by limited resources, administrative burden, and competing policy objectives (Laukka et al., 2022). More broadly, both service providers and the public reflect a perception that on-site sanitation lacks policy relevance, with concerns regarding professional competences and effectiveness of existing approaches further undermining confidence in the system (Laukka et al., 2022).

Taken together, these perceptions across diverse contexts reveal a consistent pattern: sanitation challenges are widely recognised by stakeholders, but are not translated into effective action due to structural, institutional, and economic constraints. This convergence highlights that underperformance in decentralised sanitation systems is not primarily a consequence of user behaviour or awareness, but rather systemic limitations in governance, design, and implementation, particularly under the added pressures of cold-climate conditions.

7.3 Mechanisms governing total solids removal in anaerobic systems

Total solids removal in anaerobic treatment systems is governed by the interplay of biological degradation, physical sedimentation, and their combined sensitivity to temperature (Tchobanoglous et al., 2003; Metcalf and Eddy, 2013). Understanding these mechanisms is essential for interpreting seasonal variation in septic tank performance, particularly under cold-climate conditions where both biological and physical processes

are simultaneously constrained.

Temperature exerts a direct and well-documented influence on hydrolysis kinetics, broadly following an Arrhenius-type relationship across the psychrophilic and mesophilic range (Veeken and Hamelers, 1999). As temperatures decline, enzymatic reaction rates fall, slowing the conversion of particulate substrates to soluble intermediates and reducing the overall rate of organic solids degradation (Veeken and Hamelers, 1999). This effect is compounded downstream: reduced hydrolysis rates constrain substrate availability for acetogenic and methanogenic populations, suppressing syntrophic partnerships and further limiting the mineralisation of organic matter. The net result is incomplete degradation of particulate material, accumulation of partially hydrolysed solids, and lower VS and TS removal efficiency – a pattern consistent with the seasonal performance variation observed in cold-climate septic tank studies (Luostarinen et al., 2007; Leblanc et al., 2019).

Physical sedimentation provides a parallel and partially independent mechanism of TS removal (Sperling, 2007; Ghangrekar, 2022). In septic tanks, gravitational settling of suspended solids is the primary physical removal pathway, and its efficiency is strongly influenced by temperature through the effect on fluid viscosity (Syafira et al., 2021). Water viscosity increases substantially as temperature falls, reducing the settling velocity of suspended particles and prolonging the time required for solids to sediment out of the water column (Alisawi, 2020; Ghangrekar, 2022). At the lower end of the temperature range documented in this study (4 °C), viscosity-related reductions in settling efficiency contribute independently to elevated suspended solids concentrations in the effluent, compounding the effects of reduced biological hydrolysis (Alisawi, 2020; Ghangrekar, 2022). Conversely, at higher temperatures, reduced viscosity facilitates faster particle settling, and microbial production of extracellular polymeric substances promotes floc formation, further enhancing aggregation and sedimentation of fine suspended material (Fitzpatrick et al., 2004; Alisawi, 2020).

The interaction between biological and physical mechanisms means that temperature-driven reductions in TS removal are not attributable to either process alone but arise from their simultaneous suppression under cold conditions. This coupling explains why TS, VS, COD, and BOD_5 removal exhibit similar seasonal trajectories in

laboratory-scale reactors and septic tank systems: all are ultimately governed by the rate of biological hydrolysis and the efficiency of physical solids separation, both of which decline with temperature.

7.4 Microbial ecology and seasonal dynamics

The laboratory-scale reactor experiments and field monitoring datasets together provide complementary and mutually reinforcing evidence of how temperature affects anaerobic microbial communities across controlled and real-world conditions. Across both systems, fermentative bacteria consistently constituted the dominant functional guild irrespective of temperature or season, confirming that this group represents the thermally resilient backbone of anaerobic community structure in septic tank environments, consistent with their central role in hydrolysis and acidogenesis.

Despite this shared dominance, the temperature-dependent behaviour of downstream guilds differed between laboratory and field systems in ways that are mechanistically informative. In laboratory-scale reactors, SRB and SAOB increased progressively in relative abundance from 12 °C upward while methanogens remained consistently low. In household septic tanks, SRB dominance was most pronounced in summer, consistent with the laboratory pattern, but considerably more amplified, plausibly reflecting the cumulative effect of sustained seasonal warming across multiple weeks in the field, combined with substrate and operational conditions absent from controlled incubations. The methanogenic response to temperature diverged most substantially between scales. In the laboratory-scale reactors, members of the families Methanobacteriaceae and Methanotrichaceae were consistently dominant across all incubation temperatures. In the field, methanogens showed considerably greater seasonal variation in both relative abundance and family-level composition, with a clear succession from Methanotrichaceae dominance in winter and spring to Methanocorpusculaceae dominance in summer and autumn. This divergence is consistent with the longer HRTs and more stable physical conditions characteristic of field tanks, which are more conductive to the establishment of syntrophic partnerships and methanogenic niches than short-term batch reactor incubations. The presence of Methanotrichaceae across all field seasons extends previous literature syntheses reporting no evidence of acetoclastic methanogens in septic tanks (Shaw and Dorea, 2021), demonstrating that this obligate acetoclastic family persists

under cold-climate *in situ* conditions even where it does not dominate methanogenic flux.

The results indicate that low temperatures do not simply slow anaerobic processes, but instead alter the balance between fermentative, syntrophic, methanogenic, and sulphate-reducing pathways. In this context, more stable digestion refers not to increased peak treatment efficiency, but to reduced seasonal oscillation in microbial guild composition and functional gene representation. Maintaining wastewater temperatures within a narrower range throughout the year would therefore be expected to limit the extent of seasonal microbial reassembly, preserve syntrophic partnerships, and reduce winter accumulation of partially degraded solids. While measures such as insulation, deeper tank placement, or passive thermal buffering are often recommended as general engineering practice in cold regions, this study provides microbial and functional justification for these interventions by demonstrating the sensitivity of anaerobic pathway organisation to small temperature differences.

The convergent findings from both systems indicate that the relationship between microbial community composition and system performance is strongly temperature dependent. Under psychrophilic conditions, microbial community structure is effectively decoupled from measurable performance variation, consistent with kinetically constrained dynamics. Under mesophilic conditions, the dominant physicochemical correlate shifts from solids-related inhibition at 24 °C to substrate-driven dynamics at 36 °C, indicating that, as temperatures rise from spring to summer, the primary constraint on microbial community organisation transitions from physical barriers to substrate availability. Taken together, the laboratory and field correlation analyses demonstrate that microbial guild abundance is systematically linked to treatment state in a temperature- and season-dependent manner: downstream processing guilds are suppressed under conditions of poor performance at intermediate temperatures and high solids loads, substrate-tracking dynamics characterise functionally integrated communities at upper mesophilic temperatures, and spring represents the season of greatest ecological responsiveness to physicochemical gradients – a convergent mechanistic pattern that provides direct microbiological explanation for the seasonal performance variation documented throughout this study.

In household septic tanks, the detection of significant physicochemical correlations

in winter demonstrates that between-tank design and operational differences influence community structure year-round, while the concentration of the broadest and most statistically robust correlations in spring confirms this transitional period as the phase of greatest ecological responsiveness to environmental conditions. These findings extend existing literature on temperature-dependent anaerobic community dynamics (Carballa et al., 2011; Guo et al., 2014) by providing a seasonally resolved, multi-system field dataset linking specific guild families to specific physicochemical drivers across a continuous annual temperature cycle. Comparable studies from wastewater treatment systems in Kazakhstan and Central Asia are absent from the published literature, underscoring the originality of the present dataset and the need for expanded monitoring capacity across a region where cold-climate sanitation challenges are widely shared but poorly documented at the microbiological level.

Microbial and viral communities exhibited coordinated seasonal restructuring, with spring emerging as the critical phase of ecological reorganisation during which multiple anaerobic functional pathways were reactivated concurrently and statistically robust correlations between physicochemical parameters, microbial guilds, and viral families were most pronounced. This study identifies spring as a biologically fragile recovery phase in septic tanks rather than a period of stable or optimal performance, with direct operational relevance: disruption during this phase, including maintenance interventions such as desludging, may delay biological recovery following winter constraint.

Shotgun metagenomic sequencing inherently captures DNA from both viable and non-viable cells and cannot, on its own, distinguish between active and inactive community members. However, viral replication necessarily requires metabolically active hosts (Mara and Horan, 2008; Shan et al., 2014), and the detection of structured phage–host associations inferred using iPHoP, therefore, provides indirect but robust evidence that predicted host taxa represent active components of the community rather than relic DNA. This distinction is particularly important in low-temperature environments, where biomass turnover is slow and DNA persistence may otherwise obscure biological interpretation. Complementary RNA-based approaches could further refine this interpretation by resolving transcriptionally active populations, building on the activity signals inferred here from functional genes and phage–host dynamics.

Genome-informed phage–host predictions revealed that bacteriophages were preferentially associated with dominant and persistent microbial guilds, particularly fermentative bacteria, while associations with SRB, SAOB, and methanogenic archaea were more restricted and seasonally intermittent. These patterns are most consistent with persistence-oriented, lysogeny-biased Piggyback-the-Winner dynamics, rather than classical Kill-the-Winner models characterised by rapid host turnover. This represents the first application of phage–host ecological frameworks to operational household septic tanks, extending viral ecology concepts into a decentralised sanitation context that has previously been largely unexplored. The absence of widespread host collapse, despite seasonal restructuring, indicates that viral communities are more closely linked to microbial stability than to strong top-down regulation under cold-climate, low-energy conditions.

The guild-specific and seasonally constrained pattern of predicted phage infections documented in this study has direct utility for septic tank performance assessment. The restriction of methanogen-associated phage predictions to winter and autumn, and SRB-associated predictions to spring and summer, provides a seasonally resolved map of which functional guilds are most likely to be under active viral influence at each point in the annual cycle. The host-tracking nature of phage associations – where dominant phage families shifted in concert with seasonally dominant host taxa – further indicates that the viral community responded to the same environmental drivers governing microbial guild succession, suggesting that engineering interventions designed to stabilise guild composition would correspondingly stabilise phage–host interaction networks and reduce the ecological disruption associated with seasonal community transitions.

The predominance of lysogeny-consistent phage–host dynamics has three direct implications for process stability. First, the dominant anaerobic functional guilds are unlikely to experience viral-mediated population crashes that could disrupt hydrolysis or downstream processing, contributing to the functional stability evident across all seasons despite strong community restructuring. Second, viral dynamics appear most ecologically active during community reorganisation phases, providing a viral ecological explanation for the elevated community turnover observed during spring and an additional mechanistic basis for avoiding maintenance interventions during this period. Third, the absence of predicted phage infections for Desulfobulbaceae – the dominant SRB family in

winter and autumn – despite its high relative abundance raises the possibility that this family experiences reduced viral predation pressure during the coldest periods, potentially contributing to its competitive dominance under psychrophilic conditions.

The question of whether AMGs carried by phages in these systems actively support host metabolism during the winter-to-spring recovery phase remains unresolved. The dominant viral lineages detected across both laboratory-scale reactors and household septic tanks belonged to the class Caudoviricetes, members of which have been reported to carry AMGs (Zhu et al., 2024; Yao et al., 2025). If expressed, such genes could enhance host metabolic capacity during cold-season constraint, with implications for the rate and completeness of functional recovery during spring reactivation (Fan et al., 2023; Zhu et al., 2024; Yao et al., 2025). Transcriptomic and viromic approaches targeting active phage gene expression would be required to resolve this question.

Winter enrichment of acetate activation and syntrophic pathways, combined with constrained methanogenic activity, indicates that solids accumulation during cold periods is primarily biologically driven rather than solely attributable to hydraulic loading or user behaviour. This provides new mechanistic insight into why winter solids accumulation persists even in structurally intact systems, directly explaining stakeholder reports of increased desludging frequency, and indicates that septic tanks in cold climates require sufficient sludge storage capacity to accommodate seasonal accumulation without compromising hydraulic performance. By explicitly linking microbial function to downstream engineering risk, this study extends conventional hydraulic-based design logic to incorporate biological resilience.

The results of this study highlight that septic system performance in cold climates has important implications not only for organic matter removal but also for broader environmental and public health risks. The observed temperature-dependent reduction in microbial activity and treatment efficiency, coupled with the persistence of pathogen-associated families under low-temperature conditions, indicates that septic tanks may provide limited attenuation of harmful constituents during colder periods, confirmed by the statistically significant effect of temperature on total pathogen-associated abundance in the laboratory-scale reactors (Kruskal–Wallis, $p = 0.036$). Pathogen-associated community dynamics operated through two distinct seasonal

risk pathways. During winter, psychrotolerant families persist in effluent discharged to infiltration systems while cold soils simultaneously exhibit reduced biological activity and limited capacity for pathogen attenuation during subsurface transport. During spring and summer, the risk shifts from persistence-dominated to proliferation-dominated: warming temperatures increasingly favour mesophilic pathogen-associated taxa, while spring mobilisation of accumulated winter solids creates episodic pulses of pathogen-enriched effluent coinciding with peak groundwater recharge from snowmelt.

At the level of individual taxa, families such as Clostridiaceae, Peptostreptococcaceae, and Bacteroidaceae increased markedly at higher temperatures, consistent with their anaerobic, gut-associated nature and preference for warm, nutrient-rich conditions. Enterobacteriaceae and Streptococcaceae exhibited more variable responses, suggesting broader environmental tolerance or regulation by competition, substrate availability, or redox conditions. Moraxellaceae and Pseudomonadaceae showed highly skewed distributions driven by isolated high-abundance observation, indicating episodic or stochastic proliferation rather than consistent temperature-driven behaviour.

The pathogen-associated families identified across both experimental systems have potential utility as biological indicators for routine monitoring in resource-limited settings where direct pathogen testing is not feasible. Arcobacteraceae and Aeromonadaceae, which increased consistently with warming in both systems, have been proposed as indicators of faecal contamination in groundwater monitoring programmes (Aoki et al., 2023), while Pseudomonadaceae, persistent across all temperatures and seasons, represents a thermally resilient year-round candidate. Temperature accounted for 22% of variation in overall microbial community composition, though this fell short of statistical significance (PERMANOVA, $p = 0.11$, $R^2 = 0.22$), indicating that temperature selectively promotes the proliferation of specific pathogen-associated taxa without fundamentally reorganising the broader microbial assemblage, and that taxon-specific monitoring approaches are more appropriate than community-level metrics alone for assessing pathogen-related environmental risk. In the rural Kazakhstani context, where shallow hand-dug wells frequently serve as the primary domestic water source in proximity to household sanitation systems (UNECE, 2017), this seasonal contamination pathway warrants particular attention. Establishing routine seasonal monitoring of pathogen-associated family composition in septic tank effluent and proximal groundwater

would provide the evidence base needed to evaluate whether current designs and management practices adequately protect groundwater resources in this setting.

The ecological patterns identified in this study raise the practical question of whether deliberate inoculation with cold-adapted microbial consortia could enhance functional stability during winter. The laboratory data demonstrated that anaerobic processes continue at 4 °C, with fermentative activity sustained and low-level methanogenic activity detectable, and the functional gene data reinforced this: *ackA*, *pta*, and *mcrA* genes were enriched in winter relative to summer, while WL pathway genes K00194 and K01491 showed cold-season enrichment consistent with sustained syntrophic acetate-centred metabolism. Families with documented psychrophilic members detected in this study – including Psychromonadaceae and Flavobacteriaceae (Jin et al., 2022) – represent candidate targets for inoculation, with the rationale of supplementing the resident community with taxa capable of sustaining fermentative hydrolysis below 10 °C. However, cold-resistance is a property classified at the species rather than family level, and the family-level resolution of the present analysis prevents definitive identification of suitable strains. Furthermore, introduced taxa would face susceptibility to resident phages with which they share no co-evolutionary history – a risk directly evidenced by the close coupling between resident viral communities and established dominant hosts documented in this study (Jin et al., 2022). Controlled mesocosm experiments incorporating realistic temperature cycling, representative influent composition, and phage–host tracking would therefore be required before inoculation could be recommended as a practical management intervention (Jdeed et al., 2025; Yu et al., 2026).

In addition to organic matter removal, nitrogen transformation represents an under-characterised but important dimension of septic tank performance in cold climates, with direct implications for groundwater quality. Septic tank effluent is typically rich in ammonium, generated by ammonification of urea and organic nitrogen under anaerobic conditions (Metcalf and Eddy, 2013), and its attenuation depends on nitrification and denitrification in the surrounding infiltration zone soil (Gray, 2004; Metcalf and Eddy, 2013). Both processes are strongly temperature-dependent, with autotrophic nitrification identified as the primary driver of poor nitrogen removal in cold-climate wastewater systems due to reduced abundance of ammonia- and nitrite-oxidising bacteria at low temperatures (Wild et al., 1971; Fang et al., 2018). At the winter temperatures

documented in this study, the majority of ammonium discharged from the septic tank is therefore expected to pass through the infiltration zone without conversion, entering shallow groundwater as a mobile nitrogen load. This suppression is not absolute: under prolonged cold stress, microbial communities can adapt through enrichment of heterotrophic nitrifying-aerobic denitrifying bacteria, alongside cold-adapted nitrifiers, that sustain ammonia removal efficiencies of up to 99.8% and total nitrogen removal of up to 69.9% at 3 °C to 8 °C through upregulation of extracellular electron transfer mechanisms (Tang et al., 2024). Whether such cold-adapted populations develop in the infiltration zones of the systems studied here cannot be determined from the present dataset. A further risk directly relevant to the spring thaw period – identified in this study as the phase of peak pathogen mobilisation and biological instability – is the production of nitrous oxide (N_2O) from incomplete denitrification during freeze-thaw cycling, driven by increased *nirK* and *norB* gene abundance alongside suppressed *nosZ* expression, with N_2O emissions elevated by over 20% under freeze-thaw conditions and nitrogen conversion substantially more active during thawing than freezing (Su et al., 2025). In the rural Kazakhstan context, where shallow hand-dug wells intercept groundwater in proximity to on-site sanitation systems, elevated winter ammonium loading and spring N_2O production coincide with the same seasonal window of peak pathogen risk identified in this study, constituting a compounded microbiological and chemical contamination risk not captured by conventional carbon-based performance assessments. Future work should incorporate seasonal ammonium, nitrate, and N_2O measurements alongside functional gene profiling of nitrification and denitrification pathways in infiltration zone solids, as a direct analogue to the carbon-cycling gene analyses conducted within the tanks.

7.5 Engineering Interpretation and applicability

The pronounced biological reorganisation observed during spring has direct implications for operation and maintenance practices. Desludging during early or mid-spring would coincide with peak microbial recovery and functional pathway reassembly, increasing the likelihood of removing active biomass at a critical stage and potentially prolonging post-winter instability. In contrast, late summer or early autumn corresponds to a period of comparatively stable microbial composition and functional potential. This study therefore provides the first biologically informed rationale for seasonally optimised

desludging schedules, rather than reliance on fixed annual intervals or sludge depth criteria alone. However, translating this recommendation into practice requires engagement with the specific regulatory, economic, and behavioural barriers that characterise decentralised sanitation contexts, and in Kazakhstan in particular.

At the regulatory level, the current absence of mandatory desludging intervals means that timing decisions are left entirely to household discretion, removing any enforceable mechanism for aligning maintenance behaviour with technical optimal schedules. The introduction of seasonally defined mandatory desludging windows – specifying that desludging must occur between July and October where operationally feasible – would address this gap and create a compliance standard against which municipal or regional authorities could monitor adherence. Nordic countries have demonstrated that mandatory inspection and maintenance regimes for on-site sanitation systems improve compliance rates and system performance (Laukka et al., 2022; Kinnunen et al., 2023), providing a directly applicable governance model for adaptation to the Kazakhstani context. At the economic level, reactive winter desludging – driven by system failure – currently represents the dominant service model, creating limited commercial incentive for operators to promote preventive summer desludging when winter demand is reliably high. Subsidised desludging vouchers redeemable only during the summer-autumn window, combined with unsubsidised pricing for emergency winter callouts, would create a price signal aligned with the technically optimal schedule for households, while municipal contracts guaranteeing minimum summer desludging volumes would provide operators with the commercial predictability needed to invest in planned service delivery capacity. At the behavioural level, the stakeholder survey indicates that households primarily associate system failure with visible operational problems – odour, backflow, and flooding – rather than with pathogen contamination risks. This perception gap suggests that awareness campaigns framing preventive summer desludging as failure avoidance, rather than contamination risk reduction, are more likely to motivate uptake among households with limited prior experience of system failure attributable to poor desludging timing. Engagement through the NGO networks identified in the stakeholder analysis, which already serve as trusted intermediaries between households and technical sanitation guidance, represents a realistic delivery mechanism for such campaigns within the existing institutional landscape.

It is important to note, however, that the feasibility analysis above presupposes

systems that meet minimum design standards. Where fundamental design deficiencies are present, maintenance optimisation alone is insufficient to achieve sustained treatment performance. Common deficiencies in this study region include inadequate tank sizing, insufficient burial depth relative to frost penetration, lack of thermal insulation, and poorly designed effluent discharge systems, all of which directly constrain HRT, thermal stability, and solids separation – the prerequisites for sustained anaerobic treatment. Insufficient retention time and poor solids separation limit the extent of hydrolysis and downstream microbial conversion, while inadequate insulation and shallow installation depths increase susceptibility to the thermal fluctuations that disrupt anaerobic guild balance, as documented in this study. Improving system reliability in cold climates therefore requires regulatory enforcement of construction standards as a baseline condition, alongside the behavioural and operational interventions described above – a point reinforced by the stakeholder survey findings that identified widespread reliance on self-constructed systems built without technical guidance as a primary driver of operational failure in the study region. A concise set of technical recommendations is provided in Appendix F, where key design and operational considerations are summarised for practitioners, builders, and regulators.

7.6 Interpretation scope and scientific context

To evaluate whether the focal tank selected for iPHoP analysis was representative of the broader system ensemble, the microbial community profiles of the remaining nine tanks were assessed for evidence of outlier behaviour using Aitchison distance-based beta-diversity metrics, within-tank and between-tank pairwise distance distributions, and PCoA ordination of CLR-transformed relative abundance data. No tank exhibited persistent outlier status across multiple seasons in terms of guild-level composition or diversity metrics, and inter-tank variation remained within the range expected for operational anaerobic systems subject to comparable environmental drivers. Homogeneity of multivariate dispersion assessed using *betadisper* function in R confirmed significant differences in dispersion among seasons for prokaryotic communities ($p = 0.009$) but only marginal differences for viral communities ($p = 0.056$), indicating that seasonal rather than tank-specific effects dominated compositional structuring across the dataset. PERMANOVA confirmed that season explained 21% of prokaryotic ($R^2 = 0.21$, $p =$

0.001) and 34% of viral (R^2 = 0.34, p = 0.001) community variation, while tank identity explained a smaller and secondary proportion of variance, supporting the interpretation that the ten tanks functioned as comparable seasonal replicates rather than fundamentally distinct biological systems. During winter, one tank showed a transiently elevated relative abundance of SRB families, interpreted as stochastic rather than indicative of stable ecological divergence. During spring thaw, two tanks exhibited delayed reactivation of methanogenic archaea, consistent with differences in thermal buffering capacity linked to installation depth. Neither deviation produced a systematic outlier profile that would fundamentally alter the Piggyback-the-Winner interpretation, which rested on features – absence of host collapse, persistence of dominant guilds, and host-tracking phage associations – that were consistent across the full dataset rather than specific to tank 2. Tanks 6 and 9, which showed the lowest within-tank temporal dispersion, would be expected to exhibit even more stable lysogenic dynamics than those inferred for the focal tank, reinforcing rather than challenging the overall interpretation. Some degree of inter-tank variability was nevertheless evident, particularly during transitional periods such as spring thaw, characterised by increased heterogeneity in microbial composition, functional gene profiles, and substrate availability, likely driven by differences in desludging frequency, hydraulic loading, and household usage patterns. Such variability may influence the strength and detectability of phage-host interactions across systems but remained within the expected ecological range for decentralised anaerobic treatment systems, and no persistent outliers were identified that would challenge the representativeness of the focal case study.

The results presented in this study are based on a single annual monitoring cycle and therefore capture seasonal variability but cannot resolve interannual climatic fluctuations. In cold-climate regions, year-to-year differences in winter severity, duration of freezing conditions, and the timing and intensity of spring thaw can substantially influence both microbial community dynamics and functional pathway recovery. Prolonged winters – characterised by sustained temperatures at or below 4 °C for longer than observed in this study – would be expected to delay the reactivation of syntrophic partnerships and methanogenic pathways beyond the spring recovery window documented here, potentially extending solids accumulation and incomplete organic matter mineralisation into early summer. The functional gene data from this study indicate that *mcrA* and WL pathway

genes recover from winter depletion during spring, but the rate and completeness of this recovery likely depend on the severity and duration of the preceding cold period in ways a single annual cycle cannot characterise. Rapid thaw events represent a distinct additional risk: sudden transition from frozen to above-zero conditions can mobilise accumulated winter solids rapidly, creating hydraulic and organic loading pulses that coincide with the spring community recovery phase discussed above, potentially causing transient deterioration in treatment performance and elevated pathogen-associated family abundance in effluent. Anomalously mild winters may conversely compress the period of psychrophilic community dominance, altering SRB-methanogen competitive dynamics in ways that diverge from the patterns documented here. These dynamics are particularly relevant given projected increases in inter-seasonal temperature variability for Kazakhstan (World Bank, 2026). This study establishes a baseline seasonal framework against which interannual variability can be evaluated, but multi-year monitoring incorporating years of contrasting winter severity is needed to assess whether the spring reactivation patterns documented here are consistent, delayed, or amplified under different climatic conditions, and to provide the empirical basis for climate-adaptive design standards that account for interannual performance variability.

The mechanistic relationships between temperature, microbial dynamics, and functional pathway organisation identified in this study reflect fundamental principles of anaerobic microbial ecology that are not specific to the Kazakhstan context and are expected to apply wherever comparable seasonal thermal constraints are imposed on decentralised anaerobic sanitation systems. The findings are therefore broadly relevant to other cold-climate regions characterised by strong seasonal temperature gradients, including subarctic Scandinavia, northern Canada, Russia, Mongolia, and Central Asia. In subarctic regions, operational septic systems experience analogous winter temperature suppression and spring reactivation dynamics, albeit under more regulated governance environments (Jensen et al., 2025). In permafrost-affected regions such as northern Canada and Siberia, the physical constraints are more severe – ground freezing prevents effluent infiltration and accelerates structural failure of buried components – potentially amplifying the biological effects documented here, particularly during spring thaw when hydraulic loading pulses coincide with peak ecological instability (Jensen et al., 2025). In Mongolia, where freezing temperatures similarly constrain biological treatment

and infiltration (Leblanc et al., 2019), the seasonal pathogen persistence dynamics documented in this study are directly applicable given comparable hydrogeological and socio-economic conditions. The degree to which these patterns manifest quantitatively in other settings will depend on local wastewater characteristics, organic loading rates, system design, and insulation quality, all of which modify the internal thermal environment and substrate conditions experienced by microbial communities. The core finding – that temperature governs seasonal restructuring of anaerobic communities and that spring represents a critical recovery phase requiring operational protection – nevertheless provides a transferable framework for informing cold-climate sanitation design, maintenance scheduling, and environmental risk assessment across a broad range of geographic and institutional contexts.

Taken together, these findings demonstrate that septic tanks operating in cold climates function as seasonally dynamic anaerobic systems, in which physicochemical conditions, microbial and viral community structure, functional potential, and phage–host interactions are tightly governed by temperature. Under winter-constrained conditions, reduced microbial activity and host availability favour lysogeny-dominated phage-host dynamics and suppressed syntrophic coupling; summer-stable conditions promote stronger phage-host associations alongside enhanced guild integration; and transitional periods, particularly spring thaw, are characterised by concurrent reactivation of multiple functional pathways, peak ecological responsiveness, and elevated community heterogeneity – the convergent seasonal pattern documented across both laboratory-scale reactors and household septic tanks in this study. By integrating stakeholder perceptions with field-based microbial, viral, and functional evidence, and by leveraging phage–host dynamics to infer active populations despite the limitations of shotgun metagenomics, this study advances understanding beyond simplified steady-state models and provides a robust scientific basis for climate-adaptive design, construction, and maintenance of on-site sanitation systems in extreme cold climates.

8 Conclusions

This research examined how extreme seasonal temperature variation influences microbial, viral, and functional dynamics in household septic tanks operating in cold-climate conditions. By integrating stakeholder surveys, year-long physicochemical monitoring, laboratory-scale experiments, and shotgun metagenomics, this study addressed the three research questions outlined in Chapter 1 and generated several original contributions to knowledge relevant to on-site sanitation systems in cold and highly seasonal environments.

1. *What are the stakeholder perceptions of sanitation challenges and the feasible solutions for on-site sanitation systems in the coldest cities of Kazakhstan?*

This study is the first systematic investigation to document stakeholder perceptions of on-site sanitation performance in the coldest cities of Kazakhstan and to explicitly link these perceptions to year-round field-based physicochemical and microbiological evidence. Stakeholders consistently identified extreme winter temperatures, restricted access for desludging, variable construction quality, and limited user awareness as the principal barriers to reliable on-site sanitation. Reported winter-related problems – odour emissions, wastewater backflow, and increased solids accumulation – were corroborated by field measurements showing seasonal solids accumulation and reduced treatment efficiency.

A key contribution of this work is demonstrating that stakeholder-reported failures align closely with measured seasonal system behaviour, rather than being isolated operational or behavioural issues. Feasible solutions proposed by users and service providers, including pre-winter desludging, improved insulation of tanks and pipework, deeper installation below frost penetration depth, and stricter construction and installation standards, directly address the dominant temperature-driven constraints identified through monitoring. This integrated socio-technical evidence base provides a stronger foundation for climate-adaptive sanitation policy and practice than previous studies that have considered technical or social dimensions in isolation.

The findings are broadly transferable to other cold-climate regions, including northern high-latitude and subarctic settlements, where strong seasonal temperature gradients

govern microbial processes. However, their expression will depend on local wastewater characteristics, system design, and thermal buffering capacity. In particular, regions with lower water use or permafrost constraints may experience more pronounced functional limitations. These results therefore provide a generalisable, but context-dependent, framework for understanding microbial dynamics in cold-climate sanitation systems.

2. *What methanogens and anaerobic bacteria are responsible for the biological treatment processes in septic tanks at different temperature extremes?*

This study provides the first year-round, field-based characterisation of anaerobic microbial guilds supporting biological treatment in operational household septic tanks under cold-climate conditions. Across temperature extremes, anaerobic treatment performance was supported by fermentative bacteria, including Bacteroidaceae and Carnobacteriaceae, together with SAOB such as Clostridiaceae and Sporomusaceae, which consistently dominated community composition.

Methanogenic archaea, including Methanocorpusculaceae and Methanotrichaceae, together with SRB such as Desulfobulbaceae and Desulfomicrobiaceae, were persistently present throughout the year but exhibited pronounced seasonal restructuring in both taxonomic composition and functional gene representation. Under low-temperature and transitional conditions, functional gene profiles indicated a central role for acetate activation and SAO pathways, coupled primarily to hydrogen-dependent methanogenesis, while direct acetoclastic methanogenesis was functionally constrained despite the presence of acetoclastic taxa.

As temperature increased, methanogenic activity became less constrained, with a greater contribution from direct acetoclastic methane production alongside sustained hydrogenotrophic and methylotrophic pathways. SRB remained active across all seasons, reflecting their integration within shared anaerobic niches and competition for hydrogen and organic intermediates.

Overall, these findings demonstrate that biological treatment in septic tanks at low temperature extremes is maintained by fermentative and syntrophic bacterial processes, with methanogenesis persisting via alternative pathways, while warmer conditions allow a partial shift toward direct acetoclastic methane production.

3. *What are the dynamics and interrelationships of the identified methanogens, bacteria, and bacteriophages in septic tanks at different temperatures?*

This study provides the first field-based, seasonally resolved characterisation of microbial–viral dynamics in operational household septic tanks. Across all ten monitored septic tanks, microbial and viral communities exhibited coordinated seasonal restructuring, with clear coupling in community turnover and distinct clustering by season. These dynamics demonstrate that microbial–viral interactions in septic tanks are strongly temperature dependent, rather than temporally uniform.

Statistically robust associations between physicochemical parameters, microbial guilds, and viral families were detected primarily during spring, identifying transitional temperature periods as critical windows of ecological reorganisation. During these transitions, both microbial and viral communities underwent pronounced restructuring, indicating that rising temperatures act as a key driver of community reassembly and interaction strength. In contrast, winter and summer represented more stable phases, during which microbial–viral coupling was weaker or less detectable.

Genome-informed phage–host prediction using iPHoP in a representative system (septic tank 2) further revealed that phage–host associations are guild-specific, seasonally constrained, and biased toward dominant and persistent microbial populations. Fermentative bacteria exhibited the most consistent predicted phage associations across seasons, reflecting their central role and sustained dominance within anaerobic treatment processes. In contrast, SRB, SAOB, and methanogens showed more restricted and intermittent predicted infections, indicating that viral interactions with these guilds are episodic rather than continuous.

Collectively, these findings indicate that the interrelationships between anaerobic guilds and phages in household septic tanks are best described by persistence-oriented, lysogeny-biased Piggyback-the-Winner interaction regimes, rather than classical Kill-the-Winner dynamics derived from laboratory or high-rate engineered anaerobic systems. Under cold-climate, low-energy, decentralised conditions, viruses appear to be more closely associated with microbial stability and seasonal persistence than with strong top-down regulation. This represents a substantive advance in understanding how microbial and viral communities interact under real-world operating conditions in on-site

sanitation systems.

Overall conclusions and contributions

Overall, this research demonstrates that septic tanks operating in cold climates function predominantly through bacterial fermentation and syntrophic metabolic pathways, with methanogenesis occurring under constrained and seasonally variable conditions rather than as a consistently dominant process. Seasonal temperature fluctuations drive substantial reorganisation of physicochemical conditions, microbial and viral community structure, functional potential, and phage–host associations, with spring emerging as a critical period of ecological restructuring that influences system performance and stability.

For the first time, this study links stakeholder perceptions, field-scale physicochemical monitoring, anaerobic functional gene profiles, and viral ecology within a single, seasonally resolved framework for on-site sanitation systems in cold climates. By demonstrating how microbial and viral processes respond to natural temperature forcing under real operating conditions, this work provides a robust evidence base for climate-adaptive design, improved construction quality, and seasonally informed operation and maintenance strategies. These insights are directly relevant not only to Kazakhstan, but also to other cold and highly seasonal regions to improve the sustainability and reliability of decentralised sanitation systems.

Bibliography

2GIS Mapping System. Digital navigation system, 2023. URL https://2gis.kz/.

Charisma Shonté Acey. Silence and Voice in Nigeria's Hybrid Urban Water Markets: Implications for Local Governance of Public Goods. *International Journal of Urban and Regional Research*, 43(2):313–336, March 2019. ISSN 14682427. doi: 10.1111/1468-2427.12715. Publisher: Blackwell Publishing Ltd.

A. Adegoke and T. Stenstrom. Septic systems. In J. B. Rose and B. Jiménez-Cisneros, editors, *Global Water Pathogen Project*. UNESCO, East Lansing, MI, 2019. doi: 10.14321/waterpathogens.59. URL https://www.waterpathogens.org/book/septic-systems.

Sushant Agarwal and Rakesh K. Gupta. Plastics in Buildings and Construction. In *Applied Plastics Engineering Handbook: Processing, Materials, and Applications: Second Edition*, pages 635–649. Elsevier Inc., 2017. ISBN 978-0-323-39040-8. doi: 10.1016/B978-0-323-39040-8.00030-4.

B. K. Ahring, M. Sandberg, and I. Angelidaki. Volatile fatty acids as indicators of process imbalance in anaerobic digestors. *Appl Microbiol Biotechnol*, 43:559–565, 1995.

Joseph C. Akunna, Juliette M. O'Keeffe, and Richard Allan. Reviewing factors affecting the effectiveness of decentralised domestic wastewater treatment systems for phosphorus and pathogen removal. *Desalination and Water Treatment*, 91:40–47, October 2017. ISSN 19443986. doi: 10.5004/dwt.2017.20750. URL https://linkinghub.elsevier.com/retrieve/pii/S1944398624119054.

Hussein Abed Obaid Alisawi. Performance of wastewater treatment during variable temperature. *Appl Water Sci*, 10(4):89, April 2020. ISSN 2190-5487, 2190-5495. doi: 10.1007/s13201-020-1171-x. URL http://link.springer.com/10.1007/s13201-020-1171-x.

Xavier Fonoll Almansa, Renata Starostka, Lutgarde Raskin, Grietje Zeeman, Francis De Los Reyes, Julie Waechter, Daniel Yeh, and Tanja Radu. Anaerobic Digestion

as a Core Technology in Addressing the Global Sanitation Crisis: Challenges and Opportunities. *Environmental Science and Technology*, 57(48):19078–19087, December 2023. ISSN 15205851. doi: 10.1021/acs.est.3c05291. Publisher: American Chemical Society.

American Public Health Association. *Standard Methods for the Examination of Water and Wastewater*. American Public Health Association, 23 edition, 2017.

I. Angelidaki and B. K. Ahring. Applied Microbiology Biotechnology Thermophilic anaerobic digestion of livestock waste: the effect of ammonia. *Appl Microbiol Biotechnol*, 38:560–564, 1993. doi: 10.1007/BF00242955.

Masataka Aoki, Yasuyuki Takemura, Shuji Kawakami, Wilasinee Yoochatchaval, Thao Tran P., Noriko Tomioka, Yoshitaka Ebie, and Kazuaki Syutsubo. Quantitative detection and reduction of potentially pathogenic bacterial groups of Aeromonas, Arcobacter, Klebsiella pneumoniae species complex, and Mycobacterium in wastewater treatment facilities. *PLoS ONE*, 18(9):e0291742, September 2023. ISSN 1932-6203. doi: 10.1371/journal.pone.0291742. URL `https://dx.plos.org/10.1371/journal.pone.0291742`.

Lise Appels, Jan Baeyens, Jan Degrève, and Raf Dewil. Principles and potential of the anaerobic digestion of waste-activated sludge, December 2008. ISSN: 03601285 Issue: 6 Pages: 755-781 Publication Title: Progress in Energy and Combustion Science Volume: 34.

S. S. Baisholanov, A. D. Kleshenko, G. B. Mussatayeva, A. R. Zhakiyeva, M. S. Gabbassova, E. N. Mukanov, K. A. Akshalov, and D. A. Chernov. Agroclimatic Resources of the Akmola Region. Technical report, Institute of Geography, 2017. ISBN: 9786017150839.

Elisenda Balleste, Anicet R. Blanch, Maite Muniesa, Cristina García-Aljaro, Lorena Rodríguez-Rubio, Julia Martín-Díaz, Miriam Pascual-Benito, and J. Jofre. Bacteriophages in sewage: Abundance, roles, and applications, 2022. ISSN: 26336685 Publication Title: FEMS Microbes Volume: 3.

M. S. Bedinger, A. I. Johnson, and J. S. Fleming. *Site characterization and design of on-site septic systems*. 1997.

L. N. Beisenova, S. S. Shamshedenova, and R. M. Kurbanaliyev. Comprehensive assessment of groundwater quality in the winter season in rural areas of the Karaganda region in the vicinity of the Nura River. *Eurasian Journal of Ecology*, 68(3):10–16, 2021. ISSN 2617-7358. URL https://doi.org/10.26577/EJE.2021.v68.i3.02.

Jinghao Bian, Xiaocheng Wei, Peizhen Chen, Zhiping Huang, Chunxue Zhang, Bo Yang, and Qiang Wang. Effectiveness and Risk Assessment of Septic Tank Treatment Systems in Rural Cold Regions of China. *Processes*, 12(11), November 2024. ISSN 22279717. doi: 10.3390/pr12112385. Publisher: Multidisciplinary Digital Publishing Institute (MDPI).

Gabriel Bitton. *Wastewater Microbiology*. Wiley-Liss, John Wiley & Sons, Hoboken, NJ, 3 edition, 2005.

Emma J. Bowen, Jan Dolfing, Russell J. Davenport, Fiona L. Read, and Thomas P. Curtis. Low-temperature limitation of bioreactor sludge in anaerobic treatment of domestic wastewater. *Water Science and Technology*, 69(5):1004–1013, 2014. ISSN 02731223. doi: 10.2166/wst.2013.821. Publisher: IWA Publishing.

Britannica. Atbasar, 2009. URL https://www.britannica.com/place/Atbasar. Publication Title: Encyclopedia Britannica.

Bureau of National Statistics. Population of the Akmola region, 2020. URL https://stat.gov.kz/region/akmola/population/.

Bureau of National Statistics. Population of the Republic of Kazakhstan, 2023. URL https://stat.gov.kz.

Bureau of National Statistics. Statistics of standard of living, 2025. URL https://stat.gov.kz/en/industries/labor-and-income/stat-life/.

M. Buttle and M. Smith. *Out in the Cold: Emergency water supply and sanitation for cold regions*. WEDC, Loughborough University, UK, 3rd edition, 2004.

João P. S. Cabral. Water microbiology: Bacterial pathogens and water. *International Journal of Environmental Research and Public Health*, 7(10):3657–3703, 2010. doi: 10.3390/ijerph7103657.

Marta Carballa, Marianne Smits, Claudia Etchebehere, Nico Boon, and Willy Verstraete. Correlations between molecular and operational parameters in continuous lab-scale anaerobic reactors. *Applied Microbiology and Biotechnology*, 89(2):303–314, January 2011. ISSN 01757598. doi: 10.1007/s00253-010-2858-y.

Meaghan Castledine and Angus Buckling. Critically evaluating the relative importance of phage in shaping microbial community composition, October 2024. ISSN: 18784380 Issue: 10 Pages: 957-969 Publication Title: Trends in Microbiology Volume: 32.

G. H. Chen, Mark C. M. Loosdrecht, G. A. Ekama, and Damir Brdjanovic. *Biological Wastewater Treatment*. IWA Publishing, 2 edition, 2020. ISBN 978-1-78906-036-2.

Xiaotian Chen, Wei Liu, Yehua Zhao, Huiban He, Jitao Ma, Zongjun Cui, and Xufeng Yuan. Optimization of semi-continuous dry anaerobic digestion process and biogas yield of dry yellow corn straw: Based on gradient anaerobic digestion reactor. *Bioresource Technology*, 389, December 2023. ISSN 18732976. doi: 10.1016/j.biortech.2023.129804. Publisher: Elsevier Ltd.

C. Joon Chuah and Alan D. Ziegler. Temporal Variability of Faecal Contamination from On-Site Sanitation Systems in the Groundwater of Northern Thailand. *Environmental Management*, 61(6):939–953, June 2018. ISSN 0364-152X, 1432-1009. doi: 10.1007/s00267-018-1016-7. URL `http://link.springer.com/10.1007/s00267-018-1016-7`.

Ralf Conrad. Complexity of temperature dependence in methanogenic microbial environments. *Front. Microbiol.*, 14:1232946, July 2023. ISSN 1664-302X. doi: 10.3389/fmicb.2023.1232946. URL `https://www.frontiersin.org/articles/10.3389/fmicb.2023.1232946/full`.

Allie Davis, Amy Javernick-Will, and Sherri M. Cook. A comparison of interviews, focus groups, and photovoice to identify sanitation priorities and increase success of community-based sanitation systems. *Environmental Science: Water Research and Technology*, 4(10):1451–1463, October 2018. ISSN 20531419. doi: 10.1039/c8ew00391b. Publisher: Royal Society of Chemistry.

Francesco Durazzi, Claudia Sala, Gastone Castellani, Gerardo Manfreda, Daniel Remondini, and Alessandra De Cesare. Comparison between 16S rRNA and shotgun

sequencing data for the taxonomic characterization of the gut microbiota. *Scientific Reports*, 11(1), December 2021. ISSN 20452322. doi: 10.1038/s41598-021-82726-y. Publisher: Nature Research.

Stefan Dyksma, Lukas Jansen, and Claudia Gallert. Syntrophic acetate oxidation replaces acetoclastic methanogenesis during thermophilic digestion of biowaste. *Microbiome*, 8 (1), July 2020. ISSN 20492618. doi: 10.1186/s40168-020-00862-5. Publisher: BioMed Central.

Lu Fan, Wei Peng, Haowen Duan, Fan Lü, Hua Zhang, and Pinjing He. Presence and role of viruses in anaerobic digestion of food waste under environmental variability. *Microbiome*, 11(1), December 2023. ISSN 20492618. doi: 10.1186/s40168-023-01585-z. Publisher: BioMed Central Ltd.

Dexin Fang, Gen Zhao, Xiaoyi Xu, Qian Zhang, Qiushi Shen, Zhuoyao Fang, Liping Huang, and Fangying Ji. Microbial community structures and functions of wastewater treatment systems in plateau and cold regions. *Bioresource Technology*, 249:684–693, February 2018. ISSN 09608524. doi: 10.1016/j.biortech.2017.10.063. URL `https://linkinghub.elsevier.com/retrieve/pii/S0960852417318928`.

C.S.B. Fitzpatrick, E. Fradin, and J. Gregory. Temperature effects on flocculation, using different coagulants. *Water Science and Technology*, 50(12):171–175, December 2004. ISSN 0273-1223. doi: 10.2166/wst.2004.0710. URL `https://doi.org/10.2166/wst.2004.0710`.

Ingrid H. Franke-Whittle, Andreas Walter, Christian Ebner, and Heribert Insam. Investigation into the effect of high concentrations of volatile fatty acids in anaerobic digestion on methanogenic communities. *Waste Management*, 34(11):2080–2089, November 2014. ISSN 18792456. doi: 10.1016/j.wasman.2014.07.020. Publisher: Elsevier Ltd.

Mengjiao Gao, Bing Guo, Lei Zhang, Yingdi Zhang, and Yang Liu. Microbial community dynamics in anaerobic digesters treating conventional and vacuum toilet flushed blackwater. *Water Research*, 160:249–258, September 2019. ISSN 18792448. doi: 10.1016/j.watres.2019.05.077. Publisher: Elsevier Ltd.

I. Martin Garcia, M. Mokosch, A. Soares, M. Pidou, and B. Jefferson. Impact on reactor configuration on the performance of anaerobic MBRs: Treatment of settled sewage in temperate climates. *Water Research*, 47(14):4853–4860, September 2013. ISSN 18792448. doi: 10.1016/j.watres.2013.05.008. Publisher: Elsevier Ltd.

Michael H. Gerardi. *The Microbiology of Anaerobic Digesters*. 2003. ISBN 3-17-572399-3. URL `https://onlinelibrary.wiley.com/doi/`.

Makarand M. Ghangrekar. *Wastewater to Water: Principles, Technologies and Engineering Design*. Springer Nature Singapore Pte Ltd., 2022. doi: https://doi.org/10.1007/978-981-19-4048-4_19.

H. M. Giaever. Treatment performance of a multistage constructed wetland at 69° N, Norway. *Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering*, 35(8):1377–1388, 2000. ISSN 10934529. doi: 10.1080/10934520009377041. Publisher: Marcel Dekker Inc.

Sebastián Gradilla-Hernández, Enrique J. Herrera-López, Anne Gschaedler, Marisela González-Avila, Rita Fuentes-Aguilar, and Alejandro Garcia-Gonzalez. Differential neural network identifier for parameter determination of a mixed microbial culture model. volume 51, pages 479–484. Elsevier B.V., January 2018. doi: 10.1016/j.ifacol.2018.07.323. ISSN: 24058963 Issue: 13.

N. F. Gray. *Biology of Wastewater Treatment*. Imperial College Press, 2nd edition, 2004.

Andrej M. Grjibovski, Nassikhat Nurgaliyeva, Aliya Kosbayeva, and Bettina Menne. No association between temperature and deaths from cardiovascular and cerebrovascular diseases during the cold season in Astana, Kazakhstan - The second coldest capital in the world. *International Journal of Circumpolar Health*, 71(1):1–3, 2012. ISSN 12399736. doi: 10.3402/ijch.v71i0.19769.

Xiaohui Guo, Cheng Wang, Faqian Sun, Weijing Zhu, and Weixiang Wu. A comparison of microbial characteristics between the thermophilic and mesophilic anaerobic digesters exposed to elevated food waste loadings. *Bioresource Technology*, 152:420–428, 2014. ISSN 18732976. doi: 10.1016/j.biortech.2013.11.012. Publisher: Elsevier Ltd.

HaploX. Shotgun metagenomic sequencing. Technical report, 2024.

Sharareh Harirchi, Steven Wainaina, Taner Sar, Seyed Ali Nojoumi, Milad Parchami, Mohsen Parchami, Sunita Varjani, Samir Kumar Khanal, Jonathan Wong, Mukesh Kumar Awasthi, and Mohammad J. Taherzadeh. Microbiological insights into anaerobic digestion for biogas, hydrogen or volatile fatty acids (VFAs): a review, 2022. ISSN: 21655987 Issue: 3 Pages: 6521-6557 Publication Title: Bioengineered Volume: 13.

S. Herbst, S. Benedikter, U. Koester, N. Phan, C. Berger, A. Rechenburg, and T. Kistemann. Perceptions of water, sanitation and health: A case study from the Mekong Delta, Vietnam. *Water Science and Technology*, 60(3):699–707, 2009. ISSN 02731223. doi: 10.2166/wst.2009.442.

Santiago Hernández and Martha J. Vives. Phages in anaerobic systems, October 2020. ISSN: 19994915 Issue: 10 Publication Title: Viruses Volume: 12.

R. Heyer, K. Schallert, C. Siewert, F. Kohrs, J. Greve, I. Maus, J. Klang, M. Klocke, M. Heiermann, M. Hoffmann, S. Püttker, M. Calusinska, R. Zoun, G. Saake, D. Benndorf, and U. Reichl. Metaproteome analysis reveals that syntrophy, competition, and phage-host interaction shape microbial communities in biogas plants. *Microbiome*, 7(1), April 2019. ISSN 20492618. doi: 10.1186/s40168-019-0673-y. Publisher: BioMed Central Ltd.

Tomoyuki Hori, Shin Haruta, Yoshiyuki Ueno, Masaharu Ishii, and Yasuo Igarashi. Dynamic transition of a methanogenic population in response to the concentration of volatile fatty acids in a thermophilic anaerobic digester. *Applied and Environmental Microbiology*, 72(2):1623–1630, February 2006. ISSN 00992240. doi: 10.1128/AEM.72.2.1623-1630.2006.

Sorin Istrail, Pavel Pevzner, Michael Waterman, Søren Brunak, Mikhail S. Gelfand, Thomas Lengauer, Eugene Myers, David Sankoff, Ron Shamir, Terry Speed Walter, Eliza Hall, Martin Vingron, and W. Eric Wong. Research in Computational Molecular Biology. Springer, 2016. URL `http://www.springer.com/series/5381`.

Ghadeer Jdeed, Bogdana Kravchuk, and Nina V. Tikunova. Factors Affecting Phage–Bacteria Coevolution Dynamics. *Viruses*, 17(2):235, February 2025. ISSN

1999-4915. doi: 10.3390/v17020235. URL https://www.mdpi.com/1999-4915/17/2/235.

Pernille Erland Jensen, Débora Boratto, Pekka M. Rossi, Maria Velmitskaya, Ida Beathe Øverjordet, Hrund Ólöf Andradóttir, Lisbeth Truelstrup Hansen, Inga Herrmann, Rakul Mortensen, Katrin Hoydal, Aaron Dotson, Hanne Kvitsand, Elisangela Heiderscheidt, Sarah Gewurtz, Ken Johnson, Petter D. Jenssen, Anatoly Sinitsyn, Bing Chen, and Rob Jamieson. The status of domestic wastewater treatment in the Arctic. *Environ. Sci.: Adv.*, 4(9):1373–1402, 2025. ISSN 2754-7000. doi: 10.1039/D5VA00082C. URL https://xlink.rsc.org/?DOI=D5VA00082C.

Petter D. Jenssen. Design and performance of ecological sanitation systems in Norway. In *Proceedings of the First International Conference on Ecological Sanitation*, Nanning, China, 2002. EcoSanRes.

Minxi Jiang, Audrey L. W. Wang, Nicholas A. Be, Nisha Mulakken, Kara L. Nelson, and Rose S. Kantor. Evaluation of the Impact of Concentration and Extraction Methods on the Targeted Sequencing of Human Viruses from Wastewater. *Environmental Science and Technology*, 58(19):8239–8250, May 2024. ISSN 15205851. doi: 10.1021/acs.est.4c00580. Publisher: American Chemical Society.

Shanshan Jin, Yizhe Wang, and Xihong Zhao. Cold-adaptive mechanism of psychrophilic bacteria in food and its application. *Microbial Pathogenesis*, 169:105652, August 2022. ISSN 0882-4010. doi: 10.1016/j.micpath.2022.105652. URL https://www.sciencedirect.com/science/article/pii/S0882401022002650.

Supriya Karekar, Renan Stefanini, and Birgitte Ahring. Homo-Acetogens: Their Metabolism and Competitive Relationship with Hydrogenotrophic Methanogens. *Microorganisms*, 10(2):397, February 2022. ISSN 2076-2607. doi: 10.3390/microorganisms10020397. URL https://www.mdpi.com/2076-2607/10/2/397.

Kazhydromet. Climate of Kazakshtan, 2023a. URL https://www.kazhydromet.kz/en/klimat/klimat-kazahstana/.

Kazhydromet. Climate of the Akmola region. 2023b. URL https://kazhydromet.kz/ru/klimat/klimat-kazahstana-po-oblastyam.

Juho Kinnunen, Pekka M. Rossi, Inga Herrmann, Anna Kaisa Ronkanen, and Elisangela Heiderscheidt. Factors affecting effluent quality in on-site wastewater treatment systems in the cold climates of Finland and Sweden. *Journal of Cleaner Production*, 404, June 2023. ISSN 09596526. doi: 10.1016/j.jclepro.2023.136756. Publisher: Elsevier Ltd.

C. Kinsley, K. Kennedy, and A. Crolla. A combined reed bed/freezing bed technology for septage treatment and reuse in cold climate regions. *Water Science and Technology*, 76 (2):286–293, July 2017. ISSN 02731223. doi: 10.2166/wst.2017.189. Publisher: IWA Publishing.

T. Koottatep, S. Phuphisith, T. Pussayanavin, A. Panuvatvanich, and C. Polprasert. Modeling of pathogen inactivation in thermal septic tanks. *Journal of Water Sanitation and Hygiene for Development*, 4(1):81–88, 2014a. ISSN 20439083. doi: 10.2166/wash dev.2013.190. Publisher: IWA Publishing.

T. Koottatep, P. Prapasriket, T. Pussayanavin, and C. Polprasert. Performance of up-flow thermophilic septic tank treating blackwater. *Environment, Development and Sustainability*, 22(4):3691–3700, April 2020a. ISSN 15732975. doi: 10.1007/s10668-019 -00362-z. Publisher: Springer.

Thammarat Koottatep, Rawintra Eamrat, Tatchai Pussayanavin, and Chongrak Polprasert. Hydraulic evaluation and performance of on-site sanitation systems in central Thailand. *Environmental Engineering Research*, 19(3):269–274, 2014b. ISSN 2005968X. doi: 10.4491/eer.2014.S1.004. Publisher: Korean Society of Environmental Engineers.

Thammarat Koottatep, Stephanie Connelly, Tatchai Pussayanavin, Sopida Khamyai, Wattanapong Sangchun, William Sloan, and Chongrak Polprasert. "Solar septic tank": Evaluation of innovative decentralized treatment of blackwater in developing countries. *Journal of Water Sanitation and Hygiene for Development*, 10(4):828–840, December 2020b. ISSN 24089362. doi: 10.2166/washdev.2020.168. Publisher: IWA Publishing.

Thammarat Koottatep, Tatchai Pussayanavin, and Chongrak Polprasert. Nouveau design solar septic tank: Reinvented toilet technology for sanitation 4.0. *Environmental Technology and Innovation*, 19, August 2020c. ISSN 23521864. doi: 10.1016/j.eti. 2020.100933. Publisher: Elsevier B.V.

Vuokko Laukka, Johanna Kallio, Inga Herrmann, Riikka Malila, Ritva Nilivaara, and Elisangela Heiderscheidt. Governance of on-site sanitation in Finland, Sweden and Norway. Technical report, 2022.

Maryanne Leblanc, Robert A. Reed, Martin P. Gambrill, and Diego Juan Rodriguez. Challenges and Opportunities for Improving Household Sanitation in the Ger Areas of Ulaanbaatar. Technical report, 2017. URL `www.worldbank.org/water`.

Maryanne Leblanc, Robert A. Reed, Martin Gambrill, and Diego Juan Rodriguez. Improving Sanitation in Cold Regions Catalog of Technical Options for Household-Level Sanitation. Technical report, 2019. URL `www.worldbank.org`.

Changsoo Lee, Jaai Kim, Kwanghyun Hwang, Vincent O'Flaherty, and Seokhwan Hwang. Quantitative analysis of methanogenic community dynamics in three anaerobic batch digesters treating different wastewaters. *Water Research*, 43(1):157–165, 2009. ISSN 00431354. doi: 10.1016/j.watres.2008.09.032. Publisher: Elsevier Ltd.

Carlos M. Lopez-Vazquez, Bipin Dangol, Christine M. Hooijmans, and Damir Brdjanovic. Co-treatment of Faecal Sludge in Municipal Wastewater Treatment Plants. In *Faecal Sludge ManagementSystems Approach Implementation and Operation*, pages 177–198. IWA Publishing, 2014.

Sari Luostarinen. *Anaerobic On-Site Wastewater Treatment at Low Temperatures*. PhD Thesis, University of Jyväskylä, Finland, 2005.

Sari Luostarinen, Wendy Sanders, Katarzyna Kujawa-Roeleveld, and Grietje Zeeman. Effect of temperature on anaerobic treatment of black water in UASB-septic tank systems. *Bioresource Technology*, 98(5):980–986, March 2007. ISSN 09608524. doi: 10.1016/j.biortech.2006.04.018.

D. Duncan Mara and N. J. Horan. *Handbook of water and wastewater microbiology*. Academic Press, 2008. ISBN 0-12-470100-0.

Luke J. McKay, Mensur Dlakic, Matthew W. Fields, Tom O. Delmont, A. Murat Eren, Zackary J. Jay, Korinne B. Klingelsmith, Douglas B. Rusch, and William P. Inskeep. Co-occurring genomic capacity for anaerobic methane and dissimilatory sulfur metabolisms discovered in the Korarchaeota. *Nature Microbiology*, 4(4):614–622, April

2019. ISSN 20585276. doi: 10.1038/s41564-019-0362-4. Publisher: Nature Publishing Group.

James D. McLeod, Maazuza Z. Othman, David J. Beale, and Deepak Joshi. The use of laboratory scale reactors to predict sensitivity to changes in operating conditions for full-scale anaerobic digestion treating municipal sewage sludge. *Bioresource Technology*, 189:384–390, August 2015. ISSN 18732976. doi: 10.1016/j.biortech.2015.04.049. Publisher: Elsevier Ltd.

Ran Mei and Wen-Tso Liu. Meta-Omics-Supervised Characterization of Respiration Activities Associated with Microbial Immigrants in Anaerobic Sludge Digesters. *Environ. Sci. Technol.*, 56(10):6689–6698, May 2022. ISSN 0013-936X, 1520-5851. doi: 10.1021/acs.est.2c01029. URL `https://pubs.acs.org/doi/10.1021/acs.est.2c01029`.

Metcalf and Eddy. *Wastewater Engineering: Treatment and Resource Recovery*. McGraw-Hill Education, 5th edition, 2013.

M. El Nadi, T. Sabry, K. Hassan, O. Hamdy, and M. Helmy. Improving Wastewater Effluent of Septic Tank Using Filtration Membranes. *International Journal of Science and Engineering Investigations*, 6:70, 2017. ISSN 2251-8843. URL `www.IJSEI.com`.

James Naphtali, Alexander W. Y. Chan, Faizan Saleem, Enze Li, Jacob Devries, and Herb E. Schellhorn. Comparative Metagenomics of Anaerobic Digester Communities Reveals Sulfidogenic and Methanogenic Microbial Subgroups in Conventional and Plug Flow Residential Septic Tank Systems. *Processes*, 10(3), March 2022. ISSN 22279717. doi: 10.3390/pr10030436. Publisher: MDPI.

Khanh Q. Nguyen, Célestin Mwiseneza, Khaled Mohamed, Patrice Cousin, Mathieu Robert, and Brahim Benmokrane. Long-term testing methods for HDPE pipe - advantages and disadvantages: A review. *Engineering Fracture Mechanics*, 246, April 2021. ISSN 00137944. doi: 10.1016/j.engfracmech.2021.107629. Publisher: Elsevier Ltd.

Erqi Nie, Pinjing He, Hua Zhang, Liping Hao, Liming Shao, and Fan Lü. How does temperature regulate anaerobic digestion? *Renewable and Sustainable Energy Reviews*,

150, October 2021. ISSN 18790690. doi: 10.1016/j.rser.2021.111453. Publisher: Elsevier Ltd.

S. Nonet. Requirements and test methods for on site domestic wastewater treatment plants: the European standard (prEN 12566-3) compared to other international standards. *Water Science & Technology*, 51(10):147–154, 2005. URL `http://iwaponline.com/wst/article-pdf/51/10/147/433382/147.pdf`.

Daniela Numberger, Lars Ganzert, Luca Zoccarato, Kristin Mühldorfer, Sascha Sauer, Hans Peter Grossart, and Alex D. Greenwood. Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by full-length 16S rRNA sequencing. *Scientific Reports*, 9(1), December 2019. ISSN 20452322. doi: 10.1038/s41598-019-46015-z. Publisher: Nature Publishing Group.

OECD. *Sustainable Business Models for Water Supply and Sanitation in Small Towns and Rural Settlements in Kazakhstan*. OECD, April 2016. ISBN 978-92-64-24939-4. doi: 10.1787/9789264249400-en. URL `https://www.oecd-ilibrary.org/environment/sustainable-business-models-for-water-supply-and-sanitation-in-small-towns-and-rural-settlements-in-kazakhstan_9789264249400-en`.

Derya Ozuolmez, Hyunsoo Na, Mark A. Lever, Kasper U. Kjeldsen, Bo B. Jørgensen, and Caroline M. Plugge. Methanogenic archaea and sulfate reducing bacteria co-cultured on acetate: Teamwork or coexistence? *Frontiers in Microbiology*, 6(MAY), 2015. ISSN 1664302X. doi: 10.3389/fmicb.2015.00492. Publisher: Frontiers Research Foundation.

Anmol Parajuli, Aakash Khadka, Lokesh Sapkota, and Anish Ghimire. Effect of Hydraulic Retention Time and Organic-Loading Rate on Two-Staged, Semi-Continuous Mesophilic Anaerobic Digestion of Food Waste during Start-Up. *Fermentation*, 8(11): 620, November 2022. doi: 10.3390/fermentation8110620. Publisher: MDPI AG.

B. Pathak. The role of ngos in the sanitation sector. In *Proceedings of the 22nd WEDC Conference*, page 12, 1996.

Evangelos Petropoulos, Jan Dolfing, Russell J. Davenport, Emma J. Bowen, and Thomas P. Curtis. Developing cold-adapted biomass for the anaerobic treatment of domestic wastewater at low temperatures (4, 8 and 15 °c) with inocula from cold

environments. *Water Research*, 112:100–109, 2017. ISSN 18792448. doi: 10.1016/j.watres.2016.12.009. Publisher: Elsevier Ltd.

Roya Pishgar, Dean Morin, Shane J. Young, Jon Schwartz, and Angus Chu. Characterization of domestic wastewater released from 'green' households and field study of the performance of onsite septic tanks retrofitted into aerobic bioreactors in cold climate. *Science of the Total Environment*, 755, February 2021. ISSN 18791026. doi: 10.1016/j.scitotenv.2020.142446. Publisher: Elsevier B.V.

Chongrak Polprasert, Thammarat Koottatep, and Tatchai Pussayanavin. Solar septic tanks: A new sanitation paradigm for Thailand 4.0, February 2018. ISSN: 15131874 Pages: 39-43 Publication Title: ScienceAsia Volume: 44.

Tatchai Pussayanavin, Thammarat Koottatep, Rawintra Eamrat, and Chongrak Polprasert. Enhanced sludge reduction in septic tanks by increasing temperature. *Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering*, 50(1):81–89, January 2015. ISSN 15324117. doi: 10.1080/10934529.2015.964633. Publisher: Taylor and Francis Inc.

Qiagen. AllPrep PowerViral DNA/RNA Kit Handbook. Technical report, 2018.

Emmanuelle R. J. Quemin, Petr Chlanda, Martin Sachse, Patrick Forterre, David Prangishvili, and Mart Krupovic. Eukaryotic-like virus budding in archaea. *mBio*, 7(5), September 2016. ISSN 21507511. doi: 10.1128/mBio.01439-16. Publisher: American Society for Microbiology.

Christopher Quince, Alan W. Walker, Jared T. Simpson, Nicholas J. Loman, and Nicola Segata. Shotgun metagenomics, from sampling to analysis, September 2017. ISSN: 15461696 Issue: 9 Pages: 833-844 Publication Title: Nature Biotechnology Volume: 35.

Haripriya Rama, Busiswa Ndaba, Mokhotjwa Simon Dhlamini, Nicolene Cochrane, Malik Maaza, and Ashira Roopnarain. Elucidating Key Microbial Drivers for Methane Production during Cold Adaptation and Psychrophilic Anaerobic Digestion of Cattle Manure and Food Waste. *Fermentation*, 10(7):370, July 2024. ISSN 2311-5637. doi: 10.3390/fermentation10070370.

Leticia Regueiro, Juan M. Lema, and Marta Carballa. Key microbial communities steering the functioning of anaerobic digesters during hydraulic and organic overloading shocks. *Bioresource Technology*, 197:208–216, December 2015. ISSN 18732976. doi: 10.1016/j.biortech.2015.08.076. Publisher: Elsevier Ltd.

Samia Richards, Eric Paterson, Paul J. A. Withers, and Marc Stutter. Septic tank discharges as multi-pollutant hotspots in catchments. *Science of the Total Environment*, 542:854–863, January 2016. ISSN 18791026. doi: 10.1016/j.scitotenv.2015.10.160. Publisher: Elsevier.

Samia Richards, Paul J. A. Withers, Eric Paterson, Colin W. McRoberts, and Marc Stutter. Potential tracers for tracking septic tank effluent discharges in watercourses. *Environmental Pollution*, 228:245–255, 2017. ISSN 18736424. doi: 10.1016/j.envpol.2017.05.044. Publisher: Elsevier Ltd.

Simon Roux. Integrated Phage-Host Prediction tool (iPHoP) v1.0.0, June 2022.

Daniel Gheorghe Rudaru, Irina Eugenia Lucaciu, and Ana Maria Fulgheci. Correlation between BOD5 and COD – biodegradability indicator of wastewater. *Romanian Journal of Ecology & Environmental Chemistry*, 4(2):80–86, December 2022. ISSN 26685418. doi: 10.21698/rjeec.2022.207. URL `http://www.rjeec.ro/Read/183?id=Correlation%20between%20BOD5%20and%20COD%20%E2%80%93%20biodegradability%20indicator%20of%20wastewaters`.

Stanley B. Sam, Savanna K. Smith, Robert Niederdorfer, Andreas Scheidegger, Barbara J. Ward, James Madalitso Tembo, Joel Kabika, David R. Johnson, Helmut Bürgmann, Eberhard Morgenroth, and Linda Strande. Microbial community composition reflects water usage and storage conditions in a city-wide study of non-sewered wastewater (fecal sludge). *PLOS Water*, 4(6 June), June 2025. ISSN 27673219. doi: 10.1371/journal.pwat.0000386. Publisher: Public Library of Science.

Mohammad Taghi Samadi, Ghorban Asgari, Mostafa Leili, and Sonia Chavoshi. Integrated modified septic tank and constructed wetland: an alternative green technology for phytoremediation of highly polluted leachate. *Biomass Conversion and Biorefinery*, 2023. ISSN 21906823. doi: 10.1007/s13399-023-04333-3. Publisher: Springer Science and Business Media Deutschland GmbH.

E. V. Severskiy. Geocryological Hazards of the Plans of Kazakhstan. *Questions on Geography and Geoecology*, (3):35–40, 2011. URL https://www.elibrary.ru/item.asp?id=60034103.

Jinyu Shan, Sunee Korbsrisate, Patoo Withatanung, Natalie Lazar Adler, Martha R. J. Clokie, and Edouard E. Galyov. Temperature dependent bacteriophages of a tropical bacterial pathogen. *Frontiers in Microbiology*, 5(NOV), 2014. ISSN 1664302X. doi: 10.3389/fmicb.2014.00599. Publisher: Frontiers Media S.A.

Kelsey Shaw and Caetano C. Dorea. Biodegradation mechanisms and functional microbiology in conventional septic tanks: A systematic review and meta-analysis. *Environmental Science: Water Research and Technology*, 7(1):144–155, January 2021. ISSN 20531419. doi: 10.1039/d0ew00795a. Publisher: Royal Society of Chemistry.

Kelsey A. Shaw. Revisiting the evidence: An investigation, quantification and characterization of the biodegradation mechanisms and functional microbiology in conventional septic tanks. Master of applied science thesis, University of Victoria, Victoria, Canada, 2021.

Jessica R. Sieber, Michael J. McInerney, and Robert P. Gunsalus. Genomic Insights into Syntrophy: The Paradigm for Anaerobic Metabolic Cooperation. *Annual Review of Microbiology*, 66(1):429–452, October 2012. ISSN 0066-4227. doi: 10.1146/annurev-micro-090110-102844.

Katja Sigel. Environmental sanitation in peri-urban ger areas in the city of darkhan, mongolia: A description of current status, practices, and perceptions. UFZ-Bericht 02/2010, Helmholtz Centre for Environmental Research – UFZ, Leipzig, 2010. URL https://www.econstor.eu/handle/10419/57850.

D. Sinor, D. R. Smith, G. R. G. Hambly, and E. Allworth. Kazakhstan, 2023. URL https://www.britannica.com/place/Kazakhstan. Publication Title: Encyclopedia Britannica.

Marcos von Sperling. *Basic principles of wastewater treatment*. 2007. ISBN 1-84339-162-7. URL https://www.iwapublishing.com/sites/default/files/ebooks/9781780402093.pdf.

Alfons J. M. Stams and Caroline M. Plugge. Electron transfer in syntrophic communities of anaerobic bacteria and archaea, 2009. ISSN: 17401526 Issue: 8 Pages: 568-577 Publication Title: Nature Reviews Microbiology Volume: 7.

Fei Su, Yinghua Li, Chaoqun Zhu, Geshere Abdisa Gurmesa, and Yunting Fang. Impact of freeze-thaw cycle on metagenomics in subsurface wastewater infiltration systems: Ecological implications for greenhouse gas emissions. *Journal of Environmental Management*, 395:127839, December 2025. ISSN 03014797. doi: 10.1016/j.jenvman.2025.127839. URL https://linkinghub.elsevier.com/retrieve/pii/S0301479725038150.

Somnathhdatta Subharuppguhaa. *Frontiers in Probability and the Statistical Sciences Statistical Analysis off Microbiome Data*. 2022. URL http://www.springer.com/series/11957.

K. Syafira, J. S. M. Ahmad, and S. P. Saraswati. Modification of Septic Tank for Sedimentation Optimization in the Public Toilet (Case Study at Wisdom Park UGM). In *IOP Conference Series: Earth and Environmental Science*, volume 930. IOP Publishing Ltd, December 2021. doi: 10.1088/1755-1315/930/1/012006. ISSN: 17551315 Issue: 1.

David Talavera-Cortés, Laureano E. Carpio, Patricia Serrano-Candelas, Carlos Lafita, María José Tàrrega Marti, Ángela Baeza-Serrano, Pau Granell, Rafael Gozalbes, and Eva Serrano-Candelas. Computational Characterisation of Sulphate-Reducing Bacteria Inhibitors to Overcome Methanogenic Competence and Optimise Green Biogas Production. *Applied Microbiology*, 5(4):128, November 2025. ISSN 2673-8007. doi: 10.3390/applmicrobiol5040128. URL https://www.mdpi.com/2673-8007/5/4/128.

Hideyuki Tamaki, Rui Zhang, Florent E. Angly, Shota Nakamura, Pei Ying Hong, Teruo Yasunaga, Yoichi Kamagata, and Wen Tso Liu. Metagenomic analysis of DNA viruses in a wastewater treatment plant in tropical climate. *Environmental Microbiology*, 14 (2):441–452, February 2012. ISSN 14622912. doi: 10.1111/j.1462-2920.2011.02630.x.

Taotao Tang, Zhiyong Zhao, Yin Wang, Xiaolong Zhao, Rongxin Zhu, Jiali Wu, Shuming Yang, and Maolin He. Internal mechanism of microbial nitrogen removal under prolonged low-temperature stress: A focus on microbial communities, metabolic

pathways, and microbial function. *Chemical Engineering Journal*, 500:157522, November 2024. ISSN 13858947. doi: 10.1016/j.cej.2024.157522. URL https://linkinghub.elsevier.com/retrieve/pii/S1385894724090132.

G. Tchobanoglous, F.L Burton, and H.D Stensel. *Wastewater Engineering: Treatment and Reuse*. Inc. Metcalf & Eddy, McGraw-Hill, 4th edition, 2003.

Thermo Scientific. NanoDrop 8000 Spectrophotometer V2.2 User Manual. Technical report, 2008. URL www.nanodrop.com.

Melanie J. Thomson, Tania M. Stevanin, and James W. B. Moir. Measuring nitric oxide metabolism in the pathogen neisseria meningitidis. *Methods in Enzymology*, 437: 539–558, 2008. doi: 10.1016/S0076-6879(07)37027-4.

Y. P. Thye, A. J. Effendi, P. Soewondo, D. Brdjanovic, and T. Setiadi. Understanding how people innovate for emergency sanitation: A case study of a local NGO. *Water Practice and Technology*, 10(4):704–710, December 2015. ISSN 1751231X. doi: 10.2166/wpt.2015.086. Publisher: IWA Publishing.

Gareth Trubl, Jeffrey A. Kimbrel, Jose Liquet-Gonzalez, Erin E. Nuccio, Peter K. Weber, Jennifer Pett-Ridge, Janet K. Jansson, Mark P. Waldrop, and Steven J. Blazewicz. Active virus-host interactions at sub-freezing temperatures in Arctic peat soil. *Microbiome*, 9(1), December 2021. ISSN 20492618. doi: 10.1186/s40168-021-01154-2. Publisher: BioMed Central Ltd.

UN General Assembly. The human rights to safe drinking water and sanitation : resolution, adopted by the General Assembly, 2015. URL https://digitallibrary.un.org/record/821067.

UNECE. Target indicators for the Protocol on Water and Health to the Convention on the Protection and Use of Transboundary Watercourses and International Lakes (Republic of Kazakhstan). Technical report, Astana, 2017. URL https://unece.org/sites/default/files/2025-05/Kazakhstan_analiz_proekti_pokazateley_Protokol_May_2017_RUS.pdf.

United Nations. SDG 6: Ensure availability and sustainable management of water and sanitation for all., 2015. URL https://sdgs.un.org/goals/goal6.

United Nations. Deputy Secretary-General Asks Governments to Quadruple Action on 'Neglected But Essential Issue' of Sanitation, Marking World Toilet Day, 2021.

Adrie Veeken and Bert Hamelers. Effect of temperature on hydrolysis rates of selected biowaste components. *Bioresource Technology*, (69):249–254, 1999.

Konstantina Velkushanova, Linda Strande, Mariska Ronteltap, Thammarat Koottatep, Damir Brdjanovic, and Chris Buckley. *Methods for faecal sludge analysis*. 2021. ISBN 978-1-78040-912-2.

Elena Vialkova, Svetlana Maksimova, Marina Zemlyanova, Lev Maksimov, and Anna Vorotnikova. Integrated Design Approach to Small Sewage Systems in the Arctic Climate. *Environmental Processes*, 7(2):673–690, June 2020. ISSN 21987505. doi: 10.1007/s40710-020-00427-6. Publisher: Springer.

E. Vinokurov, A. Akhunbaev, S. Chuyev, A. Adakhaev, and T. Sarsembekov. Drinking Water Supply and Sanitation in Central Asia. Technical Report 24/5, Eurasian Development Bank, Almaty, 2024.

T. Viraraghavan and K. Dickenson. Low-temperature anaerobic filtration of septic tank effluent. *Cold Regions Science and Technology*, 19(3):245–252, 1991. ISSN 0165-232X. doi: https://doi.org/10.1016/0165-232X(91)90039-J.

Dou Wang, Lei Liu, Xiaoqing Xu, Chunxiao Wang, Yulin Wang, Yu Deng, and Tong Zhang. Distributions, interactions, and dynamics of prokaryotes and phages in a hybrid biological wastewater treatment system. *Microbiome*, 12(1), December 2024. ISSN 20492618. doi: 10.1186/s40168-024-01853-6. Publisher: BioMed Central Ltd.

Long Wang, Yinzhao Wang, Xingyu Huang, Ruijie Ma, Jiangtao Li, Fengping Wang, Nianzhi Jiao, and Rui Zhang. Potential metabolic and genetic interaction among viruses, methanogen and methanotrophic archaea, and their syntrophic partners. *ISME Communications*, 2(1), December 2022. doi: 10.1038/s43705-022-00135-2. Publisher: Oxford University Press (OUP).

Mian Wang, Julia Zhu, and Xinwei Mao. Removal of Pathogens in Onsite Wastewater Treatment Systems: A Review of Design Considerations and Influencing Factors. *Water*,

13(9):1190, April 2021. ISSN 2073-4441. doi: 10.3390/w13091190. URL https://www.mdpi.com/2073-4441/13/9/1190.

Raymond R. Weil and Nyle C. Brady. *Nature and Properties of Soils*. Pearson Education, Limited, 15 edition, 2017. URL http://ebookcentral.proquest.com/lib/imperial/detail.action?docID=5186149.

WHO/UNICEF JMP. Guidance note to facilitate country consultations on JMP estimates - for WASH in households. Technical report, 2022a. URL https://www.washdata.org/.

WHO/UNICEF JMP. Trends in sewer, septic tank and latrine coverage, 2000-2022. Technical report, 2022b. URL https://washdata.org/data/household#!/dashboard/new.

WHO/UNICEF JMP. Trends in sewer, septic tank and latrine coverage, 2000-2024, 2025. URL https://washdata.org/data/household#!/dashboard/new. Publication Title: JMP Database.

H.E. Jr. Wild, C. N. Sawyer, and T. C. McMahon. Factors Affecting Nitrification Kinetics. *Journal (Water Pollution Control Federation)*, 43(9):1845–1854, 1971. URL https://www.jstor.org/stable/25037179.

Eric Williamson. Cold Climate Performance Analysis of On-Site Domestic Wastewater Treatment Systems. *Water Environment Research*, 82(6):512–518, June 2010. ISSN 15547531. doi: 10.2175/106143009x12529484815557. Publisher: Wiley.

Paul JA Withers, Philip Jordan, Linda May, Helen P. Jarvie, and Nancy E. Deal. Do septic tank systems pose a hidden threat to water quality? *Frontiers in Ecology and the Environment*, 12(2):123–130, March 2014. ISSN 1540-9295. doi: 10.1890/130131. URL https://esajournals.onlinelibrary.wiley.com/doi/epdf/10.1890/130131.

World Bank. Population data, 2022. URL https://data.worldbank.org/indicator/SP.POP.TOTL?name_desc=true.

World Bank. Climate Change Knowledge Portal: Historical Data, 2023. URL https://datacatalog.worldbank.org/search/dataset/0040276/Climate-Change-Knowledge-Portal--Historical-Data.

World Bank. Kazakhstan, 2025. URL https://www.worldbank.org/ext/en/country/kazakhstan?utm_source=chatgpt.com#tab-economy.

World Bank. Kazakhstan's Projected Climate, 2026. URL https://climateknowledgeportal.worldbank.org/country/kazakhstan/climate-data-projections.

Linwei Wu, Yunfeng Yang, Si Chen, Mengxin Zhao, Zhenwei Zhu, Sihang Yang, Yuanyuan Qu, Qiao Ma, Zhili He, Jizhong Zhou, and Qiang He. Long-term successional dynamics of microbial association networks in anaerobic digestion processes. *Water Research*, 104: 1–10, November 2016. ISSN 18792448. doi: 10.1016/j.watres.2016.07.072. Publisher: Elsevier Ltd.

Ziqi Yang, Donghai Yang, Yu Hua, Xiang Chen, Xiankai Wang, Hui Gong, Bin Dong, Xiujin Li, and Xiaohu Dai. Dual optimization in anaerobic digestion of rice straw: Effects HRT and OLR coupling on methane production in one-stage and two-stage systems. *Journal of Environmental Management*, 370, November 2024. ISSN 10958630. doi: 10.1016/j.jenvman.2024.123041. Publisher: Academic Press.

Jian Yao, Yan Zeng, Xia Hong, Meng Wang, Quan Zhang, Yating Chen, Min Gou, Zi-Yuan Xia, and Yue-Qin Tang. Phages-bacteria interactions underlying the dynamics of polyhydroxyalkanoate-producing mixed microbial cultures via meta-omics study. *mSystems*, 10(4):e00200–25, April 2025. ISSN 2379-5077. doi: 10.1128/msystems.00200-25. URL https://journals.asm.org/doi/10.1128/msystems.00200-25.

Orhan Yenigün and Burak Demirel. Ammonia inhibition in anaerobic digestion: A review, May 2013. ISSN: 13595113 Issue: 5-6 Pages: 901-911 Publication Title: Process Biochemistry Volume: 48.

Cui Yu, Zhaoxia Dong, Zihong Peng, Zhixian Zhu, Feng Zhang, Yong Li, and Cheng Zhang. Phage therapy under multiple abiotic stresses: Systematic review on mechanisms of plant disease control. *Plant Stress*, 21:101351, May 2026. ISSN 2667064X. doi: 10.1016/j.stress.2026.101351. URL https://linkinghub.elsevier.com/retrieve/pii/S2667064X2600134X.

Hui Yun, Bin Liang, Zhangwei He, Minghan Li, Simin Zong, Zhenfei Wang, Bin Ge, Peng Zhang, Xiangkai Li, and Aijie Wang. Insights into methanogenesis of

mesophilic-psychrophilic varied anaerobic digestion of municipal sludge with antibiotic stress. *Journal of Environmental Management*, 331:117278, April 2023. ISSN 03014797. doi: 10.1016/j.jenvman.2023.117278.

Dongqing Zhang, Richard M. Gersberg, and Tan Soon Keat. Constructed wetlands in China, October 2009. ISSN: 09258574 Issue: 10 Pages: 1367-1378 Publication Title: Ecological Engineering Volume: 35.

Junyu Zhang, Qun Gao, Qiuting Zhang, Tengxu Wang, Haowei Yue, Linwei Wu, Jason Shi, Ziyan Qin, Jizhong Zhou, Jiane Zuo, and Yunfeng Yang. Bacteriophage-prokaryote dynamics and interaction within anaerobic digestion processes across time and space. *Microbiome*, 5(1):57, May 2017. ISSN 20492618. doi: 10.1186/s40168-017-0272-8.

Hao Zheng, Hong Xiang Chai, Liuwei Zhao, Yong Liao, Xuekang Cao, Lihua Feng, and Fangying Ji. Pretreatment of septic tank wastewater by packed anaerobic baffled reactor: Pollutant degradation and microbial community succession in different compartments. *Environmental Research*, 223, April 2023. ISSN 10960953. doi: 10.1016/j.envres.2023.115475. Publisher: Academic Press Inc.

Hexi Zhou, Xin Li, Guoren Xu, and Huarong Yu. Overview of strategies for enhanced treatment of municipal/domestic wastewater at low temperature, December 2018. ISSN: 18791026 Pages: 225-237 Publication Title: Science of the Total Environment Volume: 643.

Dong Zhu, Shu-Yue Liu, Ming-Ming Sun, Xing-Yun Yi, Gui-Lan Duan, Mao Ye, Michael R. Gillings, and Yong-Guan Zhu. Adaptive expression of phage auxiliary metabolic genes in paddy soils and their contribution toward global carbon sequestration. *Proc. Natl. Acad. Sci. U.S.A.*, 121(49):e2419798121, December 2024. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.2419798121. URL `https://pnas.org/doi/10.1073/pnas.2419798121`.

Stephen H. Zinder and Martin Dworkin. Morphological and Physiological Diversity. In Rosenberg E, Delong E, Lory S, Stackebrandt E, and Thompson F, editors, *The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*, volume 4, page 108. Springer Berlin, Heidelberg, 2013. URL `https://link.springer.com/referencework/10.1007/978-3-642-30194-0`.

Appendices

A Ethical approval

Imperial College London

Science Engineering Technology Research Ethics Committee
Imperial College London
Room 221
Medical School BuildingSt Marys CampusLondon
W2 1PG
Tel: +44 (0)207 594 1862

researchethicscommittee@imperial.ac.uk

31 March 2023

Dear Aiman Uteyeva

Study Title: Re-inventing the septic tank

SETREC number: 6417731

The above study was reviewed by the Science Engineering Technology Research Ethics Committee on 21 February 2023.

This study was given a favourable opinion subject to revisions reviewed by the RGIT. Following the review of your amended documents submitted on 20 March 2023 the Research Governance and Integrity Team would like to grant full approval of the study on the basis described in the revised documents:

Documents
The documents reviewed were:

- Application form (v1.0 31/03/23)
- Protocol (v2.0 20/03/23)
- Participant Information Sheet (v1.0 02/02/23)
- Consent Form (v1.0 02/02/23)
- Advert (v1.0 31/03/23)

Membership of the Committee

The members of the Ethics Committee who reviewed the study were:

Chris Hughes	Chris Watkins
Ira Bhattacharya	Marek Sergot
Vivienne Kuh	Stephen Webster
Stefano Sandrone	
Becky Ward	

Yours sincerely,

Chris Hughes,
Chair, Science Engineering Technology Research Ethics Committee

Imperial College of Science, Technology and Medicine

Figure A.1: Original ethical approval letter.

Imperial College London

Imperial College Research Ethics Committee
Imperial College London
Room 221
Medical School BuildingSt Marys CampusLondon
W2 1PG
Tel: +44 (0)207 594 1862

researchethicscommittee@imperial.ac.uk

12/12/2023

Dear Aiman Uteyeva

Study Title: Re-inventing the septic tank

SETREC Reference number: 6417731

The above Notice of Amendment was reviewed by the Research Governance and Integrity Team (RGIT) on 12 December 2023.

The Research Governance and Integrity Team have reviewed the revised documents you submitted and would like to grant RGIT approval to this study on the basis described in the Notice of Amendment.

Documents
The documents reviewed were:

- Notice of Amendment
- Protocol (v3.0, 07/12/2023)
- Participant Information Sheet (v2.0, 07/12/2023)
- Advert (v2.0, 07/12/2023)
- Fieldwork Risk Assessment Form (v Aug 2022, 29/11/2023)

Yours sincerely,

Ruth Nicholson,
Head of Research Governance and Integrity,
Imperial College London

Figure A.2: First approval of the amendment by SETREC.

Imperial College London

Imperial College Research Ethics Committee
Imperial College London
Room 221
Medical School BuildingSt Marys CampusLondon
W2 1PG
Tel: +44 (0)207 594 1862

researchethicscommittee@imperial.ac.uk

27/02/2024

Dear Aiman Uteyeva

Study Title: Re-inventing the septic tank

SETREC Reference number: 6417731

The above Notice of Amendment was reviewed by the Research Governance and Integrity Team (RGIT) on 26 February 2024.

The Research Governance and Integrity Team have reviewed the revised documents you submitted and would like to grant RGIT approval to this study on the basis described in the Notice of Amendment.

Documents
The documents reviewed were:

- Notice of Amendment (v1, 26/02/2024)
- Protocol (v4.0, 14/02/2024)
- Participant Information Sheet (v3.0, 21/02/2024)
- Consent form (v2.0, 21/02/2024)
- Advert (v3.0, 14/02/2024)
- Questionnaire (v1.0, 23/02/2024)

Yours sincerely,

Ruth Nicholson,
Head of Research Governance and Integrity,
Imperial College London

Figure A.3: Second (and final) approval of the amendment by SETREC.

B Survey design

Table B.1: Characteristics of the categories of the main sanitation stakeholders.

Stakeholders	Main responsibilities	Number of respondents (n=30)	
		Astana (n=15)	**Atbasar (n=15)**
Ministry of Water Resources and Irrigation	• Development and enforcement of regulations and standards related to the use and protection of the water fund, water supply, sanitation, and irrigation.	1	-
Ministry of Ecology and Natural Resources (Committee for Environmental Regulation and Control)	• Monitoring compliance with environmental regulations related to water and sanitation.	1	-
Ministry of Industry and Infrastructure Development (Committee for Construction and Housing and Communal Services)	• Oversight of the public utilities • Coordination, monitoring, and analysis of water and sanitation provision.	1	-
Municipal authorities	• Oversight of the publicly owned water utilities.	1	1

Stakeholders	Main responsibilities	Number of respondents (n=30) **Astana (n=15)**	**Atbasar (n=15)**
Water Utilities	• Collection and management of sanitation fees • Water and wastewater management. • Provision of centralised water and sanitation services • Faecal sludge collection and management from on-site sanitation facilities.	1	1
Association of enterprises for water supply and sanitation	• Capacity building in water utilities • Uphold the rights of water utilities.	1	-
NGOs	• Support environment-related initiatives (incl. sanitation and wastewater management) • Raise public awareness regarding environmental issues.	1	2
Political parties	• Raise awareness of the water and sanitation-related problems • Provide recommendations on the government legislature.	2	-
Small-scale faecal sludge management businesses	• Provide water and sanitation-related services (transport, collection, installation of on-site sanitation facilities, etc.).	1	1

Stakeholders	Main responsibilities	Number of respondents (n=30) **Astana (n=15)**	**Atbasar (n=15)**
Households	• Provision of timely payments for water and sanitation services • Manage water resources and sanitation responsibly.	5	10

Dear survey participant,

You are given this document because you have agreed to participate in the *Re-Inventing the Septic Tank* project, read the Participant Information Sheet for Sanitation Stakeholders, received all answers to your questions, and signed the Consent Form for Sanitation Stakeholders. Please inform us how you wish the interview to be conducted – either in-person or indirectly by filling out the survey and sending it back by email (to: a.uteyeva22@imperial.ac.uk)

Please feel free to answer the questions from your professional (if working in the sanitation sector) or personal (if households) perspective. All answers will be kept strictly confidential.

QUESTIONNAIRE

Personal Information:

Gender: ☐ Male ☐ Female	Age:	Residence: ☐ Astana ☐ Atbasar

Sanitation-related questions:

1. Have you heard about the Sustainable Development Goals? If so, which SDG targets have been identified as priorities for your projects?

2. From your perspective, what is the sanitation situation in Kazakhstan? What are the cold weather-related challenges in the sanitation sector?

3. From your perspective, how aware are ministries and the public of the need to address the sanitation-related problems?

4. What existing activities being undertaken by the country/your organisation to implementing the sanitation-related projects and what lessons have been learnt?

5. What plans do you have in relation to improving the sanitation situation in Kazakhstan? What sustainable solutions to improve the sanitation situation would you propose?

6. Do you have limited access to safe sanitation during the cold season? Do you have limited access to safe sanitation during all seasons?

Figure B.1: Questionnaire published as supplementary material in Paper 1 (page 1)(see Section 1.7).

7. What barriers or challenges do you consider the most important in achieving safe sanitation in the cold climate environment?

Barriers or Challenges (check if applicable)	Explanation
_ Political support	
_ Public support	
_ Financial support	
_ Understanding of the importance of sanitation issues and their impact on the drinking water quality and the environment	
_ Coordination within/between relevant stakeholders	
_ Availability of sanitation experts	
_ Scientific evidence base (information/data) to support arguments	
_ Economic evidence base (information/data) to support arguments	
_ Other reasons	

If you are a representative of the household, please additionally answer the following questions:

8. Do you find aging infrastructure as the main problem for your household?

9. Is your sanitation facility connected to centralised wastewater treatment facilities?

10. Does your household have any problems with the operation and maintenance of existing sanitation infrastructure?

Thank you for your participation!

Figure B.2: Questionnaire published as supplementary material in Paper 1 (page 2)(see Section 1.7).

C Septic tank sampling and laboratory analysis

C.1 Definitions of physicochemical parameters measured in the study

- **pH:** Determined by the concentration of hydrogen ions in solution, pH represents a critical control variable affecting both chemical equilibria and biological activity in anaerobic systems (Metcalf and Eddy, 2013). A range of 6.8-7.2 is generally considered optimal for stable reactor performance (Koottatep et al., 2020c,b). Deviations outside this range can suppress microbial activity and reduce process efficiency (Metcalf and Eddy, 2013; Franke-Whittle et al., 2014).

- **TS** – Measured as the residual matter remaining after evaporation and drying at 103 °C to 105 °C, TS provides an overall indicator of wastewater composition and concentration (Metcalf and Eddy, 2013; American Public Health Association, 2017).

- **VS** – Determined as the organic fraction of TS lost upon ignition at 550 °C, VS represents the biodegradable organic content of the sample (Metcalf and Eddy, 2013; American Public Health Association, 2017).

- **VS/TS ratio (organic fraction of solids)** – Calculated as an index of the proportion of biodegradable material within the total solids, this ratio serves as a practical indicator of anaerobic digestion efficiency (Koottatep et al., 2020c).

- BOD_5 - Measured as the oxygen consumed by microbial activity during a five-day incubation period, BOD_5 quantifies the biologically degradable fraction of organic matter in wastewater (American Public Health Association, 2017). While widely used, its accuracy may be reduced in samples containing non-biodegradable or inhibitory compounds (Lopez-Vazquez et al., 2014).

- **COD** – Reflects the total oxygen equivalent required to chemically oxidise both organic and inorganic compounds in the sample, providing a comprehensive measure of total pollutant load (American Public Health Association, 2017).
- **BOD_5/COD ratio** – Used as an indicator of wastewater biodegradability, with higher ratios signifying a greater proportion of readily degradable organic matter (Samadi et al., 2023).
- **VFAs** – Quantified as the sum of short-chain organic acids (mainly acetic, propionic, and butyric acids) formed during the acidogenesis stage of anaerobic digestion (Metcalf and Eddy, 2013). VFA accumulation indicates process imbalance between acidogenic and methanogenic microbial groups (Hori et al., 2006; Metcalf and Eddy, 2013; Franke-Whittle et al., 2014).

Figure C.1: Monthly sampling from one of the septic tanks.

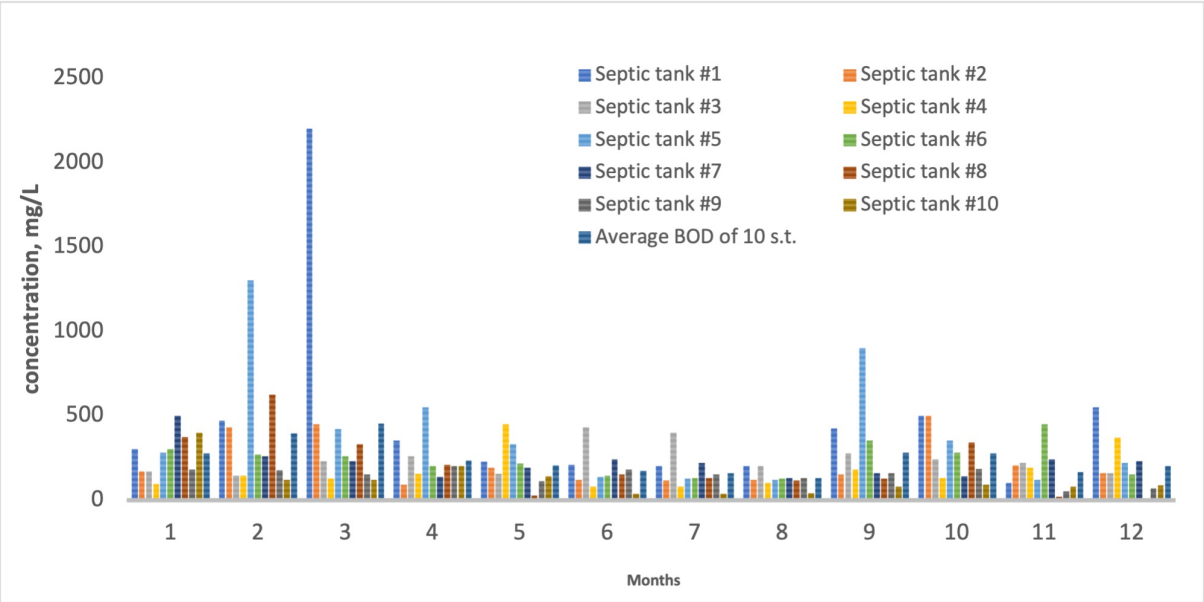

Figure C.2: Annual variation in BOD_5 concentration in wastewater samples collected from the first compartments of ten conventional septic tanks.

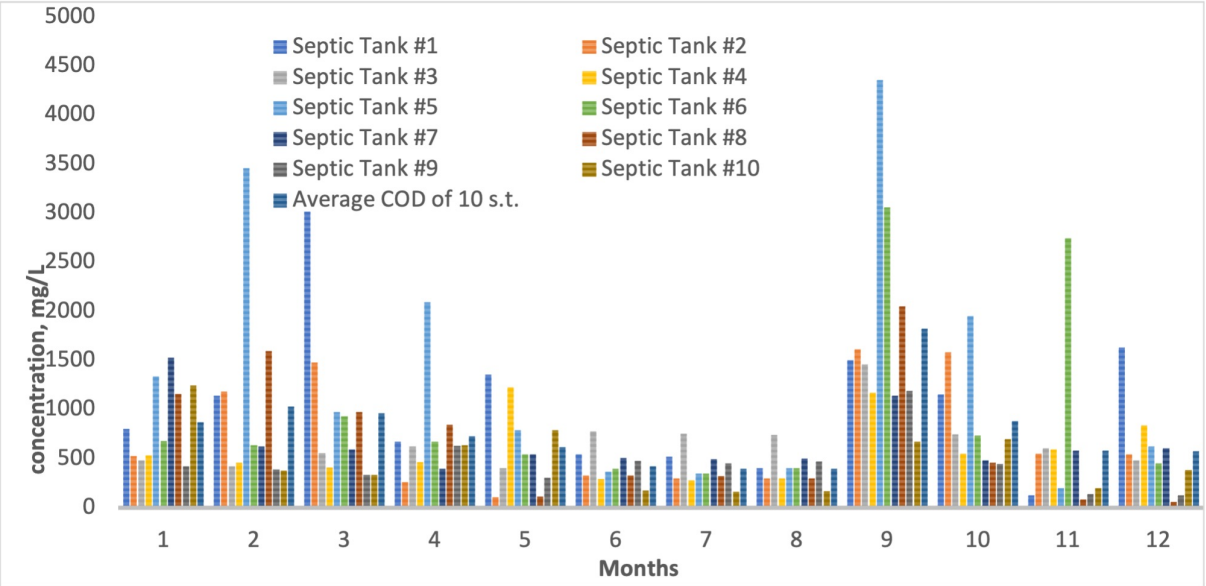

Figure C.3: Annual variation in COD concentration in wastewater samples collected from the first compartments of ten conventional septic tanks.

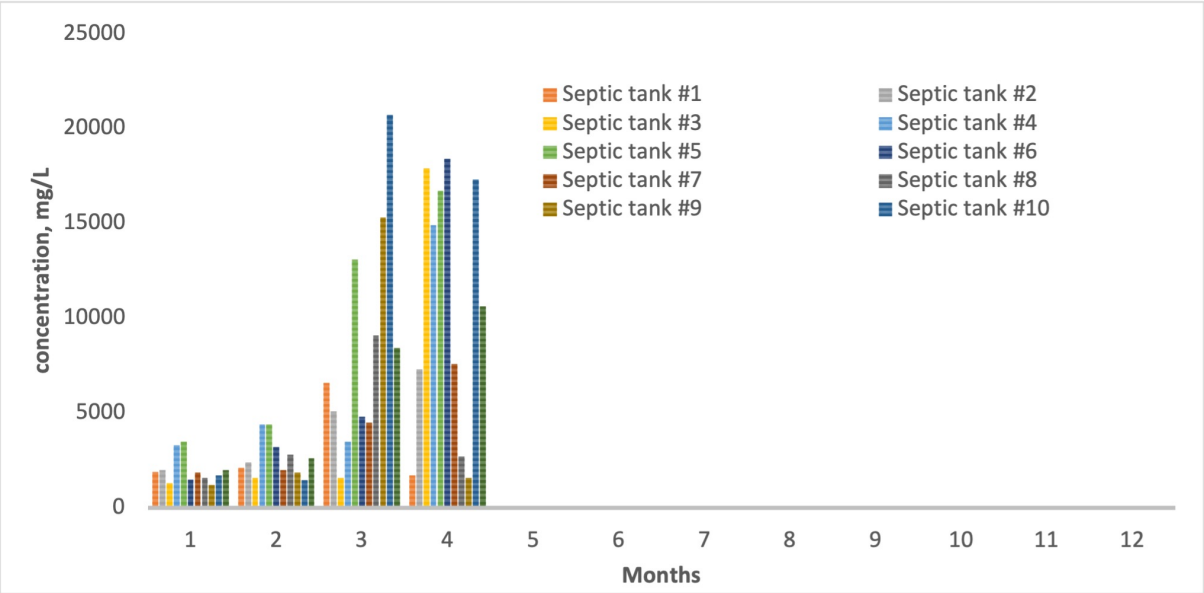

Figure C.4: Annual variation in TS concentration in wastewater samples collected from the first compartments of ten conventional septic tanks. TS concentrations were close to zero in May-December.

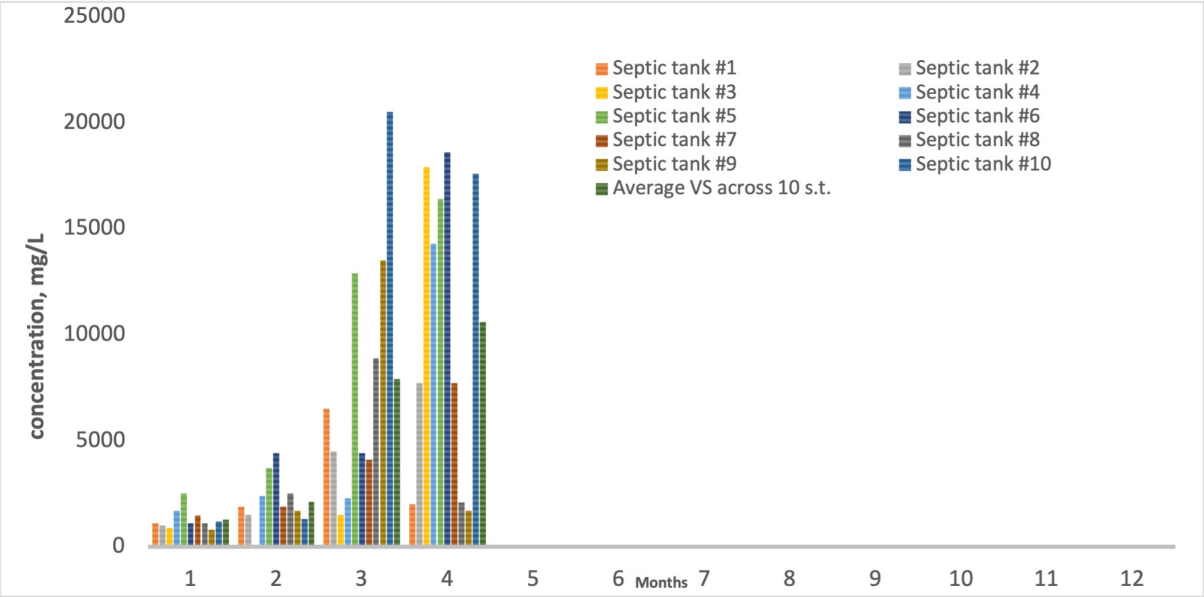

Figure C.5: Annual variation in VS concentration in wastewater samples collected from the first compartments of ten conventional septic tanks. VS concentrations were close to zero in May-December.

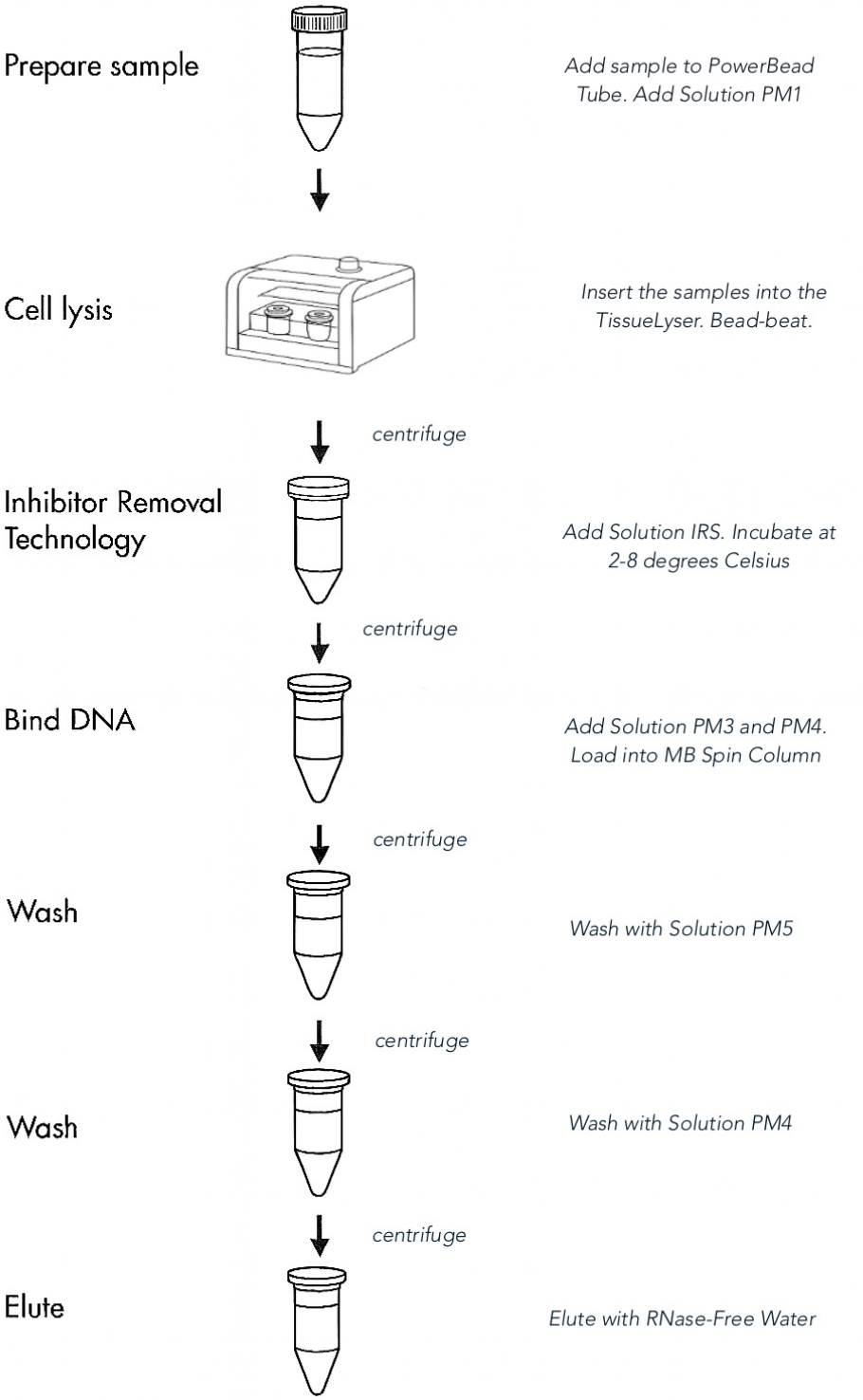

Figure C.6: DNA isolation protocol using Qiagen AllPrep Power Viral DNA/RNA kit, adapted from Qiagen (2018).

Table C.1: Summary of household septic tank samples included in metagenomic analyses (sampling year: 2024).

#	Sample ID	Septic Tank ID	Month	Season
1	A1	Tank2	January	Winter
2	S2A5	Tank5	February	Winter
3	S2A10	Tank10	February	Winter
4	A2	Tank3	April	Spring
5	A3	Tank7	April	Spring
6	S4A8	Tank8	April	Spring
7	A4	Tank8	April	Spring
8	S4A10	Tank10	April	Spring
9	Sample_G55	Tank 1	May	Spring
10	Sample_G1	Tank 1	May	Spring
11	Sample_G2	Tank2	May	Spring
12	Sample_G3	Tank3	May	Spring
13	Sample_G57	Tank4	May	Spring
14	Sample_G5	Tank4	May	Spring
15	Sample_G6	Tank5	May	Spring
16	Sample_G7	Tank5	May	Spring
17	Sample_G8	Tank5	May	Spring
18	Sample_G10	Tank6	May	Spring
19	Sample_G11	Tank7	May	Spring
20	Sample_G14	Tank8	May	Spring
21	Sample_G15	Tank9	May	Spring
22	Sample_G17	Tank10	May	Spring
23	Sample_G25	Tank 1	June	Summer
24	Sample_G26	Tank 1	June	Summer
25	Sample_G28	Tank2	June	Summer
26	Sample_G29	Tank2	June	Summer
27	Sample_G31	Tank3	June	Summer
28	Sample_G34	Tank4	June	Summer

#	Sample ID	Septic Tank ID	Month	Season
29	Sample_G37	Tank5	June	Summer
30	Sample_G40	Tank6	June	Summer
31	Sample_G43	Tank7	June	Summer
32	Sample_G44	Tank7	June	Summer
33	Sample_G46	Tank8	June	Summer
34	Sample_G49	Tank9	June	Summer
35	Sample_G52	Tank10	June	Summer
36	Sample_G53	Tank10	June	Summer
37	Sample_G22	Tank 1	September	Autumn
38	A7	Tank 1	September	Autumn
39	S9A2	Tank2	September	Autumn
40	A8	Tank3	September	Autumn
41	A9	Tank5	September	Autumn
42	S9A51	Tank5	September	Autumn
43	A10	Tank2	November	Autumn
44	A11	Tank4	November	Autumn
45	A12	Tank5	November	Autumn
46	A13	Tank7	November	Autumn
47	Sample_G20	Tank7	November	Autumn
48	A14	Tank 1	December	Winter
49	Sample_G19	Tank 1	December	Winter
50	A5	Tank 5	December	Winter

D Statistical analysis

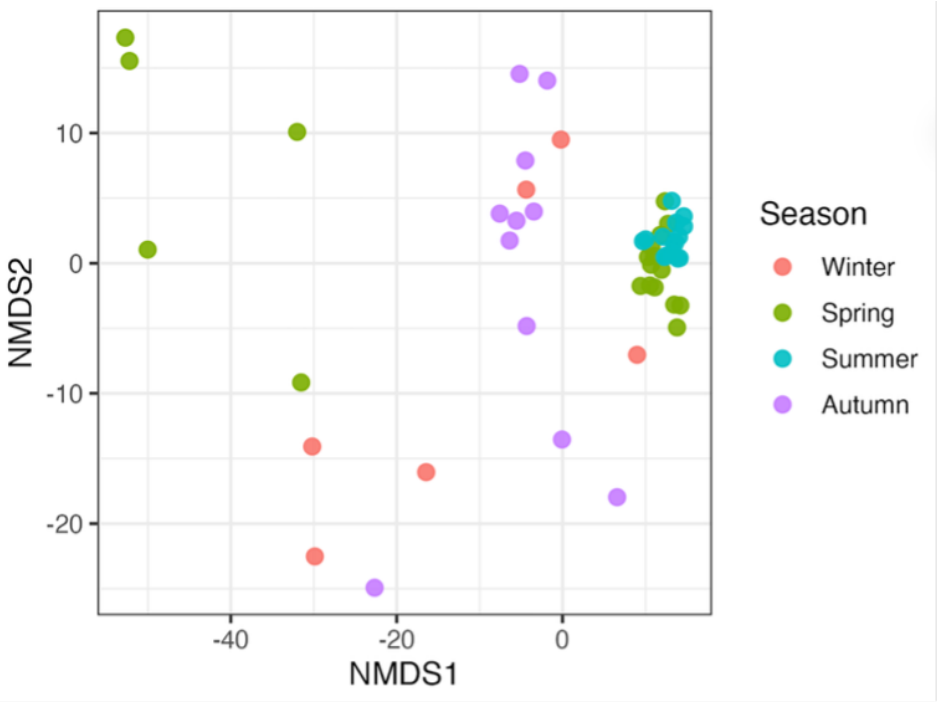

Figure D.1: NMDS ordination of prokaryotic communities based on Aitchison distances (stress = 0.092) across all ten septic tanks. Points represent samples coloured by season.

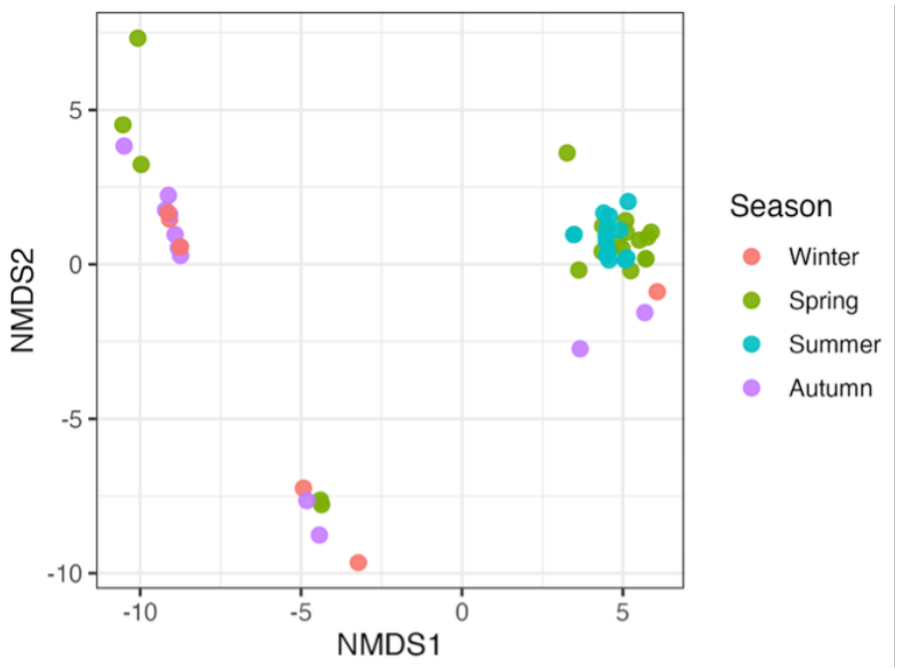

Figure D.2: NMDS ordination of viral communities based on Aitchison distances (stress = 0.035) across all ten septic tanks. Points represent samples coloured by season.

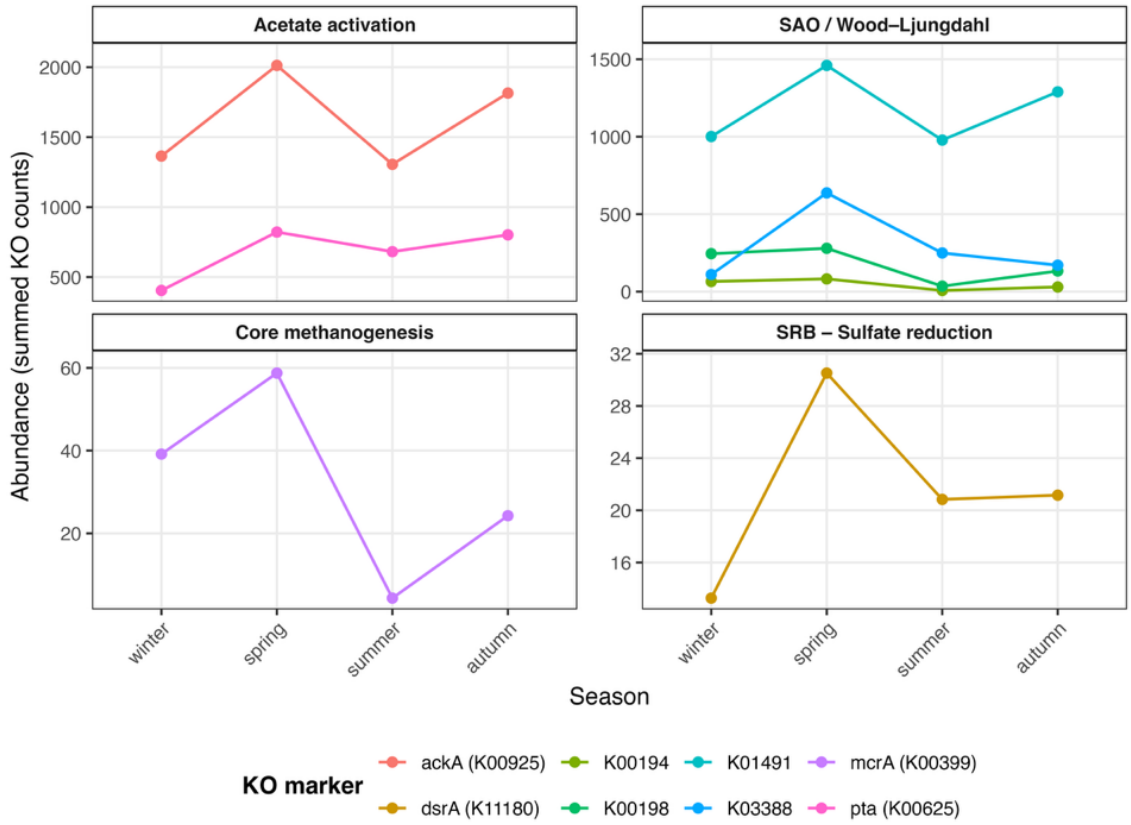

Figure D.3: Seasonal variation in marker gene abundances associated with acetate activation, Wood-Ljungdahl pathway, core methanogenesis, and sulphate reduction. Values represent cumulative KO counts per season across all ten septic tanks.

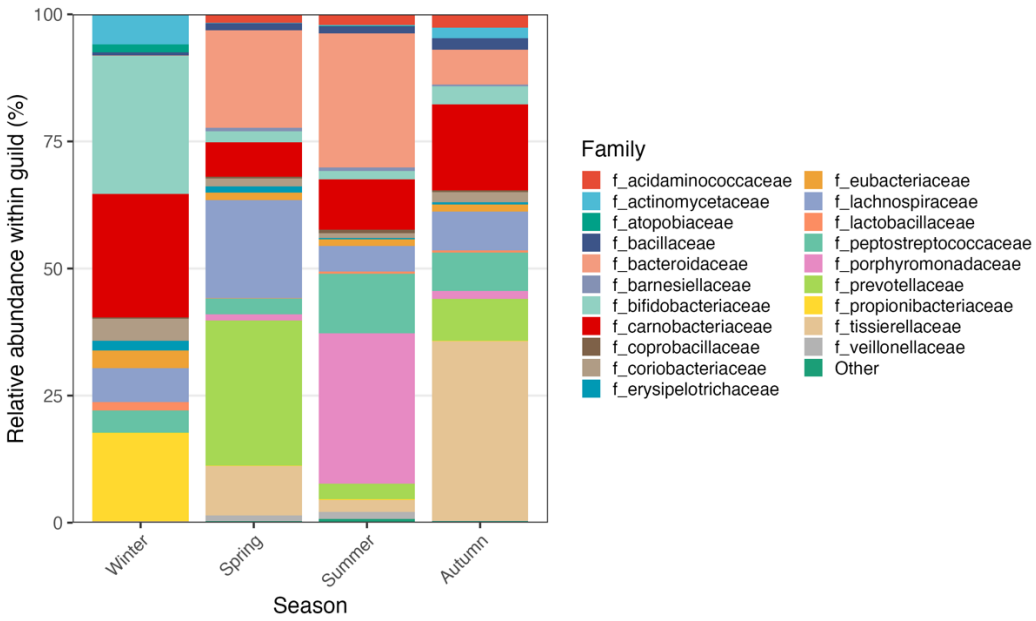

Figure D.4: Seasonal relative abundance (100%-stacked) of the top 20 fermentative bacterial families in household septic tank 2.

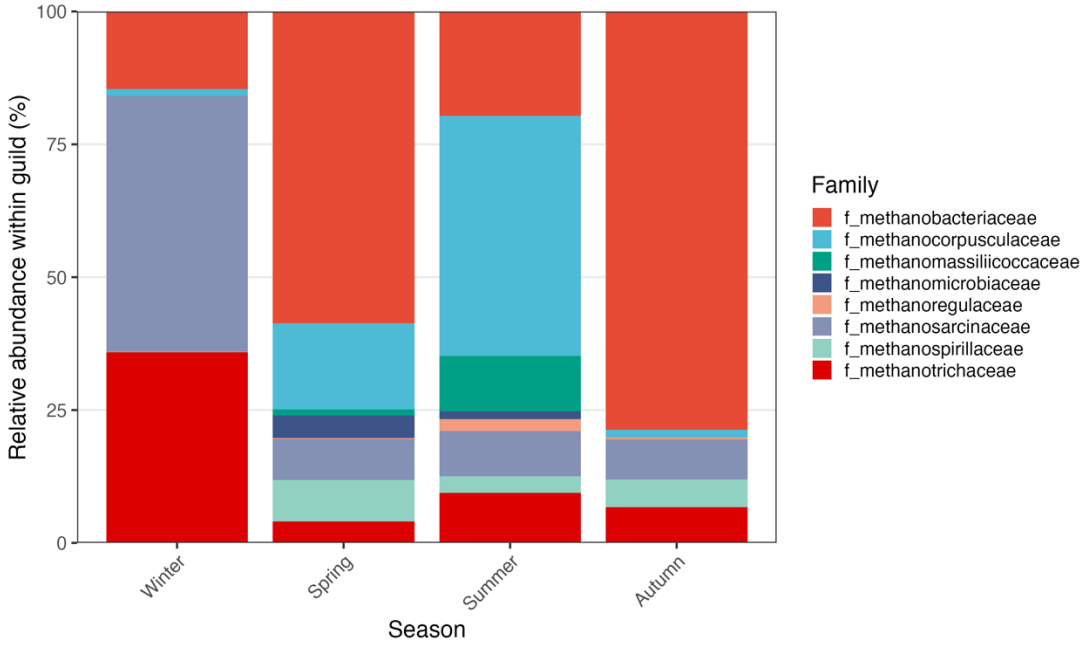

Figure D.5: Seasonal relative abundance (100%-stacked) of methanogenic families in household septic tank 2.

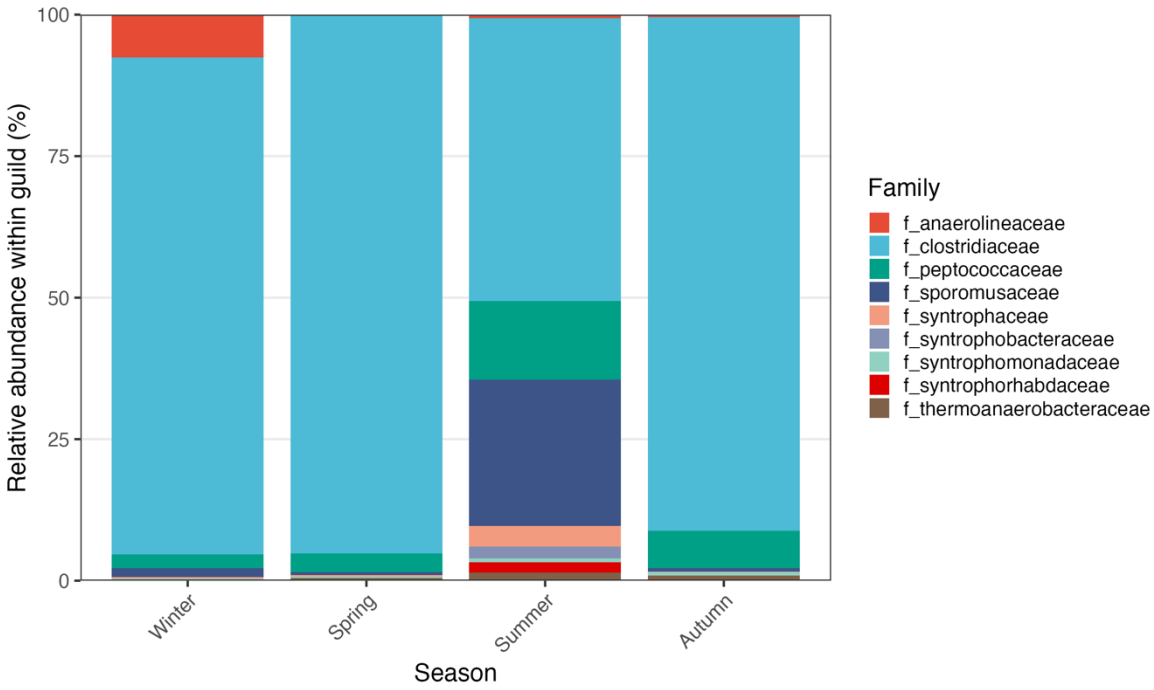

Figure D.6: Seasonal relative abundance (100%-stacked) of syntrophic acetate-oxidising bacterial families in household septic tank 2.

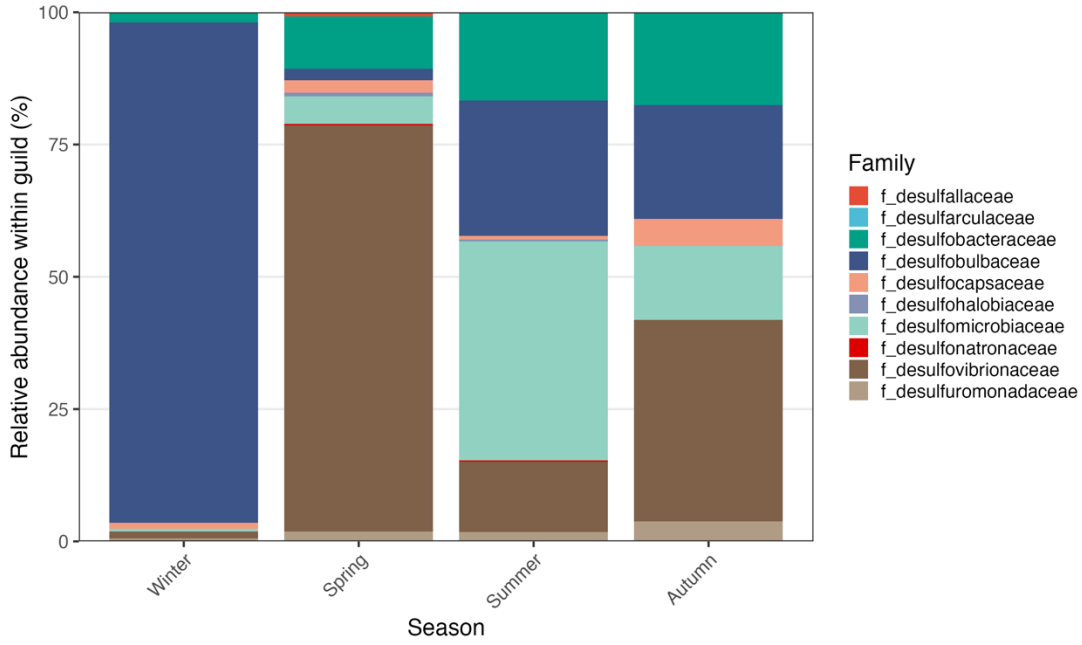

Figure D.7: Seasonal relative abundance (100%-stacked) of sulphate-reducing bacterial families in household septic tank 2.

E Custom R and Python scripts

In line with the Statement of originality, ChatGPT-5.2 (OpenAI, `https://chatgpt.com`) was used to assist with drafting, refining, and modifying R and Python scripts. All scripts presented in thesis were reviewed, executed, and finalised by the author.

E.1 Diversity and ordination (in R)

```
library(tidyverse)

library(vegan)

library(compositions)

library(ape)

library(patchwork)

library(reshape2)

HAS_ZCOMP <- requireNamespace("zCompositions", quietly = TRUE)

meta_file <- "/Users/au321/PhD/meta.csv"

prok_file <- "/Users/au321/PhD/prokaryote_tax.csv"

virus_file <- "/Users/au321/PhD/virus_abundance.csv"

out_dir <- "/Users/au321/PhD/figures_diversity_beta"

if (!dir.exists(out_dir)) dir.create(out_dir, recursive = TRUE)

sample_col <- "SampleID"

meta <- read.csv(meta_file, check.names = FALSE)

prok <- read.csv(prok_file, check.names = FALSE)

virus <- read.csv(virus_file, check.names = FALSE)

stopifnot(

sample_col %in% names(meta),
```

```
sample_col %in% names(prok),
sample_col %in% names(virus)
)
common_samples <- Reduce(
intersect,
list(meta[[sample_col]], prok[[sample_col]], virus[[sample_col]])
)
meta <- meta |> filter(.data[[sample_col]] %in% common_samples)
prok <- prok |> filter(.data[[sample_col]] %in% common_samples)
virus <- virus |> filter(.data[[sample_col]] %in% common_samples)
meta <- meta[match(common_samples, meta[[sample_col]]), ]
prok <- prok[match(common_samples, prok[[sample_col]]), ]
virus <- virus[match(common_samples, virus[[sample_col]]), ]
if ("Season" %in% names(meta)) {
meta$Season <- factor(meta$Season,
levels = c("Winter", "Spring", "Summer", "Autumn"))
}
if ("TankID" %in% names(meta)) {
meta$TankID <- factor(meta$TankID)
}
prok_mat <- prok |> select(-all_of(sample_col)) |> as.matrix()
virus_mat <- virus |> select(-all_of(sample_col)) |> as.matrix()
rownames(prok_mat) <- prok[[sample_col]]
rownames(virus_mat) <- virus[[sample_col]]
calc_alpha <- function(mat) {
```

```
S <- specnumber(mat)
H <- diversity(mat, index = "shannon")
invD <- diversity(mat, index = "invsimpson")
J <- ifelse(S > 0, H / log(S), NA_real_)
tibble(
!!sample_col := rownames(mat),
richness = S,
shannon = H,
invsimpson = invD,
evenness = J
)
}
alpha_prok <- calc_alpha(prok_mat) |> rename_with(~ paste0(.x, "_prok"),
-all_of(sample_col))
alpha_virus <- calc_alpha(virus_mat) |> rename_with(~ paste0(.x, "_virus"),
-all_of(sample_col))
alpha_all <- meta |>
select(all_of(sample_col), Season, TankID) |>
left_join(alpha_prok, by = sample_col) |>
left_join(alpha_virus, by = sample_col)
safe_row_prop <- function(mat) {
rs <- rowSums(mat, na.rm = TRUE)
rs[rs <= 0] <- 1
out <- sweep(mat, 1, rs, "/")
out[!is.finite(out)] <- 0
out
```

```
}
aitchison_dist <- function(mat) {
mat_rel <- safe_row_prop(mat)
if (HAS_ZCOMP) {
mat_nz <- zCompositions::cmultRepl(mat_rel, method = "CZM")
mat_nz <- sweep(mat_nz, 1, rowSums(mat_nz), "/")
} else {
eps <- 1e-6
mat_nz <- ifelse(mat_rel == 0, eps, mat_rel)
mat_nz <- sweep(mat_nz, 1, rowSums(mat_nz), "/")
}
clr_mat <- clr(mat_nz)
dist(clr_mat, method = "euclidean")
}
dist_prok_ait <- aitchison_dist(prok_mat)
dist_virus_ait <- aitchison_dist(virus_mat)
mantel_pv <- mantel(dist_prok_ait, dist_virus_ait,
method = "spearman", permutations = 999)
print(mantel_pv)
if (all(c("Season", "TankID") %in% names(meta))) {
perma_prok <- adonis2(dist_prok_ait ~ Season, data = meta,
permutations = 999, strata = meta$TankID)
perma_virus <- adonis2(dist_virus_ait ~ Season, data = meta,
permutations = 999, strata = meta$TankID)
cat("\nPERMANOVA (prok, Season) with TankID strata:\n")
```

```
print(perma_prok)
cat("\nPERMANOVA (virus, Season) with TankID strata:\n")
print(perma_virus)
cat("\nDispersion check (betadisper) — prok:\n")
bd_prok <- betadisper(dist_prok_ait, group = meta$Season)
print(permutest(bd_prok, permutations = 999))
cat("\nDispersion check (betadisper) — virus:\n")
bd_virus <- betadisper(dist_virus_ait, group = meta$Season)
print(permutest(bd_virus, permutations = 999))
}
cor_rich <- suppressWarnings(
cor.test(alpha_all$richness_prok, alpha_all$richness_virus,
method = "spearman", exact = FALSE)
)
p_a <- ggplot(alpha_all,
aes(x = richness_prok, y = richness_virus, colour = Season)) +
geom_point(size = 2.5, alpha = 0.9) +
geom_smooth(method = "lm", se = TRUE, colour = "black") +
theme_bw(base_size = 12) +
labs(
title = "a) Richness: viruses vs prokaryotes",
x = "Richness of prokaryotes",
y = "Richness of viruses"
) +
annotate("text",
```

```
x = Inf, y = -Inf,
hjust = 1.1, vjust = -0.5,
label = paste0("Spearman ρ = ",
round(cor_rich$estimate, 2),
", P = ",
signif(cor_rich$p.value, 2)),
size = 3)
ggsave(file.path(out_dir, "Fig_a_richness_prok_vs_virus.png"),
p_a, width = 5, height = 4, dpi = 300)
cor_shan <- suppressWarnings(
cor.test(alpha_all$shannon_prok, alpha_all$shannon_virus,
method = "spearman", exact = FALSE)
)
p_b <- ggplot(alpha_all,
aes(x = shannon_prok, y = shannon_virus, colour = Season)) +
geom_point(size = 2.5, alpha = 0.9) +
geom_smooth(method = "lm", se = TRUE, colour = "black") +
theme_bw(base_size = 12) +
labs(
title = "b) Shannon diversity: viruses vs prokaryotes",
x = "Shannon index of prokaryotes",
y = "Shannon index of viruses"
) +
annotate("text",
x = Inf, y = -Inf,
```

```
hjust = 1.1, vjust = -0.5,
label = paste0("Spearman ρ = ",
round(cor_shan$estimate, 2),
", P = ",
signif(cor_shan$p.value, 2)),
size = 3)
ggsave(file.path(out_dir, "Fig_b_shannon_prok_vs_virus.png"),
p_b, width = 5, height = 4, dpi = 300)
cor_invs <- suppressWarnings(
cor.test(alpha_all$invsimpson_prok, alpha_all$invsimpson_virus,
method = "spearman", exact = FALSE)
)
p_c <- ggplot(alpha_all,
aes(x = invsimpson_prok, y = invsimpson_virus, colour = Season)) +
geom_point(size = 2.5, alpha = 0.9) +
geom_smooth(method = "lm", se = TRUE, colour = "black") +
theme_bw(base_size = 12) +
labs(
title = "",
x = "Inverse Simpson of prokaryotes",
y = "Inverse Simpson of viruses"
) +
annotate("text",
x = Inf, y = -Inf,
hjust = 1.1, vjust = -0.5,
```

```
label = paste0("Spearman ρ = ",
round(cor_invs$estimate, 2),
", P = ",
signif(cor_invs$p.value, 2)),
size = 3)
ggsave(file.path(out_dir, "Fig_c_invsimpson_prok_vs_virus.png"),
p_c, width = 5, height = 4, dpi = 300)
beta_df <- tibble(
dist_prok_pair = as.numeric(dist_prok_ait),
dist_virus_pair = as.numeric(dist_virus_ait)
)
p_d <- ggplot(beta_df, aes(x = dist_prok_pair, y = dist_virus_pair)) +
geom_point(size = 1.5, alpha = 0.5) +
geom_smooth(method = "lm", se = TRUE, colour = "black") +
theme_bw(base_size = 12) +
labs(
title = "d) Pairwise Aitchison distances: viruses vs prokaryotes",
x = "Pairwise Aitchison distance (prokaryotes)",
y = "Pairwise Aitchison distance (viruses)"
) +
annotate("text",
x = Inf, y = -Inf,
hjust = 1.1, vjust = -0.5,
label = paste0("Mantel r = ",
round(mantel_pv$statistic, 2),
```

```
", P = ",
signif(mantel_pv$signif, 2)),
size = 3)
ggsave(file.path(out_dir, "Fig_d_pairwise_aitchison_prok_vs_virus.png"),
p_d, width = 5, height = 4, dpi = 300)
dist_df <- as.matrix(dist_prok_ait)
mean_upper <- function(d) {
v <- d[upper.tri(d)]
if (length(v) == 0) return(NA_real_)
mean(v, na.rm = TRUE)
}
mean_within_season <- function(season_name) {
samples <- meta$SampleID[meta$Season == season_name]
d <- dist_df[samples, samples, drop = FALSE]
mean_upper(d)
}
mean_between_seasons <- function(s1, s2) {
g1 <- meta$SampleID[meta$Season == s1]
g2 <- meta$SampleID[meta$Season == s2]
if (length(g1) == 0 || length(g2) == 0) return(NA_real_)
mean(dist_df[g1, g2, drop = FALSE], na.rm = TRUE)
}
seasons <- levels(meta$Season)
within_season <- sapply(seasons, mean_within_season)
print(within_season)
```

```
between_season <- outer(seasons, seasons, Vectorize(mean_between_seasons))
colnames(between_season) <- seasons
rownames(between_season) <- seasons
print(between_season)
mean_within_tank <- function(tank_name) {
samples <- meta$SampleID[meta$TankID == tank_name]
d <- dist_df[samples, samples, drop = FALSE]
mean_upper(d)
}
mean_between_tanks <- function(g1, g2) {
s1 <- meta$SampleID[meta$TankID == g1]
s2 <- meta$SampleID[meta$TankID == g2]
if (length(s1) == 0 || length(s2) == 0) return(NA_real_)
mean(dist_df[s1, s2, drop = FALSE], na.rm = TRUE)
}
tanks <- levels(meta$TankID)
within_tank <- sapply(tanks, mean_within_tank)
print(within_tank)
between_tank <- outer(tanks, tanks, Vectorize(mean_between_tanks))
colnames(between_tank) <- tanks
rownames(between_tank) <- tanks
print(between_tank)
dist_mat <- as.matrix(dist_prok_ait)
dist_long <- as_tibble(as.table(dist_mat), .name_repair = "minimal")
names(dist_long) <- c("Sample1", "Sample2", "distance")
```

```
dist_long <- dist_long |>
filter(Sample1 < Sample2) |>
left_join(meta |> select(Sample1 = SampleID, Season1 = Season, Tank1 = TankID),
by = "Sample1") |>
left_join(meta |> select(Sample2 = SampleID, Season2 = Season, Tank2 = TankID),
by = "Sample2") |>
mutate(
same_season = Season1 == Season2,
same_tank = Tank1 == Tank2
)
dist_long_season <- dist_long |>
mutate(season_group = ifelse(same_season,
as.character(Season1),
"Between seasons"))
p_beta_season <- ggplot(dist_long_season,
aes(x = season_group, y = distance)) +
geom_boxplot() +
theme_bw(base_size = 12) +
labs(
title = "β-diversity of prokaryotic communities within and between seasons",
x = "Season comparison",
y = "Aitchison distance"
)
ggsave(file.path(out_dir, "Fig_g_beta_within_between_seasons_prok.png"),
p_beta_season, width = 6, height = 4, dpi = 300)
```

```
dist_long_tank <- dist_long |>
mutate(tank_group = ifelse(same_tank,
as.character(Tank1),
"Between tanks"))
p_beta_tank <- ggplot(dist_long_tank,
aes(x = tank_group, y = distance)) +
geom_boxplot() +
theme_bw(base_size = 12) +
labs(
title = "β-diversity of prokaryotic communities within and between tanks",
x = "Tank comparison",
y = "Aitchison distance"
)
ggsave(file.path(out_dir, "Fig_h_beta_within_between_tanks_prok.png"),
p_beta_tank, width = 7, height = 4, dpi = 300)
pcoa_virus <- pcoa(dist_virus_ait)
pcoa_virus_df <- tibble(
!!sample_col := rownames(virus_mat),
Axis1 = pcoa_virus$vectors[, 1],
Axis2 = pcoa_virus$vectors[, 2]
) |>
left_join(meta[, c(sample_col, "Season")], by = sample_col)
var_v1 <- round(pcoa_virus$values$Relative_eig[1] * 100, 1)
var_v2 <- round(pcoa_virus$values$Relative_eig[2] * 100, 1)
p_e <- ggplot(pcoa_virus_df,
```

```
aes(x = Axis1, y = Axis2, colour = Season)) +
geom_point(size = 2.5, alpha = 0.9) +
theme_bw(base_size = 12) +
labs(
title = "e) PCoA of viral communities (Aitchison)",
x = paste0("PCoA1 (", var_v1, "%)"),
y = paste0("PCoA2 (", var_v2, "%)")
)
ggsave(file.path(out_dir, "Fig_e_pcoa_virus_aitchison.png"),
p_e, width = 5, height = 4, dpi = 300)
pcoa_prok <- pcoa(dist_prok_ait)
pcoa_prok_df <- tibble(
!!sample_col := rownames(prok_mat),
Axis1 = pcoa_prok$vectors[, 1],
Axis2 = pcoa_prok$vectors[, 2]
) |>
left_join(meta[, c(sample_col, "Season")], by = sample_col)
var_p1 <- round(pcoa_prok$values$Relative_eig[1] * 100, 1)
var_p2 <- round(pcoa_prok$values$Relative_eig[2] * 100, 1)
p_f <- ggplot(pcoa_prok_df,
aes(x = Axis1, y = Axis2, colour = Season)) +
geom_point(size = 2.5, alpha = 0.9) +
theme_bw(base_size = 12) +
labs(
title = "f) PCoA of prokaryotic communities (Aitchison)",
```

```
x = paste0("PCoA1 (", var_p1, "%)"),
y = paste0("PCoA2 (", var_p2, "%)")
)
ggsave(file.path(out_dir, "Fig_f_pcoa_prok_aitchison.png"),
p_f, width = 5, height = 4, dpi = 300)
set.seed(123)
nmds_prok <- metaMDS(dist_prok_ait, k = 2, trymax = 200, autotransform =
FALSE)
nmds_prok_df <- as.data.frame(nmds_prok$points)
nmds_prok_df[[sample_col]] <- rownames(nmds_prok_df)
nmds_prok_df <- nmds_prok_df |>
left_join(meta[, c(sample_col, "Season")], by = sample_col)
stress_prok <- round(nmds_prok$stress, 3)
p_nmds_prok <- ggplot(nmds_prok_df,
aes(x = MDS1, y = MDS2, colour = Season)) +
geom_point(size = 2.5, alpha = 0.9) +
theme_bw(base_size = 12) +
labs(
title = paste0("h) NMDS of prokaryotic communities (Aitchison, stress = ",
stress_prok, ")"),
x = "NMDS1",
y = "NMDS2"
)
ggsave(file.path(out_dir, "Fig_h_nmds_prok_aitchison.png"),
p_nmds_prok, width = 5, height = 4, dpi = 300)
nmds_virus <- metaMDS(dist_virus_ait, k = 2, trymax = 200, autotransform =
```

```
FALSE)
nmds_virus_df <- as.data.frame(nmds_virus$points)
nmds_virus_df[[sample_col]] <- rownames(nmds_virus_df)
nmds_virus_df <- nmds_virus_df |>
left_join(meta[, c(sample_col, "Season")], by = sample_col)
stress_virus <- round(nmds_virus$stress, 3)
p_nmds_virus <- ggplot(nmds_virus_df,
aes(x = MDS1, y = MDS2, colour = Season)) +
geom_point(size = 2.5, alpha = 0.9) +
theme_bw(base_size = 12) +
labs(
title = paste0("i) NMDS of viral communities (Aitchison, stress = ",
stress_virus, ")"),
x = "NMDS1",
y = "NMDS2"
)
ggsave(file.path(out_dir, "Fig_i_nmds_virus_aitchison.png"),
p_nmds_virus, width = 5, height = 4, dpi = 300)
```

E.2 Relative abundances of AD functional guilds (in R)

```
library(tidyverse)
base_dir <- "/Users/au321/PhD"
guild_file <- file.path(base_dir, "guild_map.csv")
comm_file <- file.path(base_dir, "community_wide.csv")
meta_file <- file.path(base_dir, "meta.csv")
outdir <- file.path(base_dir, "figures_results")
```

```
dir.create(outdir, showWarnings = FALSE, recursive = TRUE)
guild_map <- readr::read_csv(guild_file) # Taxon, Guild
comm_wide <- readr::read_csv(comm_file) # SampleID + taxa columns
meta <- readr::read_csv(meta_file) # SampleID, Season
comm_long <- comm_wide %>%
pivot_longer(
cols = -SampleID,
names_to = "Taxon",
values_to = "RelAbundance"
)
## Convert % to proportions
if (max(comm_long$RelAbundance, na.rm = TRUE) > 1.5) {
comm_long <- comm_long %>% mutate(RelAbundance = RelAbundance / 100)
}
comm_guild <- comm_long %>%
left_join(guild_map, by = "Taxon") %>%
mutate(
Guild = ifelse(is.na(Guild), "Others", Guild),
Guild = factor(
Guild,
levels = c("Fermenters", "SAOB", "SRB", "Methanogens", "Others")
)
)
sample_guild <- comm_guild %>%
group_by(SampleID, Guild) %>%
```

```
summarise(
guild_rel = sum(RelAbundance, na.rm = TRUE),
.groups = "drop"
)
sample_guild_season <- sample_guild %>%
left_join(meta, by = "SampleID") %>%
filter(!is.na(Season)) %>%
mutate(Season = factor(Season, levels = c("Winter","Spring","Summer","Autumn")))
season_guild <- sample_guild_season %>%
group_by(Season, Guild) %>%
summarise(mean_rel = mean(guild_rel), .groups = "drop")
print(table(season_guild$Guild))
season_guild_4 <- season_guild %>%
filter(Guild != "Others")
guild_cols_full <- c(
"Fermenters" = "#D55E00",
"SAOB" = "#E69F00",
"SRB" = "#0072B2",
"Methanogens" = "#009E73",
"Others" = "grey70"
)
guild_cols_4 <- guild_cols_full[names(guild_cols_full) != "Others"]
p4 <- ggplot(season_guild_4, aes(x = Season, y = mean_rel, fill = Guild)) +
geom_col(position = "fill", color = "white", linewidth = 0.2) +
scale_y_continuous(labels = scales::percent) +
```

```
scale_fill_manual(values = guild_cols_4) +
labs(
x = "Season",
y = "Relative abundance of AD guilds (%)",
fill = "Guild",
title = ""
) +
theme_bw(base_size = 13) +
theme(
panel.grid.major.x = element_blank(),
plot.title = element_text(face = "bold", hjust = 0.5)
)
ggsave(
file.path(outdir, "05_guilds_stackbar_4guilds.png"),
p4, width = 5, height = 4, dpi = 300
)
p_full <- ggplot(season_guild, aes(x = Season, y = mean_rel, fill = Guild)) +
geom_col(position = "fill", color = "white", linewidth = 0.2) +
scale_y_continuous(labels = scales::percent) +
scale_fill_manual(values = guild_cols_full) +
labs(
x = "Season",
y = "Relative abundance of AD guilds (%)",
fill = "Guild",
title = ""
```

```
) +
theme_bw(base_size = 13) +
theme(
panel.grid.major.x = element_blank(),
plot.title = element_text(face = "bold", hjust = 0.5)
)
ggsave(
file.path(outdir, "05_guilds_stackbar_with_others.png"),
p_full, width = 5, height = 4, dpi = 300
)
```

E.3 Spearman correlation (in R)

```
library(tidyverse)
library(ComplexHeatmap)
library(circlize)
library(grid)
commfile <- "/Users/au321/PhD/community_wide.csv"
metafile <- "/Users/au321/PhD/meta.csv"
outdir <- "/Users/au321/PhD/figures_corr_heatmaps_FINAL"
dir.create(outdir, showWarnings = FALSE, recursive = TRUE)
comm_tax <- readr::read_csv(commfile)
meta <- readr::read_csv(metafile)
comm_all <- comm_tax %>%
dplyr::left_join(meta, by = "SampleID")
if ("Season" %in% names(comm_all)) {
comm_all <- comm_all %>%
```

```
dplyr::mutate(season = Season)
}
# Methanogens
Meth_taxa <- c(
"f_methanobacteriaceae",
"f_methanocorpusculaceae",
"f_methanomassiliicoccaceae",
"f_methanomicrobiaceae",
"f_methanoregulaceae",
"f_methanosarcinaceae",
"f_methanospirillaceae",
"f_methanotrichaceae"
)
# SRBs
SRB_taxa <- c(
"f_desulfallaceae",
"f_desulfarculaceae",
"f_desulfobacteraceae",
"f_desulfobulbaceae",
"f_desulfocapsaceae",
"f_desulfohalobiaceae",
"f_desulfomicrobiaceae",
"f_desulfonatronaceae",
"f_desulfovibrionaceae",
"f_desulfuromonadaceae"
```

```
)
# SAOBs
SAOB_taxa <- c(
"f_anaerolineaceae",
"f_clostridiaceae",
"f_peptococcaceae",
"f_sporomusaceae",
"f_syntrophaceae",
"f_syntrophobacteraceae",
"f_syntrophomonadaceae",
"f_syntrophorhabdaceae",
"f_thermoanaerobacteraceae"
)
# Fermenters (top 10)
Ferm_taxa <- c(
"f_propionibacteriaceae",
"f_prevotellaceae",
"f_tissierellaceae",
"f_coriobacteriaceae",
"f_lachnospiraceae",
"f_carnobacteriaceae",
"f_porphyromonadaceae",
"f_bacteroidaceae",
"f_bifidobacteriaceae",
"f_barnesiellaceae"
```

```
)
Prok_taxa_all <- c(Meth_taxa, SRB_taxa, SAOB_taxa, Ferm_taxa)
Prok_taxa_all <- intersect(Prok_taxa_all, names(comm_all))
prok_guild <- c(
setNames(rep("Methanogens", length(Meth_taxa)), Meth_taxa),
setNames(rep("SRB", length(SRB_taxa)), SRB_taxa),
setNames(rep("SAOB", length(SAOB_taxa)), SAOB_taxa),
setNames(rep("Fermenters", length(Ferm_taxa)), Ferm_taxa)
)
prok_guild <- prok_guild[Prok_taxa_all]
virus_cols <- grep("viridae|phage|virus", names(comm_all),
ignore.case = TRUE, value = TRUE)
virus_cols <- setdiff(virus_cols,
c("f_mimiviridae"))
virus_cols <- intersect(virus_cols, names(comm_all))
candidate_abiotic <- c("COD","BOD5","TS","VS","pH","temp_ww")
Abiotic_vars <- intersect(candidate_abiotic, names(comm_all))
if (length(Abiotic_vars) == 0) {
stop("Abiotic_vars is empty – check abiotic column names in meta.csv.")
}
compute_corr <- function(df, rows, cols, method = "spearman") {
rows <- intersect(rows, names(df))
cols <- intersect(cols, names(df))
if (length(rows) == 0 || length(cols) == 0) {
stop("No overlapping variables between rows/cols and the dataframe.")
```

```
}
df_num <- df %>% dplyr::select(where(is.numeric))
rows <- intersect(rows, names(df_num))
cols <- intersect(cols, names(df_num))
rho_mat <- matrix(NA_real_, nrow = length(rows), ncol = length(cols),
dimnames = list(rows, cols))
p_mat <- matrix(NA_real_, nrow = length(rows), ncol = length(cols),
dimnames = list(rows, cols))
for (i in seq_along(rows)) {
for (j in seq_along(cols)) {
x <- df_num[[rows[i]]]
y <- df_num[[cols[j]]]
ok <- is.finite(x) & is.finite(y)
if (sum(ok) > 2 &&
length(unique(x[ok])) > 1 &&
length(unique(y[ok])) > 1) {
ct <- suppressWarnings(cor.test(x[ok], y[ok], method = method))
rho_mat[i, j] <- unname(ct$estimate)
p_mat[i, j] <- ct$p.value
}
}
}
p_vec <- as.vector(p_mat)
padj_vec <- stats::p.adjust(p_vec, method = "BH")
padj_mat <- matrix(padj_vec,
```

```
nrow = nrow(p_mat),
ncol = ncol(p_mat),
dimnames = dimnames(p_mat))
list(rho = rho_mat, padj = padj_mat)
}
p_to_stars <- function(p) {
ifelse(is.na(p), "",
ifelse(p < 0.001, "***",
ifelse(p < 0.01, "**",
ifelse(p < 0.05, "*", ""))))
}
col_fun <- circlize::colorRamp2(
c(-1, -0.5, 0, 0.5, 1),
c("#053061", "#2166ac", "#f7f7f7", "#b2182b", "#67001f")
)
make_heatmap_object <- function(rho_mat, padj_mat,
main_title = "",
prok_group_vector = NULL,
show_guild_annotation = FALSE) {
star_mat <- matrix(p_to_stars(padj_mat),
nrow = nrow(padj_mat),
ncol = ncol(padj_mat),
dimnames = dimnames(padj_mat))
group_factor <- NULL
top_anno <- NULL
```

```
if (show_guild_annotation && !is.null(prok_group_vector)) {
cols <- colnames(rho_mat)
guild <- prok_group_vector[cols]
group_factor <- factor(
guild,
levels = c("SRB", "SAOB", "Methanogens", "Fermenters")
)
guild_colors <- c(
"SRB" = "#1f78b4",
"SAOB" = "#33a02c",
"Methanogens" = "#fb9a99",
"Fermenters" = "grey70"
)
present_lvls <- levels(group_factor)[levels(group_factor) %in% unique(group_factor)]
label_map <- c(
"SRB" = "SRB",
"SAOB" = "SAOB",
"Methanogens" = "Methanogens",
"Fermenters" = "Fermenters"
)
block_labels <- label_map[present_lvls]
top_anno <- HeatmapAnnotation(
block = anno_block(
labels = block_labels,
labels_gp = gpar(fontsize = 10, fontface = "bold")
```

```
),
Guild = group_factor,
col = list(Guild = guild_colors),
which = "column",
simple_anno_size = unit(3, "mm"),
show_legend = FALSE
)
}
Heatmap(
rho_mat,
name = "Spearman (rho)",
col = col_fun,
cluster_rows = FALSE,
cluster_columns = FALSE,
show_row_dend = FALSE,
show_column_dend = FALSE,
top_annotation = top_anno,
column_split = group_factor,
cluster_column_slices = FALSE,
column_title = main_title,
column_title_gp = gpar(fontsize = 10, fontface = "bold"),
row_names_gp = gpar(fontsize = 6),
column_names_gp = gpar(fontsize = 6, rot = 90),
column_gap = unit(2, "mm"),
heatmap_legend_param = list(
```

```
direction = "horizontal",
labels_gp = gpar(fontsize = 7),
title_gp = gpar(fontsize = 8),
legend_width = unit(4, "cm")
),
cell_fun = function(j, i, x, y, width, height, fill) {
lab <- star_mat[i, j]
if (lab != "") {
grid.text(
lab, x = x, y = y,
gp = gpar(col = "white",
fontsize = 7,
fontface = "bold")
)
}
}
)
}
season_levels <- c("Winter", "Spring", "Summer", "Autumn")
present_seasons <- intersect(season_levels, unique(comm_all$season))
for (s in present_seasons) {
message("Phage vs prok – season: ", s)
df_s <- comm_all %>% dplyr::filter(season == s)
corr_phage_prok <- compute_corr(df_s,
rows = virus_cols,
```

```
cols = Prok_taxa_all)
ht <- make_heatmap_object(
rho_mat = corr_phage_prok$rho,
padj_mat = corr_phage_prok$padj,
main_title = paste("Phages vs prokaryotic guilds –", s),
prok_group_vector = prok_guild,
show_guild_annotation = TRUE
)
pdf(file.path(outdir,
paste0("heatmap_phage_vs_prok_", s, ".pdf")),
width = 8, height = 5.2)
draw(
ht,
merge_legend = TRUE,
heatmap_legend_side = "bottom"
)
dev.off()
}
ht_list_prok_env <- list()
for (s in present_seasons) {
message("Prok vs abiotic – season: ", s)
df_s <- comm_all %>% dplyr::filter(season == s)
corr_prok_env <- compute_corr(df_s,
rows = Prok_taxa_all,
cols = Abiotic_vars)
```

```
ht_list_prok_env[[s]] <- make_heatmap_object(
rho_mat = corr_prok_env$rho,
padj_mat = corr_prok_env$padj,
main_title = s,
prok_group_vector = NULL,
show_guild_annotation = FALSE
)
}
if (length(ht_list_prok_env) > 0) {
ht_prok_env_combined <- Reduce('+', ht_list_prok_env[present_seasons])
pdf(file.path(outdir,
"heatmap_prok_vs_abiotic_allSeasons.pdf"),
width = 11, height = 5.5)
draw(
ht_prok_env_combined,
merge_legend = TRUE,
heatmap_legend_side = "bottom"
)
dev.off()
}
ht_list_virus_env <- list()
for (s in present_seasons) {
message("Phage vs abiotic – season: ", s)
df_s <- comm_all %>% dplyr::filter(season == s)
corr_virus_env <- compute_corr(df_s,
```

```
rows = virus_cols,
cols = Abiotic_vars)
ht_list_virus_env[[s]] <- make_heatmap_object(
rho_mat = corr_virus_env$rho,
padj_mat = corr_virus_env$padj,
main_title = s,
prok_group_vector = NULL,
show_guild_annotation = FALSE
)
}
if (length(ht_list_virus_env) > 0) {
ht_virus_env_combined <- Reduce('+', ht_list_virus_env[present_seasons])
pdf(file.path(outdir,
"heatmap_phage_vs_abiotic_allSeasons.pdf"),
width = 11, height = 5.5)
draw(
ht_virus_env_combined,
merge_legend = TRUE,
heatmap_legend_side = "bottom"
)
dev.off()
}
message("All heatmaps generated.")
```

E.4 Seasonal dynamics of microbial guilds and viruses (in R)

```
library(tidyverse)

meta_file <- "/Users/au321/PhD/meta.csv"

prok_file <- "/Users/au321/PhD/prokaryote_tax.csv"

guild_file <- "/Users/au321/PhD/guild_map.csv"

virus_file <- "/Users/au321/PhD/virus_abundance.csv"

out_dir <- "/Users/au321/PhD/figures_virus_microbe"

if (!dir.exists(out_dir)) dir.create(out_dir, recursive = TRUE)

sample_col <- "SampleID"

meta <- read.csv(meta_file, check.names = FALSE)

prok <- read.csv(prok_file, check.names = FALSE)

guild <- read.csv(guild_file, check.names = FALSE)

nature_palette <- c(

"#E64B35","#4DBBD5","#00A087","#3C5488","#F39B7F","#8491B4",

"#91D1C2","#DC0000","#7E6148","#B09C85","#0099B4","#EEA236",

"#8DA0CB","#FC8D62","#66C2A5","#E78AC3","#A6D854","#FFD92F",

"#E5C494","#B3B3B3","#1B9E77","#D95F02","#7570B3","#E7298A"

)

stopifnot(

sample_col %in% names(meta),

sample_col %in% names(prok),

"Taxon" %in% names(guild),

"Guild" %in% names(guild)

)

prok_taxa_cols <- setdiff(names(prok), sample_col)
```

```
prok_mat <- prok |>
select(all_of(prok_taxa_cols)) |>
as.matrix()
prok_rel <- sweep(prok_mat, 1, rowSums(prok_mat, na.rm = TRUE), "/")
prok_rel[is.na(prok_rel)] <- 0
prok_rel_df <- as.data.frame(prok_rel)
prok_rel_df[[sample_col]] <- prok[[sample_col]]
prok_rel_df <- prok_rel_df |>
left_join(meta[, c(sample_col, "Season")], by = sample_col) |>
filter(!is.na(Season))
prok_rel_df$Season <- factor(
prok_rel_df$Season,
levels = c("Winter", "Spring", "Summer", "Autumn")
)
rel_long <- prok_rel_df |>
pivot_longer(
cols = all_of(prok_taxa_cols),
names_to = "Taxon",
values_to = "Abundance_rel"
) |>
left_join(guild, by = "Taxon") |>
filter(!is.na(Guild))
plot_guild_100 <- function(df_all, guild_name) {
df_g <- df_all |>
filter(Guild == guild_name, !is.na(Season))
```

```
if (nrow(df_g) == 0) {
message("No taxa for guild: ", guild_name)
return(invisible(NULL))
}
seasonal <- df_g |>
group_by(Season, Taxon) |>
summarise(
mean_rel = mean(Abundance_rel, na.rm = TRUE),
.groups = "drop"
) |>
mutate(
mean_rel = ifelse(is.na(mean_rel) | is.nan(mean_rel), 0, mean_rel)
)
guild_totals <- seasonal |>
group_by(Season) |>
summarise(
guild_total = sum(mean_rel, na.rm = TRUE),
.groups = "drop"
)
seasonal_guild <- seasonal |>
left_join(guild_totals, by = "Season") |>
mutate(
guild_prop = ifelse(guild_total > 0, mean_rel / guild_total, 0)
)
top_taxa <- seasonal_guild |>
```

```
group_by(Taxon) |>
summarise(total_prop = sum(guild_prop), .groups = "drop") |>
arrange(desc(total_prop)) |>
slice_head(n = 20) |>
pull(Taxon)
seasonal_norm <- seasonal_guild |>
mutate(
Taxon_plot = ifelse(Taxon %in% top_taxa, Taxon, "Other")
) |>
group_by(Season, Taxon_plot) |>
summarise(
prop_raw = sum(guild_prop, na.rm = TRUE),
.groups = "drop"
) |>
group_by(Season) |>
mutate(
season_sum = sum(prop_raw, na.rm = TRUE),
prop = ifelse(season_sum > 0, prop_raw / season_sum, 0)
) |>
ungroup()
check <- seasonal_norm |>
group_by(Season) |>
summarise(sum_prop = sum(prop), .groups = "drop")
print(paste("Guild:", guild_name))
print(check)
```

```
p <- ggplot(seasonal_norm,
aes(x = Season,
y = prop * 100,
fill = Taxon_plot)) +
geom_col(width = 0.85) + scale_fill_manual(values = nature_palette) +
scale_y_continuous(
limits = c(0, 100),
expand = c(0, 0),
oob = scales::squish
) +
labs(
title = paste("Top 20", guild_name, "families by season (100% stacked)"),
x = "Season",
y = "Relative abundance within guild (%)",
fill = "Family"
) +
theme_bw(base_size = 12) +
theme(
axis.text.x = element_text(angle = 45, hjust = 1),
legend.position = "right",
legend.key.size = unit(0.4, "cm"),
panel.grid.minor = element_blank()
)
outfile <- file.path(
out_dir,
```

```
paste0("top20_", tolower(gsub(" ", "_", guild_name)), "_by_season_100pct.png")
)
ggsave(outfile, p, width = 8, height = 5, dpi = 300)
message("Saved: ", outfile)
}
guilds <- c("Methanogens", "Fermenters", "SRB", "SAOB")
for (g in guilds) {
plot_guild_100(rel_long, g)
}
virus <- read.csv(virus_file, check.names = FALSE)
stopifnot(sample_col %in% names(virus))
virus_taxa_cols <- setdiff(names(virus), sample_col)
virus_mat <- virus |>
select(all_of(virus_taxa_cols)) |>
as.matrix()
virus_rel <- sweep(virus_mat, 1, rowSums(virus_mat, na.rm = TRUE), "/")
virus_rel[is.na(virus_rel)] <- 0
virus_rel_df <- as.data.frame(virus_rel)
virus_rel_df[[sample_col]] <- virus[[sample_col]]
virus_rel_df <- virus_rel_df |>
left_join(meta[, c(sample_col, "Season")], by = sample_col) |>
filter(!is.na(Season))
virus_rel_df$Season <- factor(
virus_rel_df$Season,
levels = c("Winter", "Spring", "Summer", "Autumn")
```

```
)
virus_long <- virus_rel_df |>
pivot_longer(
cols = all_of(virus_taxa_cols),
names_to = "Taxon",
values_to = "Abundance_rel"
)
virus_seasonal <- virus_long |>
group_by(Season, Taxon) |>
summarise(
mean_rel = mean(Abundance_rel, na.rm = TRUE),
.groups = "drop"
) |>
mutate(
mean_rel = ifelse(is.na(mean_rel) | is.nan(mean_rel), 0, mean_rel)
)
virus_totals <- virus_seasonal |>
group_by(Season) |>
summarise(
virus_total = sum(mean_rel, na.rm = TRUE),
.groups = "drop"
)
virus_seasonal2 <- virus_seasonal |>
left_join(virus_totals, by = "Season") |>
mutate(
```

```
virus_prop = ifelse(virus_total > 0, mean_rel / virus_total, 0)
)
virus_top20 <- virus_seasonal2 |>
group_by(Taxon) |>
summarise(total_prop = sum(virus_prop), .groups = "drop") |>
arrange(desc(total_prop)) |>
slice_head(n = 20) |>
pull(Taxon)
virus_norm <- virus_seasonal2 |>
mutate(
Taxon_plot = ifelse(Taxon %in% virus_top20, Taxon, "Other")
) |>
group_by(Season, Taxon_plot) |>
summarise(
prop_raw = sum(virus_prop, na.rm = TRUE),
.groups = "drop"
) |>
group_by(Season) |>
mutate(
season_sum = sum(prop_raw, na.rm = TRUE),
prop = ifelse(season_sum > 0, prop_raw / season_sum, 0)
) |>
ungroup()
print("Viruses:")
print(virus_norm |>
```

```
group_by(Season) |>
summarise(sum_prop = sum(prop), .groups = "drop"))
p_virus <- ggplot(virus_norm,
aes(x = Season,
y = prop * 100,
fill = Taxon_plot)) +
geom_col(width = 0.85) + scale_fill_manual(values = nature_palette) +
scale_y_continuous(
limits = c(0, 100),
expand = c(0, 0),
oob = scales::squish
) +
labs(
title = "Top 20 viral families by season (100% stacked)",
x = "Season",
y = "Relative abundance within viral community (%)",
fill = "Viral family"
) +
theme_bw(base_size = 12) +
theme(
axis.text.x = element_text(angle = 45, hjust = 1),
legend.position = "right",
legend.key.size = unit(0.4, "cm"),
panel.grid.minor = element_blank()
)
```

```
virus_outfile <- file.path(out_dir, "top20_viruses_by_season_100pct.png")

ggsave(virus_outfile, p_virus, width = 8, height = 5, dpi = 300)

message("Saved: ", virus_outfile)
```

E.5 Anaerobic pathways statistics (in R)

```
#heatmap

library(tidyverse)

library(ComplexHeatmap)

library(circlize)

base <- "/Users/au321/PhD"

file_6 <- file.path(base,
"Unigene_Analysis/6.Unigene_KEGG_TPM_AD_markers.csv")

file_45 <- file.path(base,
"Unigene_Analysis/45.Unigene_KEGG_TPM_AD_markers.csv")

file_58 <- file.path(base,
"Unigene_Analysis/58.Unigene_KEGG_TPM_AD_markers.csv")

meta_file <- file.path(base, "meta.csv")

outdir <- file.path(base, "figures_results")

dir.create(outdir, showWarnings = FALSE, recursive = TRUE)

meta <- readr::read_csv(meta_file) %>%

dplyr::select(HaploxID, Season) %>%

dplyr::mutate(

HaploxID = as.character(HaploxID),

Season = factor(Season, levels = c("Winter", "Spring", "Summer", "Autumn"))

)

season_lookup <- meta %>%
```

```
dplyr::filter(!is.na(HaploxID))

df6 <- readr::read_csv(file_6)

df45 <- readr::read_csv(file_45)

df58 <- readr::read_csv(file_58)

to_long <- function(df) {

# numeric columns are assumed to be sample TPM columns

sample_cols <- names(df)[sapply(df, is.numeric)]

df_long <- df %>%

dplyr::select(KO, dplyr::all_of(sample_cols)) %>%

tidyr::pivot_longer(

cols = -KO,

names_to = "HaploxID",

values_to = "TPM"

) %>%

dplyr::mutate(

HaploxID = as.character(HaploxID),

KO = as.character(KO)

)

df_long %>%

dplyr::semi_join(season_lookup, by = "HaploxID")

}

ko_long <- list(df6, df45, df58) %>%

purrr::map_dfr(to_long)

ko_long <- ko_long %>%

dplyr::left_join(season_lookup, by = "HaploxID") %>%
```

```
dplyr::filter(!is.na(Season))

target_kos <- tribble(

~KO, ~GeneLabel, ~Pathway,

"K00925", "ackA (K00925)", "AA", # Acetate activation

"K00625", "pta (K00625)", "AA",

"K00194", "K00194", "SAO/WL", # Wood–Ljungdahl / SAO

"K00198", "K00198", "SAO/WL",

"K01491", "K01491", "SAO/WL",

"K03388", "K03388", "SAO/WL",

"K11180", "dsrA (K11180)", "SR", # Sulphate reduction

"K11181", "dsrB (K11181)", "SR",

"K00399", "mcrA (K00399)", "Core-M" # Core methanogenesis

)

ko_long <- ko_long %>%

dplyr::filter(KO %in% target_kos$KO)

ko_season <- ko_long %>%

dplyr::group_by(KO, Season) %>%

dplyr::summarise(

mean_tpm = mean(TPM, na.rm = TRUE),

.groups = "drop"

)

ko_mat_df <- ko_season %>%

tidyr::pivot_wider(

names_from = Season,

values_from = mean_tpm
```

```
) %>%
dplyr::arrange(match(KO, target_kos$KO))
ko_ids <- ko_mat_df$KO
ko_mat <- ko_mat_df %>%
dplyr::select(-KO) %>%
as.matrix()
# ensure column order: Winter, Spring, Summer, Autumn (if present)
wanted_cols <- c("Winter", "Spring", "Summer", "Autumn")
present_cols <- intersect(wanted_cols, colnames(ko_mat))
ko_mat <- ko_mat[, present_cols, drop = FALSE]
ko_log <- log10(ko_mat + 1)
scale_rows <- function(x) {
t(scale(t(x)))
}
ko_z <- scale_rows(ko_log)
ko_z[is.na(ko_z)] <- 0
row_labels <- target_kos$GeneLabel[match(ko_ids, target_kos$KO)]
row_groups <- target_kos$Pathway[match(ko_ids, target_kos$KO)]
rownames(ko_z) <- row_labels
col_fun <- circlize::colorRamp2(
c(-1, -0.5, 0, 0.5, 1),
c("#053061", "#2166ac", "#f7f7f7", "#b2182b", "#67001f")
)
ht <- Heatmap(
ko_z,
```

```
name = "Z-score\n(log10 TPM)",
col = col_fun,
cluster_rows = FALSE,
cluster_columns = FALSE,
row_split = factor(
row_groups,
levels = c("AA", "SAO/WL", "SR", "Core-M")
),
row_title_gp = grid::gpar(fontsize = 9, fontface = "bold"),
row_title_rot = 0,
row_names_gp = grid::gpar(fontsize = 9),
column_names_gp = grid::gpar(fontsize = 10, fontface = "bold"),
heatmap_legend_param = list(
title = "Z-score",
at = c(-1, -0.5, 0, 0.5, 1)
)
)
pdf(file.path(outdir, "Fig_AD_KO_heatmap.pdf"), width = 6, height = 4)
draw(ht)
dev.off()
png(file.path(outdir, "Fig_AD_KO_heatmap.png"),
width = 2000, height = 1600, res = 300)
draw(ht)
dev.off()
#plot for Appendix D
```

```
library(tidyverse)
infile <- "/Users/au321/PhD/KO_abundance_by_season.csv"
outdir <- "/Users/au321/PhD/figures_ADpathways"
outfile <- file.path(outdir, "Fig5xx_KO_abundance_by_season_faceted_line.png")
dir.create(outdir, showWarnings = FALSE, recursive = TRUE)
ko <- readr::read_csv(infile)
ko <- ko %>%
mutate(
season = factor(
season,
levels = c("winter", "spring", "summer", "autumn")
),
pathway_block = factor(
pathway_block,
levels = c("acetate_activation",
"SAO_WoodLjungdahl",
"hydrogenotrophic",
"core_methanogenesis"),
labels = c("Acetate activation",
"SAO / Wood–Ljungdahl",
"Hydrogenotrophic methanogenesis",
"Core methanogenesis")
)
)
ko <- ko %>%
```

```
mutate(
KO_label = dplyr::case_when(
KO == "K00399" ~ "mcrA (K00399)",
KO == "K00200" ~ "K00200",
KO == "K00194" ~ "K00194",
KO == "K00198" ~ "K00198",
KO == "K01491" ~ "K01491",
KO == "K03388" ~ "K03388",
KO == "K00625" ~ "pta (K00625)",
KO == "K00925" ~ "ackA (K00925)",
TRUE ~ KO
)
)
p <- ggplot(ko,
aes(x = season,
y = abundance_sum,
group = KO_label,
colour = KO_label)) +
geom_line(linewidth = 0.7) +
geom_point(size = 2) +
facet_wrap(~ pathway_block, scales = "free_y") +
scale_y_continuous(expand = expansion(mult = c(0.05, 0.1))) +
labs(
x = "Season",
y = "Abundance (summed KO counts)",
```

```
colour = "KO marker"
) +
theme_bw(base_size = 12) +
theme(
panel.grid.minor = element_blank(),
axis.text.x = element_text(angle = 45, hjust = 1),
strip.background = element_rect(fill = "white", colour = "black"),
strip.text = element_text(face = "bold"),
legend.position = "bottom",
legend.title = element_text(face = "bold")
)
ggsave(filename = outfile,
plot = p,
width = 8,
height = 5,
dpi = 600)
```

E.6 Phage–host association analyses (in Python): Virsorter2

```
#!/usr/bin/env bash
set -euo pipefail
# Uses: run_virsorter_one.sh (single-sample runner)
# Assumes: conda env vs2_x86 is already active
SCRIPT="/Users/au321/PhageHostPipeline/run_virsorter_one.sh"
CONTIG_DIR="/Users/au321/PhageHostPipeline/contigs_all"
PROPHAGE_DIR="/Users/au321/PhageHostPipeline/prophage_x86"
# Samples to process
```

```
SAMPLES=(
"Sample_G1"
"Sample_G2"
"Sample_G3"
"Sample_G4"
"Sample_G5"
"Sample_G6"
"Sample_G7"
"Sample_G8"
"Sample_G10"
"Sample_G11"
"Sample_G12"
"Sample_G13"
"Sample_G14"
"Sample_G15"
"Sample_G16"
"Sample_G17"
"Sample_G19"
"Sample_G20"
"Sample_G22"
"Sample_G25"
"Sample_G26"
"Sample_G28"
"Sample_G29"
"Sample_G31"
```

"Sample_G34"

"Sample_G37"

"Sample_G40"

"Sample_G43"

"Sample_G44"

"Sample_G46"

"Sample_G49"

"Sample_G52"

"Sample_G53"

"Sample_G55"

"Sample_G57"

"A1"

"A2"

"A3"

"A4"

"A5"

"A7"

"A8"

"A9"

"A10"

"A11"

"A12"

"A13"

"A14"

"S2A5"

```
"S2A10"
"S4A8"
"S4A10"
"S9A2"
"S9A51"
)
echo "== VirSorter2 batch run for samples: ${SAMPLES[*]} =="
echo
for s in "${SAMPLES[@]}"; do
contigs="${CONTIG_DIR}/${s}.contigs.fa"
outdir="${PROPHAGE_DIR}/${s}"
echo "----------------------------"
echo "Sample: $s"
echo "Contigs: $contigs"
echo "Workdir: $outdir"
if [[ ! -f "$contigs" ]]; then
echo " WARNING: contigs file not found for $s: $contigs"
echo " Skipping."
echo
continue
fi
# Skip if VirSorter2 has already finished for this sample
if [[ -f "${outdir}/final-viral-combined.fa" ]]; then
echo " final-viral-combined.fa already exists for $s"
echo " -> Skipping (already done)."
```

```
echo

continue

fi

echo " Running VirSorter2..."

"${SCRIPT}" "$s" "$contigs"

echo " Finished $s."

echo

done

echo "== VirSorter2 batch run complete =="
```

E.7 Phage–host association analyses (in Python): CheckV

```
micromamba activate checkv_env

DB="/Users/au321/checkv-db-full/checkv-db-v1.5"

INPUT_DIR="/Users/au321/PhageHostPipeline/prophage_x86"

OUT_DIR="/Users/au321/PhageHostPipeline/checkv"

mkdir -p "$OUT_DIR"

for dir in "$INPUT_DIR"/*; do

NAME=$(basename "$dir")

IN_FASTA="$dir/final-viral-combined.fa"

OUT_SUB="$OUT_DIR/$NAME"

if [[ ! -f "$IN_FASTA" ]]; then

echo “Missing $IN_FASTA, skipping $NAME"

continue

fi

echo "----------------------------"

echo "Processing sample: $NAME"
```

```
echo "----------------------------"

rm -rf "$OUT_SUB"

mkdir -p "$OUT_SUB"

checkv end_to_end "$IN_FASTA" "$OUT_SUB" -d "$DB" -t 1

done
```

E.8 Phage–host association analyses (in Python): iPHoP

iPHoP was executed on a Slurm-managed HPC system; the job submission script is provided below for reproducibility (Roux, 2022):

```
#!/bin/bash

#SBATCH --job-name=iphop_run

#SBATCH --cpus-per-task=8

#SBATCH --mem=64G

#SBATCH --time=48:00:00

#SBATCH --output=iphop_all.out

#SBATCH --error=iphop_all.err

source ~/.bashrc

conda activate iphop_env

IN_BASE="/Volumes/ExtraCopy/checkv"

IPHOP_DB="/Volumes/ExtraCopy/Jun_2025_pub_rw"

OUT_BASE="/Volumes/ExtraCopy/iphop_out"

mkdir -p "$OUT_BASE"

for f in "$IN_BASE"/*/viruses.fna; do

s=$(basename "$(dirname "$f")")

OUT_DIR="$OUT_BASE/$s"

mkdir -p "$OUT_DIR"
```

```
iphop predict "$f" \
--db_dir "$IPHOP_DB" \
--out_dir "$OUT_DIR" \
--threads "${SLURM_CPUS_PER_TASK}"
done
```

E.9 Phage–Host Analysis (in R)

```
library(tidyverse)
iphop_dir <- "/Users/au321/PhageHostPipeline/Tank2"
files <- list(
Winter = file.path(iphop_dir, "Tank2_Winter_Host_prediction_to_genus_m90.csv"),
Spring = file.path(iphop_dir, "Tank2_Spring_Host_prediction_to_genus_m90.csv"),
Summer = file.path(iphop_dir,
"Tank2_Summer_Host_prediction_to_genus_m90.csv"),
Autumn = file.path(iphop_dir,
"Tank2_Autumn_Host_prediction_to_genus_m90.csv")
)
seasons_all <- c("Winter", "Spring", "Summer", "Autumn")
stopifnot(dir.exists(iphop_dir))
stopifnot(all(file.exists(unlist(files, use.names = FALSE))))
Meth_taxa <- c(
"f_methanobacteriaceae",
"f_methanocorpusculaceae",
"f_methanomassiliicoccaceae",
"f_methanomicrobiaceae",
"f_methanoregulaceae",
```

```
"f_methanosarcinaceae",
"f_methanospirillaceae",
"f_methanotrichaceae"
)
SRB_taxa <- c(
"f_desulfallaceae",
"f_desulfarculaceae",
"f_desulfobacteraceae",
"f_desulfobulbaceae",
"f_desulfocapsaceae",
"f_desulfohalobiaceae",
"f_desulfomicrobiaceae",
"f_desulfonatronaceae",
"f_desulfovibrionaceae",
"f_desulfuromonadaceae"
)
SAOB_taxa <- c(
"f_anaerolineaceae",
"f_clostridiaceae",
"f_peptococcaceae",
"f_sporomusaceae",
"f_syntrophaceae",
"f_syntrophobacteraceae",
"f_syntrophomonadaceae",
"f_syntrophorhabdaceae",
```

```
"f_thermoanaerobacteraceae"
)
Ferm_taxa <- c(
"f_propionibacteriaceae",
"f_prevotellaceae",
"f_tissierellaceae",
"f_coriobacteriaceae",
"f_lachnospiraceae",
"f_carnobacteriaceae",
"f_porphyromonadaceae",
"f_bacteroidaceae",
"f_bifidobacteriaceae",
"f_barnesiellaceae"
)
guild_levels <- c("Fermenters", "SAOB", "SRB", "Methanogens")
guild_lookup <- bind_rows(
tibble(Host_family = tolower(Meth_taxa), Guild = "Methanogens"),
tibble(Host_family = tolower(SRB_taxa), Guild = "SRB"),
tibble(Host_family = tolower(SAOB_taxa), Guild = "SAOB"),
tibble(Host_family = tolower(Ferm_taxa), Guild = "Fermenters")
) %>%
distinct(Host_family, .keep_all = TRUE) %>%
mutate(Guild = factor(Guild, levels = guild_levels))
iphop_all <- purrr::map_df(names(files), function(season) {
readr::read_csv(files[[season]], show_col_types = FALSE) %>%
```

```
mutate(Season = season)
})
required_cols <- c("Virus", "Host genus")
miss <- setdiff(required_cols, names(iphop_all))
if (length(miss) > 0) stop("Missing required column(s): ", paste(miss, collapse = ", "))
conf_col <- if ("Confidence score" %in% names(iphop_all)) {
"Confidence score"
} else if ("Confidence" %in% names(iphop_all)) {
"Confidence"
} else {
stop("Could not find confidence column. Expected 'Confidence score' or 'Confidence'.")
}
iphop_all <- iphop_all %>%
mutate(
Host_family = stringr::str_extract(‘Host genus‘, "f__[^;]+"),
Host_family = stringr::str_replace(Host_family, "f__", "f_"),
Host_family = stringr::str_to_lower(Host_family),
Season = factor(Season, levels = seasons_all)
)
iphop_guild <- iphop_all %>%
left_join(guild_lookup, by = "Host_family") %>%
filter(!is.na(Guild)) %>%
mutate(Guild = factor(Guild, levels = guild_levels))
out_dir <- file.path(iphop_dir, "thesis_figures_colours_matched")
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)
```

```
pal_meth <- c(
"f_methanotrichaceae" = "#E04830",
"f_methanosarcinaceae" = "#8090B0",
"f_methanospirillaceae" = "#90D0C0",
"f_methanoregulaceae" = "#F09880",
"f_methanomicrobiaceae" = "#385080",
"f_methanomassiliicoccaceae" = "#009880",
"f_methanocorpusculaceae" = "#48B8D0",
"f_methanobacteriaceae" = "#D94A3A"
)
pal_saob <- c(
"f_clostridiaceae" = "#48B8D0",
"f_sporomusaceae" = "#385080",
"f_peptococcaceae" = "#009880",
"f_anaerolineaceae" = "#E04830",
"f_syntrophaceae" = "#F09880",
"f_syntrophobacteraceae" = "#8090B0",
"f_syntrophomonadaceae" = "#90D0C0",
"f_syntrophorhabdaceae" = "#B0402C",
"f_thermoanaerobacteraceae" = "#786048"
)
pal_srb <- c(
"f_desulfobulbaceae" = "#385080",
"f_desulfomicrobiaceae" = "#90D0C0",
"f_desulfovibrionaceae" = "#786048",
```

```
"f_desulfobacteraceae" = "#009880",
"f_desulfuromonadaceae" = "#C8B090",
"f_desulfohalobiaceae" = "#8090B0",
"f_desulfocapsaceae" = "#F09880",
"f_desulfonatronaceae" = "#E04830",
"f_desulfallaceae" = "#D94A3A",
"f_desulfarculaceae" = "#48B8D0"
)
pal_ferm_key <- c(
"f_bacteroidaceae" = "#F09880",
"f_lachnospiraceae" = "#8090B0",
"f_prevotellaceae" = "#B6D64A",
"f_bifidobacteriaceae" = "#90D0C0",
"f_porphyromonadaceae" = "#D790C8",
"f_propionibacteriaceae" = "#F2D14B",
"f_tissierellaceae" = "#C8B090",
"f_carnobacteriaceae" = "#E00000",
"f_coriobacteriaceae" = "#786048",
"Other" = "#009880"
)
plot_guild_100pct <- function(df, guild_name, palette, out_file, seasons_all, top_n =
NULL) {
palette <- palette[!duplicated(names(palette))]
base <- df %>%
filter(Guild == guild_name, !is.na(Host_family)) %>%
mutate(Season = factor(Season, levels = seasons_all)) %>%
```

```
count(Season, Host_family, name = "n")

if (!is.null(top_n)) {

top_fams <- base %>%

group_by(Host_family) %>%

summarise(Total = sum(n), .groups = "drop") %>%

slice_max(Total, n = top_n, with_ties = FALSE) %>%

pull(Host_family)

d <- df %>%

filter(Guild == guild_name, !is.na(Host_family)) %>%

mutate(

Season = factor(Season, levels = seasons_all),

Host_family = if_else(Host_family %in% top_fams, Host_family, "Other")

) %>%

count(Season, Host_family, name = "n")

} else {

d <- base

}

families_all <- unique(c(names(palette), as.character(unique(d$Host_family))))

d <- d %>%

mutate(

Season = factor(Season, levels = seasons_all),

Host_family = as.character(Host_family)

) %>%

tidyr::complete(

Season = factor(seasons_all, levels = seasons_all),
```

```
Host_family = families_all,
fill = list(n = 0)
)
# If a season has total = 0, Percent becomes NA -> set to 0
d <- d %>%
group_by(Season) %>%
mutate(
total = sum(n),
Percent = if_else(total > 0, n / total * 100, 0)
) %>%
ungroup()
missing_levels <- setdiff(unique(d$Host_family), names(palette))
if (length(missing_levels) > 0) {
extra <- rep("#BDBDBD", length(missing_levels))
names(extra) <- missing_levels
palette <- c(palette, extra)
}
d <- d %>%
mutate(Host_family = factor(Host_family, levels = names(palette)))
p <- ggplot(d, aes(x = Season, y = Percent, fill = Host_family)) +
geom_col(width = 0.8) +
scale_fill_manual(values = palette, drop = FALSE) +
scale_x_discrete(drop = FALSE) +
labs(
x = "Season",
```

```
y = "Relative abundance within guild (%)",
fill = "Family"
) +
theme_classic(base_size = 12) +
theme(
axis.text.x = element_text(angle = 45, hjust = 1),
legend.position = "right"
)
ggsave(out_file, p, width = 8.5, height = 4.8, dpi = 300)
return(p)
}
p_ferm <- plot_guild_100pct(
iphop_guild, "Fermenters", pal_ferm_key,
file.path(out_dir, "Fig_iPHoP_100pct_Fermenters.png"),
seasons_all = seasons_all, top_n = 10
)
p_meth <- plot_guild_100pct(
iphop_guild, "Methanogens", pal_meth,
file.path(out_dir, "Fig_iPHoP_100pct_Methanogens.png"),
seasons_all = seasons_all
)
p_saob <- plot_guild_100pct(
iphop_guild, "SAOB", pal_saob,
file.path(out_dir, "Fig_iPHoP_100pct_SAOB.png"),
seasons_all = seasons_all
```

```
)
p_srb <- plot_guild_100pct(
iphop_guild, "SRB", pal_srb,
file.path(out_dir, "Fig_iPHoP_100pct_SRB.png"),
seasons_all = seasons_all
)
readr::write_csv(guild_lookup %>% arrange(Guild, Host_family),
file.path(out_dir, "guild_lookup_family_to_guild.csv"))
season_family <- iphop_guild %>%
count(Guild, Season, Host_family, name = "n") %>%
arrange(Guild, Season, desc(n))
readr::write_csv(season_family, file.path(out_dir, "season_by_family_counts.csv"))
cat("Saved 100% stacked guild plots + tables to:\n", out_dir, "\n")
multi_host_phages <- iphop_all %>%
filter(!is.na(`Host genus`)) %>%
distinct(Virus, Season, `Host genus`, .keep_all = TRUE) %>%
group_by(Virus) %>%
mutate(n_hosts = n_distinct(`Host genus`)) %>%
ungroup() %>%
filter(n_hosts > 1) %>%
arrange(desc(n_hosts), Virus)
out_dir <- file.path(iphop_dir, "multi_host_results")
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)
readr::write_csv(
multi_host_phages,
```

```
file.path(out_dir, "multi_host_phage_predictions.csv")
)
mh_dir <- file.path(iphop_dir, "multi_host_results")
dir.create(mh_dir, showWarnings = FALSE, recursive = TRUE)
multi_host_phages <- multi_host_phages %>%
mutate(
Host_family = stringr::str_extract(‘Host genus‘, "f___[^;]+"),
Host_family = stringr::str_replace(Host_family, "f___", "f_"),
Host_family = stringr::str_to_lower(Host_family)
)
mh_by_virus <- multi_host_phages %>%
distinct(Virus, n_hosts)
p_mh_hist <- ggplot(mh_by_virus, aes(x = n_hosts)) +
geom_bar(width = 0.75) +
scale_x_continuous(breaks = sort(unique(mh_by_virus$n_hosts))) +
labs(
x = "Number of predicted host genera per viral contig",
y = "Number of viral contigs",
title = "Distribution of predicted host multiplicity (iPHoP)"
) +
theme_classic(base_size = 12)
ggsave(
filename = file.path(mh_dir, "Fig_multi_host_hist_n_hosts.png"),
plot = p_mh_hist, width = 7.5, height = 4.5, dpi = 300
)
```

```
# Bin n_hosts for readability
mh_bins <- mh_by_virus %>%
mutate(
host_bin = dplyr::case_when(
n_hosts <= 3 ~ "2–3 hosts",
n_hosts <= 5 ~ "4–5 hosts",
TRUE ~ "≥6 hosts"
)
)
mh_family <- multi_host_phages %>%
distinct(Virus, Host_family, n_hosts) %>%
left_join(mh_bins, by = c("Virus", "n_hosts")) %>%
filter(!is.na(Host_family)) %>%
count(Host_family, host_bin, name = "n_contigs") %>%
group_by(Host_family) %>%
mutate(total_family = sum(n_contigs)) %>%
ungroup() %>%
arrange(desc(total_family), Host_family)
topN <- 15
top_fams <- mh_family %>%
distinct(Host_family, total_family) %>%
slice_max(total_family, n = topN, with_ties = FALSE) %>%
pull(Host_family)
mh_family_plot <- mh_family %>%
mutate(
```

```
Host_family_plot = if_else(Host_family %in% top_fams, Host_family,
"other_families")
) %>%
group_by(Host_family_plot, host_bin) %>%
summarise(n_contigs = sum(n_contigs), .groups = "drop") %>%
group_by(Host_family_plot) %>%
mutate(total_family = sum(n_contigs)) %>%
ungroup()
p_mh_family <- ggplot(
mh_family_plot,
aes(
x = reorder(Host_family_plot, total_family),
y = n_contigs,
fill = host_bin
)
) +
geom_col() +
coord_flip() +
labs(
x = "Host family",
y = "Number of multi-host viral contigs",
fill = "Host multiplicity",
title = "Multi-host viral contigs by host family (binned host multiplicity)"
) +
theme_classic(base_size = 12)
ggsave(
```

```
filename = file.path(mh_dir, "Fig_multi_host_by_family_binned.png"),
plot = p_mh_family, width = 8.5, height = 6.0, dpi = 300
)
mh_matrix <- multi_host_phages %>%
filter(Host_family %in% top_fams) %>%
distinct(Virus, Host_family) %>%
count(Host_family, Virus, name = "n_links") %>%
group_by(Host_family) %>%
mutate(total_family = sum(n_links)) %>%
ungroup()
top_viruses <- multi_host_phages %>%
distinct(Virus, n_hosts) %>%
slice_max(n_hosts, n = 30, with_ties = FALSE) %>%
pull(Virus)
mh_matrix_plot <- mh_matrix %>%
filter(Virus %in% top_viruses) %>%
mutate(
Host_family = factor(Host_family, levels = mh_matrix %>%
distinct(Host_family, total_family) %>%
arrange(desc(total_family)) %>% pull(Host_family)),
Virus = factor(Virus, levels = top_viruses)
)
p_mh_heat <- ggplot(mh_matrix_plot, aes(x = Virus, y = Host_family, fill =
n_links)) +
geom_tile() +
labs(
```

```
x = "Viral contig (top 30 by n_hosts)",
y = "Host family (top families)",
fill = "n links",
title = "Multi-host associations (virus Œ host family)"
) +
theme_classic(base_size = 11) +
theme(
axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1)
)
ggsave(
filename = file.path(mh_dir, "Fig_multi_host_heatmap_virus_by_family.png"),
plot = p_mh_heat, width = 10.5, height = 5.5, dpi = 300
)
readr::write_csv(mh_by_virus, file.path(mh_dir,
"multi_host_by_virus_n_hosts.csv"))
readr::write_csv(mh_family, file.path(mh_dir,
"multi_host_by_family_binned_counts_full.csv"))
readr::write_csv(mh_family_plot, file.path(mh_dir,
"multi_host_by_family_binned_counts_topN.csv"))
cat("Saved multi-host figures + tables to:\n", mh_dir, "\n")
```

F Technical Guidance for the Design and Operation of Septic Systems in Cold Climates

F.1 Purpose and Rationale

This appendix provides a concise set of technical recommendations for the design, construction, and operation of septic tank systems in cold-climate regions. The guidance is informed by the findings of this study, which demonstrate that many system failures are attributable to design deficiencies and non-compliance with construction standards rather than user behaviour. In particular, inadequate consideration of thermal conditions, hydraulic performance, and solids management has been shown to compromise microbial processes and overall treatment efficiency.

The objective of this guidance is to translate research findings into practical recommendations for engineers, builders, inspectors, and regulators, thereby supporting improved system reliability, environmental protection, and public health outcomes.

F.2 Key Design Requirements

F.2.1 Tank Sizing and Hydraulic Retention Time

1. Septic tanks must be sized to ensure adequate hydraulic retention time (HRT), which is critical for solids settling and anaerobic digestion.
2. Recommended minimum HRT under cold conditions: 48–72 hours
3. Tank volume should be based on:
    - Number of users
    - Daily wastewater generation

- Anticipated sludge accumulation rates

Undersized tanks reduce effective treatment by limiting solids separation and microbial degradation processes.

F.2.2 Installation Depth and Frost Protection

1. To maintain stable operating temperatures, septic tanks should be installed below the local frost penetration depth.
2. Tanks should be placed below frost line (typically 1.5–2.5 m in cold regions, site-specific)
3. Shallow installations should be avoided due to high risk of freezing and performance failure
4. Where deep installation is not feasible, additional insulation must be provided.

F.2.3 Thermal Insulation

1. Thermal protection is essential to sustain microbial activity during cold periods.
2. Recommended measures:
    - Insulation of tank covers and access points
    - Use of insulating materials (e.g. expanded polystyrene, felt) above the tank
    - Minimisation of exposed surfaces
    - Maintaining wastewater temperatures above approximately 4 °C to 8 °C significantly improves treatment performance.

F.2.4 Internal Configuration and Solids Management

- Effective solids separation is required for stable system operation.
- Install inlet and outlet baffles or tees to prevent short-circuiting
- Consider multi-chamber designs to enhance settling and digestion

- Ensure sufficient sludge storage capacity
- Poor internal configuration can lead to solids carryover and reduced treatment efficiency.

F.2.5 Effluent Discharge and Infiltration Systems

- Effluent should be discharged to properly designed subsurface infiltration systems.

Key considerations:

- Soil permeability and infiltration capacity
- Even distribution of effluent (e.g. perforated pipes, distribution boxes)
- Protection against freezing of pipes and drainage layers
- In cold climates, infiltration systems should be designed to maintain functionality during winter conditions and prevent clogging.

F.3 Operational Recommendations

F.3.1 Preventive Desludging

- Desludging should be conducted as a preventive measure rather than in response to system failure.
- Recommended timing: late summer or early autumn
- Avoid desludging during winter, when system access and biological recovery are limited
- Regular desludging prevents excessive sludge accumulation and maintains effective tank volume.

F.3.2 Monitoring and Maintenance

- Routine inspection is necessary to ensure continued system performance.

Recommended practices:

- Periodic measurement of sludge depth
- Inspection of inlet and outlet structures
- Identification of signs of malfunction (e.g. odour, overflow, slow drainage)

F.3.3 User Practices

Although design deficiencies are a primary cause of failure, user behaviour remains important.

Users should:

- Avoid disposal of non-biodegradable materials
- Limit excessive water loading
- Be informed about maintenance requirements

F.4 Environmental Protection Considerations

F.4.1 Pathogen Risk

- Low temperatures may reduce pathogen attenuation within septic systems.
- Ensure adequate soil treatment zones
- Avoid direct discharge to surface waters
- Consider additional treatment where risk is high

F.4.2 Nitrogen and Groundwater Protection

- Septic tanks primarily convert nitrogen to ammonium, with limited removal occurring within the tank.
- Cold conditions may further reduce nitrogen transformation
- Effluent may contribute to groundwater contamination

System design should consider:

- Distance from groundwater sources
- Soil filtration capacity
- Potential for nutrient leaching

F.5 Implications for Regulation and Implementation

The findings of this study indicate the need for stronger enforcement of design standards and improved technical literacy among practitioners.

Recommended actions include:

- Development of region-specific design guidelines for cold climates
- Establishment of minimum construction standards
- Introduction of scheduled desludging requirements
- Certification or inspection systems for installation quality
- Improved alignment between engineering practice, regulatory frameworks, and environmental conditions is essential to ensure reliable performance of septic systems in cold regions.

G Pathogen-associated families identified in this study

Table G.1: Pathogen-associated bacterial families identified in this study and their reported clinical and environmental characteristics.

Family	Representative genera/species	Associated diseases	Typical environments/sources	Temperature preference	Reference
Aeromonadaceae	*Aeromonas* spp.	Gastroenteritis; wound infections	Aquatic environments; wastewater	Mesophilic; psychrotolerant	Bitton (2005); Cabral (2010)
Arcobacteraceae	*Arcobacter* spp.	Diarrhoeal disease	Wastewater	Mesophilic; psychrotolerant	Bitton (2005)
Bacteroidaceae	*Bacteroides* spp.	Opportunistic infections	Human gut	Mesophilic	Bitton (2005); Mara and Horan (2008); Cabral (2010)
Brucellaceae	*Brucella* spp.	Brucellosis	Livestock; animal waste	Mesophilic	Bitton (2005)
Campylobacteraceae	*Campylobacter* spp.	Gastroenteritis; Guillain–Barré syndrome	Animal faeces; wastewater	Mesophilic	Mara and Horan (2008); Bitton (2005)
Clostridiaceae	*Clostridium botulinum*; *C. tetani*	Tetanus; botulism; diarrhoeal disease	Soil; faeces; anaerobic environments	Psychrophilic to thermophilic	Bitton (2005); Mara and Horan (2008); Cabral (2010)
Corynebacteriaceae	*Corynebacterium* spp.	Diphtheria; opportunistic infections	Skin; mucosa; wastewater	Mesophilic	Mara and Horan (2008)
Enterobacteriaceae	*Escherichia coli*; *Salmonella typhi*; *Shigella sonnei*; *Klebsiella pneumoniae*	Gastrointestinal infections; typhoid fever; bacillary dysentery	Faecal matter; wastewater	Mesophilic; some taxa, such as *Yersinia*, can grow at 4 °C	Bitton (2005); Mara and Horan (2008); Cabral (2010)
Enterococcaceae	*Enterococcus* spp.	Urinary tract infections	Human gut; wastewater	Mesophilic	Mara and Horan (2008); Cabral (2010)
Fusobacteriaceae	*Fusobacterium prausnitzii*	Periodontal infections	Human gut	Mesophilic	Mara and Horan (2008); Cabral (2010)
Helicobacteraceae	*Helicobacter* spp.	Gastric infections; stomach cancer	Human stomach; wastewater; biofilms	Mesophilic	Bitton (2005); Mara and Horan (2008); Cabral (2010)
Legionellaceae	*Legionella pneumophila*	Legionnaires' disease; Pontiac fever	Water systems; biofilms	Mesophilic to thermophilic	Mara and Horan (2008); Bitton (2005)

Continued on next page

Table G.1 continued from previous page

Family	Representative genera/species	Associated diseases	Typical environments/sources	Temperature preference	Reference
Moraxellaceae	*Acinetobacter calcoaceticus*; *Moraxella osloensis*	Respiratory infections; opportunistic infections	Soil; water; wastewater	Psychrotolerant	Mara and Horan (2008); Bitton (2005)
Mycobacteriaceae	*Mycobacterium avium*; *M. tuberculosis*	Tuberculosis; Crohn's disease	Soil; water; biofilms	Mesophilic	Bitton (2005); Mara and Horan (2008); Cabral (2010)
Neisseriaceae	*Neisseria gonorrhoeae*; *N. meningitidis*	Gonorrhoea; meningitis	Human mucosa	Mesophilic	Thomson et al. (2008)
Pasteurellaceae	*Pasteurella* spp.; *Haemophilus* spp.	Respiratory infections	Animal microbiota; human mucosa; wastewater	Mesophilic	Mara and Horan (2008)
Peptostreptococcaceae	*Peptostreptococcus productus*	Anaerobic infections	Human gut; anaerobic environments	Mesophilic	Mara and Horan (2008); Cabral (2010)
Pseudomonadaceae	*Pseudomonas aeruginosa*	Opportunistic infections	Soil; water; wastewater	Mesophilic	Bitton (2005); Mara and Horan (2008); Cabral (2010)
Rickettsiaceae	*Rickettsia* spp.	Typhus fever	Arthropod vectors	Mesophilic	Cabral (2010)
Staphylococcaceae	*Staphylococcus aureus*	Skin infections	Skin; wastewater	Mesophilic	Mara and Horan (2008); Bitton (2005)
Streptococcaceae	*Streptococcus bovis*	Respiratory infections; meningitis	Human mucosa; wastewater	Mesophilic	Mara and Horan (2008); Bitton (2005)
Treponemataceae	*Treponema* spp.	Syphilis; oral infections	Human and animal microbiota	Mesophilic	Mara and Horan (2008)
Vibrionaceae	*Vibrio cholerae*	Cholera; gastroenteritis	Aquatic environments	Mesophilic	Bitton (2005); Mara and Horan (2008); Cabral (2010)